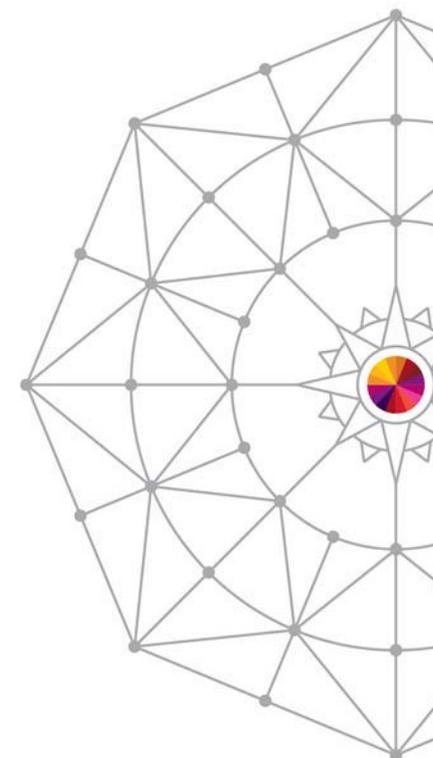


Understanding z/VM 6.3 Through New Performance Toolkit Reports

Version 1.2 – July 10, 2014

Bill Bitner
z/VM Development Client Focus and Care
bitnerb@us.ibm.com



SHARE August 2014
Session 15741

#SHAREorg



SHARE is an independent volunteer-run information technology association
that provides **education, professional networking and industry influence.**

Copyright (c) 2014 by SHARE Inc.  Except where otherwise noted, this work is licensed under
<http://creativecommons.org/licenses/by-nc-sa/3.0/>

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

IBM*	System z10*	System z196
IBM Logo*	Tivoli*	System z114
DB2*	z10 BC	System zEC12
Dynamic Infrastructure*	z9*	System zBC12
GDPS*	z/OS*	
HiperSockets	z/VM*	
Parallel Sysplex*	z/VSE	
RACF*	zEnterprise*	
System z*		

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

OpenSolaris, Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
INFINIBAND, InfiniBand Trade Association and the INFINIBAND design marks are trademarks and/or service marks of the INFINIBAND Trade Association.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

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.

Notice Regarding Specialty Engines (e.g., zIIPs, zAAPs and IFLs):

Any information contained in this document regarding Specialty Engines ("SEs") and SE eligible workloads provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs). IBM authorizes customers to use IBM SE only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the "Authorized Use Table for IBM Machines" provided at www.ibm.com/systems/support/machine_warranties/machine_code/aut.html ("AUT").

No other workload processing is authorized for execution on an SE.

IBM offers SEs at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

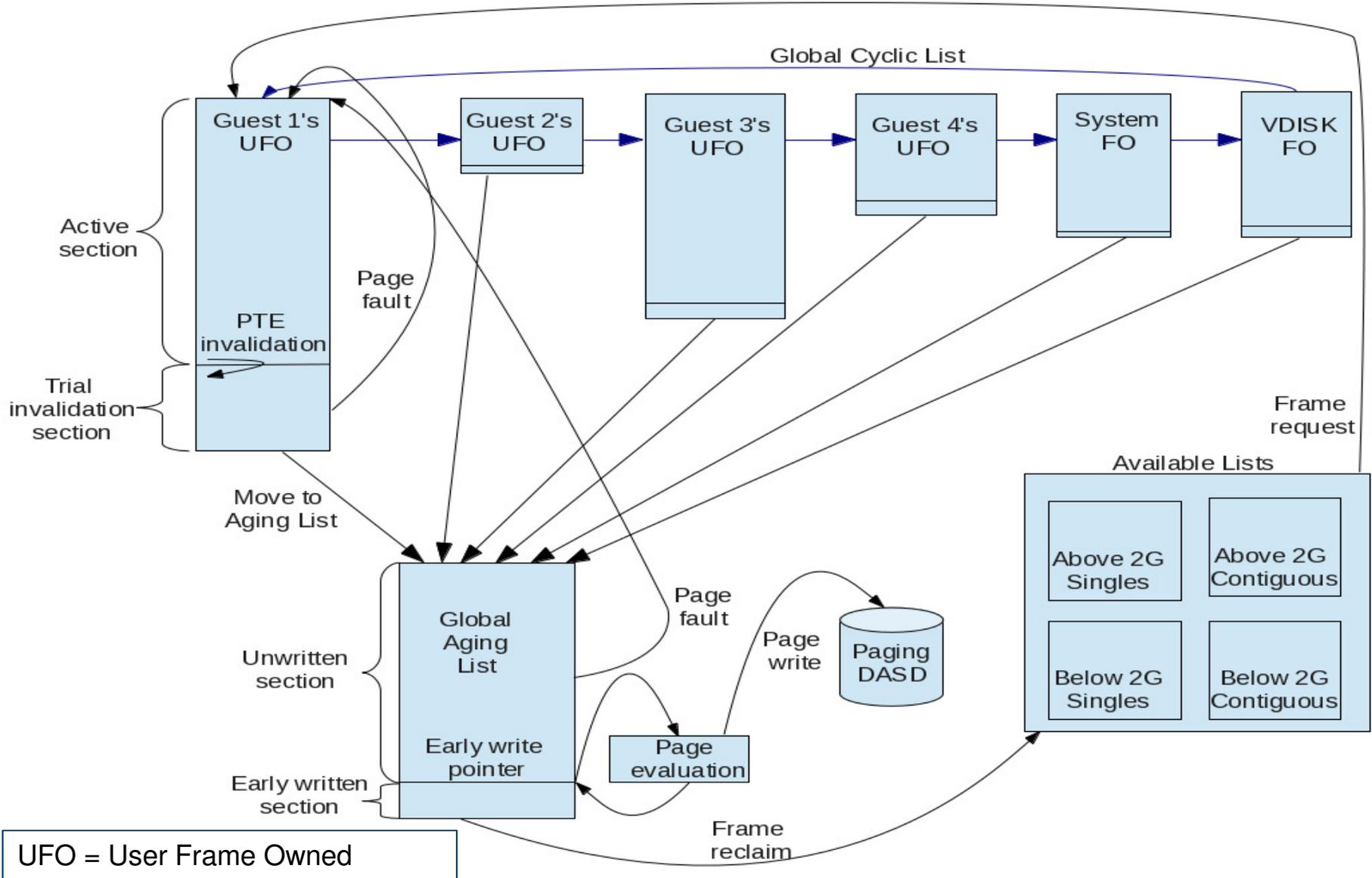
Agenda

- Focus on two main areas changed during z/VM 6.3
 - HiperDispatch
 - Memory Management

- For each area:
 - Review the technology
 - Reference monitor records changed
 - Look at the key metrics
 - Discuss the new relevant Performance Toolkit Reports

Large Memory

Memory Management Algorithm Visualization



UFO = User Frame Owned

Large Memory CP Monitor Changes

Domain	Record	Name	Type	Title	Fields, N / D / C
D0	R3	MRSYTRSG	sample	Real Storage Data (Global)	D C
D0	R4	MRSYTRSP	sample	Real Storage Data (Per Processor)	D
D0	R6	MRSYTASG	sample	Auxiliary Storage (Global)	N C
D0	R7	MRSYTSHS	sample	Shared Storage Data	D
D0	R23	MRSYTLCK	sample	Formal Spin Lock Data	N C
D1	R7	MRMTRMEM	config	Memory Configuration Data	N
D1	R15	MRMTRUSR	config	Logged on User	C
D2	R4	MRSCCLADL	event	Add User to Dispatch List	D C
D2	R5	MRSCCLDDL	event	Drop User from Dispatch List	D C
D2	R6	MRSCCLAEL	event	Add User to Eligible List	C
D2	R8	MRSCCLSTP	event	System Timer Pop	D
D3	R1	MRSTORSG	sample	Real Storage Management (Global)	N D C
D3	R2	MRSTORSP	sample	Real Storage Activity (Per Processor)	D
D3	R3	MRSTOSHR	sample	Shared Storage Management	N C
D3	R14	MRSTOASI	sample	Address Space Information Record	N C
D3	R15	MRSTOSHL	event	NSS/DCSS/SSP Loaded into Storage	N
D3	R16	MRSTOSHD	event	NSS/DCSS/SSP Removed From Storage	N C
D4	R2	MRUSELOF	event	User Logoff Data	N D C
D4	R3	MRUSEACT	sample	User Activity Data	N D C
D4	R9	MRUSEATE	event	User Activity Data at Transaction End	D C

z/VM Performance Toolkit: Highlights

- **Changed screens:**

- FCX102 SYSTEM, Some Internal System Counters
- FCX103 STORAGE, General Storage Utilization
- FCX133 NSS, NSS and DCSS Utilization and Paging Activity
- FCX146 AUXLOG, Auxiliary Storage Utilization, by Time
- FCX147 VDISKS, Virtual Disks in Storage
- FCX265 LOCKLOG, Spin Lock Log, by Time

- **Deleted screens:**

- FCX254 AVAILLOG, Available List Management, by Time
- FCX259 DEMNDLOG, Demand Scan Details, by Time

- **New screens:**

- FCX290 UPGACT, User Page Activity
- FCX291 UPGACTLG, User Page Activity (benchmarks a user)
- FCX292 UPGUTL, User Page Utilization Data
- FCX293 UPGUTLLG, User Page Utilization Data (benchmarks a user)
- FCX294 AVLB2GLG, Available List Data Below 2G, by Time
- FCX295 AVLA2GLG, Available List Data Above 2G, by Time
- FCX296 STEALLOG, Steal Statistics, by Time
- FCX297 AGELLOG, Age List Log, by Time

page state transition rates

page residency counts

available list counts

steal algorithm activity

global aging list activity

Key Considerations

- Do I have enough page space?
- Should Early Write be ON (default) or OFF?
- Do I have eligible lists forming?
- How much memory are virtual machines really using?
- How is SET RESERVE working?
- How effective is the local Invalid But Resident section?
- How effective is the global Age List?

z/VM Performance Toolkit: New Columns and Concepts

New Field	What this means
Inst	<i>Instantiations</i> : the rate at which valid memory is being created <i>Instantiated</i> : the amount of valid memory
Relse	<i>Releases</i> : the rate at which memory is being released
Inval	<i>Invalidations</i> : the rate at which demand scan is marking memory invalid as a way to determine whether it is being touched
Reval	<i>Revalidations</i> : the rate at which invalid pages are being made valid because somebody touched them
Ready	<i>Ready reclaims</i> or <i>ready steals</i> : the frame was found and selected for reclaim and had already been prewritten to auxiliary storage
Not Ready	<i>Notready reclaims</i> or <i>notready steals</i> : the frame was selected for reclaim but we had to wait for the auxiliary write (DASD) to finish before we could take it

z/VM Performance Toolkit: New Columns and Concepts

New Field	What this means
PNR	<i>Private, not referenced:</i> the page was read from aux as part of a block read, but it is still marked invalid because nobody has touched it yet
$x < 2G$ or $x > 2G$	<i>Below 2 GB or Above 2 GB:</i> tells where the real backing frames are in real central
Sing	<i>Singles:</i> free frames surrounded by in-use frames (cannot coalesce)
Cont	<i>Contigs:</i> free frames in strings of two or more
Prot	<i>Protect threshold:</i> number of frames a singles-obtain must leave on a contigs-list

Page Utilization – FCX109 – DEV CPOWN

FCX109 Data for 2014/02/03 Interval 07:28:00 – 07:29:00 Monitor Scan

Page / SPOOL Allocation Summary

PAGE slots available	235865k	SPOOL slots available	4808160
PAGE slot utilization	17%	SPOOL slot utilization	59%
T-Disk space avail. (MB)	DUMP slots available	0
T-Disk space utilization	...%	DUMP slot utilization	..%

< Device Descr. ->						<----- Rate/s ----->					
Addr	Devtyp	Volume Serial	Area Type	Area Extent	Used %	<--Page-->		<--Spool-->		Total	SSCH +RSCH
						P-Rds	P-Wrt	S-Rds	S-Wrt		
1020	3390-9	H2PG00	PAGE	5896620	17	23.4	13.2	36.6	5.7
1021	3390-9	H2PG01	PAGE	5896620	17	20.3	14.0	34.3	5.2
1022	3390-9	H2PG02	PAGE	5896620	17	20.5	13.1	33.6	5.2
1023	3390-9	H2PG03	PAGE	5896620	17	25.7	11.3	37.0	6.0
1024	3390-9	H2PG04	PAGE	5896620	17	26.2	11.7	37.9	6.5
1025	3390-9	H2PG05	PAGE	5896620	17	24.8	13.2	38.0	6.8
1027	3390-9	H2PG07	PAGE	5896620	17	22.7	12.0	34.7	5.8
1028	3390-9	H2PG08	PAGE	5896620	17	22.3	12.6	35.0	6.5

Page Utilization History – FCX146 - AUXLOG

FCX146 Data for 2014/02/03 Interval 07:28:00 - 07:33:00 Monitor Scan										
Interval	<Page Slots>		<Spool Slots>		<Dump Slots>		<----- Spool Files ----->			
	Total	Used	Total	Used	Total	Used	<--Created-->		<--Purged-->	
End Time	Slots	%	Slots	%	Slots	%	Total	/s	Total	/s
>>Mean>>	235865k	17	4808160	59	0	..	0	.00	0	.00
07:29:00	235865k	17	4808160	59	0	..	0	.00	0	.00
07:30:00	235865k	17	4808160	59	0	..	0	.00	0	.00
07:31:00	235865k	17	4808160	59	0	..	0	.00	0	.00
07:32:00	235865k	17	4808160	59	0	..	0	.00	0	.00
07:33:00	235865k	17	4808160	59	0	..	0	.00	0	.00

Early Writes? – FCX297 – AGELLOG (Age List Log)

```

FCX297      Data for 2013/10/15  Interval 09:28:00 - 09:29:00  Monitor Scan

                <----- Storage ----->
                        <-- Steal Ready ----> <--- Not Ready ---->
Interval      Size S E <-List Size--> <--RefOnly--> <--Changed--> <Evaluating->
End Time      %DPA Z W Target Current  NoWrt   Write   Write PndWrt  Refd Change
>>Mean>>     2.0 V Y  7787M   7787M   299M    0    480M  3884M   24K    0
09:29:00     2.0 V Y  7787M   7787M   300M    0    479M  3874M   48K    0

```

- Running with default 2% of DPA
- Early Writes is ON (“Y”)

Early Writes? – Write vs. Read – FCX143 - PAGELOG

```
FCX143      Data for 2013/10/15  Interval 09:28:00 - 09:29:00
```

```
<----- Paging to DASD ----->
```

```
      <-Single Reads-->
```

Reads	Write	Total	Shrd	Guest	System	Total
/s	/s	/s	/s	/s	/s	/s
981.3	603.3	1585	46.9	302.2	1.1	303.3

- Compare Writes/Second to Reads/Second
 - Reads can be > Writes if pages aren't being changed
 - Writes can be > Reads if the pages aren't being re-referenced and sit idle on DASD
 - Writes can be >> Reads if written during early write, but revalidated before actually stolen

Early Writes Revalidated – FCX297 - AGELLOG

```

FCX297      Data for 2013/10/15  Interval 09:28:00 - 09:29:00  Monitor Scan
<----- Storage ----->          <----- Revalidation ----->
<-- Steal Ready ---> <--- Not Ready ---->    %Of <----- Storage/s ----->
<--RefOnly--> <--Changed--> <Evaluating-> Pages <--RefOnly--> <--Changed-->
  NoWrt  Write  Write PndWrt   Refd Change  Eval  NoWrt  Write  NoWrt  Write
    299M    0   480M  3884M    24K    0    10 560742    .0  2303K  21026

```

- You see above that most of the revalidated pages are pages that were not written yet. Though the majority of those were ones that would have been written.

Eligible Lists Forming? – FCX145 - SCHEDLOG

```
FCX145      Data for 2013/10/15  Interval 09:28:00 - 10:05:00
<- In Eligible List -->
           <Loading->
    E1  E2  E3  E1  E2  E3
    .0  .0  .0  .0  .0  .0
    .0  .0  .0  .0  .0  .0
```

- Subtle changes in “Loading Users” in z/VM 6.3 can cause inadvertent eligible lists.
- Keep an eye on SCHEDLOG and the subset of users in eligible list that are “Loading Users”

Eligible Lists Forming? – FCX154 - SYSSET

FCX154	Data for 2013/10/15	System	Settings	Monitor Scan
Initial Scheduler Settings: 2013/10/15 at 09:27:50				
DSPSLICE (minor)	5.000 msec.	IABIAS	Intensity	90 Percent
Hotshot T-slice	1.999 msec.	IABIAS	Duration	2 Minor T-slices
DSPBUF Q1	32767 Openings	STORBUF	Q1 Q2 Q3	300 % Main storage
DSPBUF Q1 Q2	32767 Openings	STORBUF	Q2 Q3	300 % Main storage
DSPBUF Q1 Q2 Q3	32767 Openings	STORBUF	Q3	300 % Main storage
LDUBUF Q1 Q2 Q3	100 % Paging exp.	Max. working set	9999	% Main storage
LDUBUF Q2 Q3	95 % Paging exp.	Loading user	5	Pgrd / T-slice
LDUBUF Q3	85 % Paging exp.	Loading capacity	47	Paging expos.

- Review LDUBUF settings and Loading capacity
- From above example, 40 loading users in Q3 would cause an eligible list to form.
 - $.85 \times 47 = 39.95$

Virtual Machine Memory Usage – FCX292 - UPGUTL

```

FCX292      Data for 2013/10/15  Interval 10:04:00 - 10:05:00  Monitor Scan
-----
              <----- Storage ----->
                <----- Resident ----->
          Data          <----- Invalid But Resident ----->          Base
        Spaces          <---- Total ----> <-Locked--> <-- UFO --> <-- PNR --> <-AgeList->          Space
Userid   Owned   WSS  Inst  Resvd  T_All  T<2G  T>2G  L<2G  L>2G  U<2G  U>2G  P<2G  P>2G  A<2G  A>2G  XSTOR  AUX  Size
>>Mean>>   .9 1807M 2669M 86780 1529M 7588K 1522M 7567 504K 2378 550K 76557 11M 168K 33M .0 2222M 3315M
DJS LA101   0 5120M 5113M   0 4404M 19M 4384M   0 208K   0 960K 16K 11M 280K 55M 0 3434M 5120M

```

Data

```

          Spaces
Userid   Owned   WSS  Inst  Resvd
>>Mean>>   .9 1807M 2669M 86780
DJS LA305   0 3100M 6728M   0

```

- “Inst” = pages guest has interacted with in some way which requires z/VM to back the page.
 - Up to the size of the virtual machine
 - Often less than sum of (Resident+XSTOR+AUX) because of pages kept on DASD and in real memory

Reserved? – FCX292 - UPGUTL

```

FCX292      Data for 2013/10/15  Interval 10:04:00 - 10:05:00  Monitor Scan
-----
              <----- Storage ----->
                <----- Resident ----->
              Data
              Spaces
              <---- Total ----> <-Locked--> <--- UFO ---> <--- PNR ---> <-AgeList->
Userid      Owned   WSS  Inst Resvd T_All  T<2G  T>2G  L<2G  L>2G  U<2G  U>2G  P<2G  P>2G  A<2G  A>2G  XSTOR  AUX   Base
WJBLA101    0 5120M 5113M  20M 4404M  19M 4384M   0 208K   0 960K  16K  11M 280K  55M   0 3434M 5120M
    
```

```

          Data
          Spaces
Userid    Owned   WSS  Inst Resvd
>>Mean>>   .9 1807M 2669M 86780
WJBLA101   0 5120M 5113M  20M
    
```

- “Resvd” = Amount of pages reserved. May be larger than number of resident pages if virtual machine has not instantiated that memory yet.
- Note that memory is now in bytes (suffixed) not pages.

Virtual Machine Activity – FCX292 - UPGUTL

```

FCX292      Data for 2013/10/15  Interval 10:04:00 - 10:05:00  Monitor Scan
-----
          <----- Storage ----->
                <----- Resident ----->
          Data          <----- Invalid But Resident ----->          Base
          Spaces          <----- Total -----> <-Locked--> <-- UFO --> <-- PNR --> <-AgeList->          Space
Userid   Owned   WSS  Inst Resvd T_All  T<2G  T>2G  L<2G  L>2G  U<2G  U>2G  P<2G  P>2G  A<2G  A>2G  XSTOR  AUX  Size
WJBLA101  0 5120M 5113M 20M 4404M 19M 4384M 0 208K 0 960K 16K 11M 280K 55M 0 3434M 5120M
    
```

```

          <----- Resident ----->
                <----- Invalid But Resident ----->
          <----- Total -----> <-Locked--> <-- UFO --> <-- PNR --> <-AgeList->
Userid   T_All  T<2G  T>2G  L<2G  L>2G  U<2G  U>2G  P<2G  P>2G  A<2G  A>2G
WJBLA101 4404M 19M 4384M 0 208K 0 960K 16K 11M 280K 55M
    
```

- Get an understanding of where in the lists pages reside:
 - IBR = Invalid But Resident
 - UFO = User Framed Owned section
 - PNR = Private Not Referenced
 - AgeList = part of global age list, but still associated with virtual machine.

z/VM Performance Toolkit: New Report FCX295 AVLA2GLG

```

FCX295  Run 2013/04/10 07:38:36          AVLA2GLG          Page 25
                                     Available List Data Above 2G, by Time
From 2013/04/09 16:02:10                SYSTEMID
To   2013/04/09 16:13:10                CPU 2817-744   SN A6D85
For   660 Secs 00:11:00                 "This is a performance report for SYSTEM XYZ" z/VM   V.6.3.0 SLU 0000

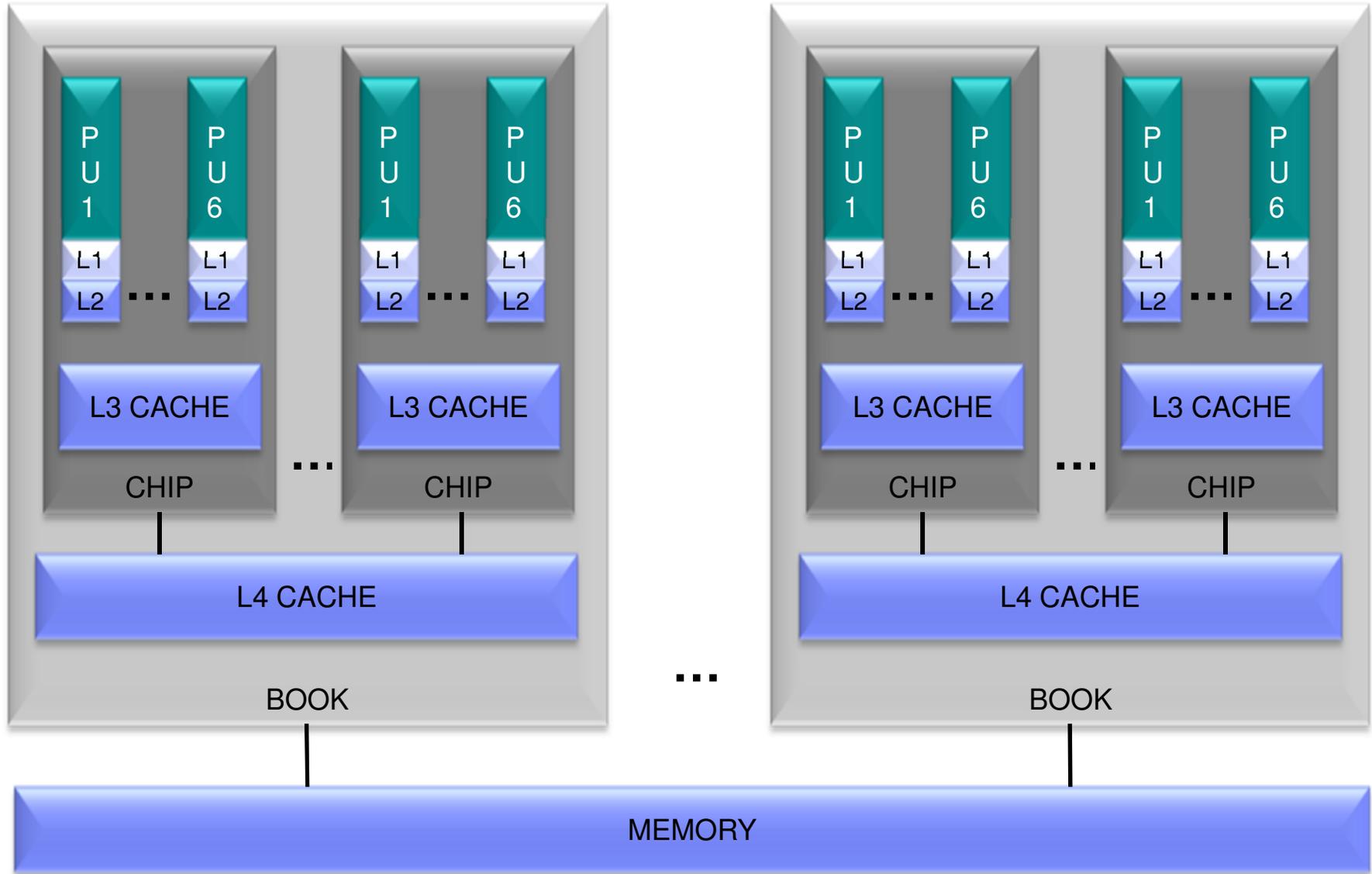
```

Interval	<----- Storage ----->				<--Times-->		<-Frame Thresh-->				
	<Available>		<Requests/s>		<Returns/s>		<-Empty/s->		Sing	<-Contigs->	
End Time	Sing	Cont	Sing	Cont	Sing	Cont	Sing	Cont	Low	Low	Prot
>>Mean>>	23M	267M	47M	59M	47M	51M	.0	.0	1310	15	15
16:02:40	0	938M	32M	126M	502K	30310	.0	.0	1332	15	15
16:03:10	152K	4556K	50M	89M	49M	59M	.0	.0	1168	15	15
16:03:40	400K	4824K	68M	82M	71M	79M	.0	.0	1321	15	15
16:04:10	0	5896K	49M	72M	52M	70M	.0	.0	2409	15	15
16:04:40	0	2124K	40M	60M	41M	59M	.0	.0	1308	15	15
16:05:10	876K	3488K	54M	52M	55M	51M	.0	.0	1118	15	15
16:05:40	0	3624K	53M	58M	54M	57M	.0	.0	1409	15	15
16:06:10	2016K	4464K	49M	57M	51M	56M	.0	.0	1273	15	15

- Look for the new concepts: Singles Contigs Prot
- Amounts are in bytes, suffixed. Not page counts!
- FCX254 AVAILLOG is no longer produced.

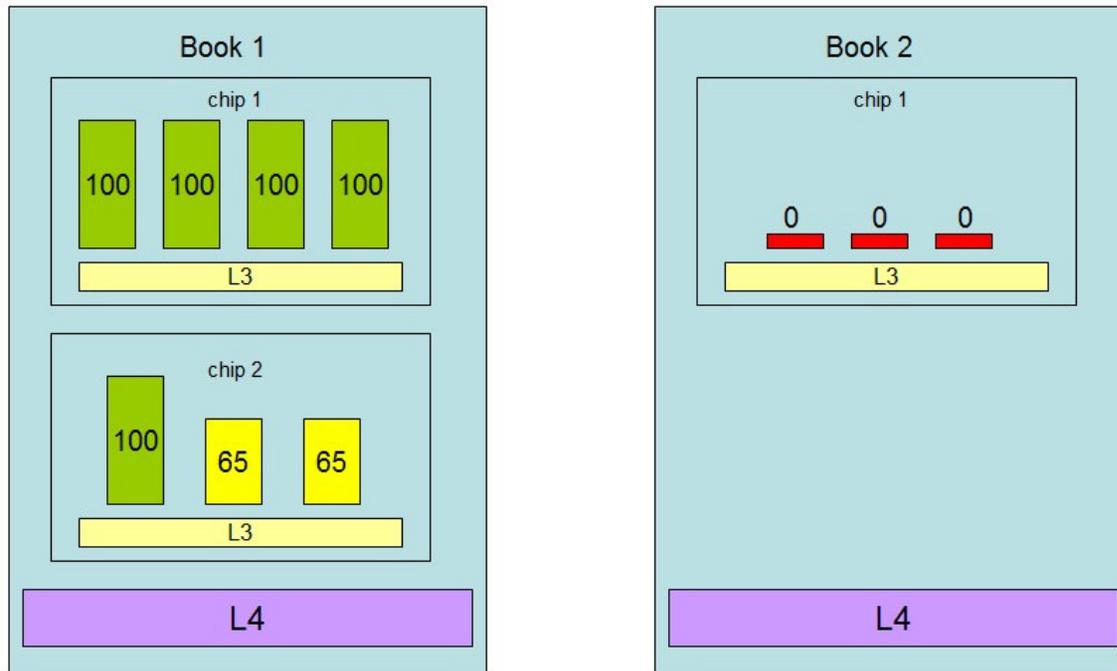
HiperDispatch

HiperDispatch- Dispatching Affinity



IBM System z: The Partition Knows Its Placement

Partition Topology



In vertical partitions:

- Sense your placement
- Run work smartly in light of your placement
- Sense unentitled power
- Use LPUs smartly in light of unentitled power

Notice PR/SM has given this partition a “quiet place” to do its work, provided the partition runs its work on its Vh LPUs.

Key Considerations

- Running Horizontal or Vertical? Other related SRM Settings?
- Is Global Performance Data enabled?
- What's my Topology?
- What's my entitlement?
- Are processors being parked?

z/VM HiperDispatch: CP Monitor Records

Domain	Record	Name	Type	Description of Change
D0	R2	MRSYTPRP	sample	Added polarity, entitlement, and park-time accumulator
D0	R16	MRSYTCUP	sample	Added partition current weight
D0	R23	MRSYTLCK	sample	Added the HCPDSVTL topology lock
D1	R4	MRMTRSYS	config	Added bit indicating whether system is horizontal or vertical
D1	R5	MRMTRPRP	config	Added park state, polarization, entitlement, and topological location
D1	R16	MRMTRSCH	config	Added h/v bit, CPUPAD settings, and EXCESSUSE settings
D2	R7	MRSCLSRM	event	Added h/v bit, CPUPAD settings, and EXCESSUSE settings
D4	R2	MRUSELOF	event	Added rebalance results and steal results
D4	R3	MRUSEACT	sample	Added rebalance results and steal results
D5	R2	MRPRCVOF	event	Added park/unpark failure as reason varied off
D5	R3	MRPRCVON	event	Added parked as a state; use if neither D5 R17 nor D5 R18 are seen
D5	R15 (new)	MRPRCDSV	event	Records assignment of processors to dispatch vectors
D5	R16 (new)	MRPRCPUP	event	Records park/unpark decision
D5	R17 (new)	MRPRCRCD	sample	Records processor's VMDBK steal behavior
D5	R18 (new)	MRPRCDHF	sample	Records PLDV population trends

As usual, the monitor records will be on www.vm.ibm.com at GA.

z/VM HiperDispatch: z/VM Performance Toolkit

- Themes in the changes in existing Perfkit screens
 - CPU entitlement appears in sensible places, e.g. FCX100 CPU
 - Percent-parked appears in sensible places, e.g. FCX100 CPU
 - Parked time is correctly accounted for, e.g. FCX126 LPAR %Susp
 - SRM settings are reported where they ought to be, e.g. FCX154 SYSSET
 - Interesting events are reported in FCX180 SYSCONF as they should
 - Number of unparked CPUs appears in sensible places, e.g. FCX225 SYSSUMLG
 - Counts of new monitor records appear in FCX155 MONDATA as they should
 - Obsolete data is compatibly deleted in certain places, e.g. FCX144 PROCLOG

- New reports sure to attract interest:
 - FCX287 TOPOLOG shows a log of partition topology, container-major
 - FCX298 PUORGLOG shows a log of partition topology, CPU-major
 - FCX299 PUCFGLOG shows a log of the park/unpark state
 - FCX301 DSVBKACT replaces the PLDV emptiness columns on FCX144 PROCLOG
 - FCX302 PHYSLOG shows a physical CPU utilization log of the CEC by type pool
 - FCX303 DSVSLOG replaces the PLDV steal columns on FCX144 PROCLOG
 - FCX304 PRCLOG is where you should now look instead of FCX144 PROCLOG
 - FCX306 LSHARACT reports the partitions' entitlements vs. logical CPU counts

- Obsolete reports
 - FCX144 PROCLOG is still there for now, but start using FCX304 PRCLOG instead

Vertical? SRM Settings? – FCX154 - SYSSET

```

FCX154          Data for 2013/10/15          System  Settings          Monitor Scan

Initial Scheduler Settings: 2013/10/15 at 09:27:50
DSPSLICE (minor) 5.000 msec.          IABIAS Intensity          90 Percent
Hotshot T-slice  1.999 msec.          IABIAS Duration           2 Minor T-slices
DSPBUF Q1         32767 Openings       STORBUF Q1 Q2 Q3         300 % Main storage
DSPBUF Q1 Q2      32767 Openings       STORBUF Q2 Q3           300 % Main storage
DSPBUF Q1 Q2 Q3   32767 Openings       STORBUF Q3              300 % Main storage
LDUBUF Q1 Q2 Q3   100 % Paging exp.     Max. working set        9999 % Main storage
LDUBUF Q2 Q3      100 % Paging exp.     Loading user             5 Pgrd / T-slice
LDUBUF Q3         100 % Paging exp.     Loading capacity         47 Paging expos.
LIMITHARD algorithm Consumption
DSPWD method      Reshuffle
Polarization      Vertical
Global Perf. Data ON
EXCESSUSE: CP    MEDIUM  CPUPAD: CP    100%
              ZAAP MEDIUM  ZAAP    100%
              IFL  MEDIUM  IFL    1000%
              ICF  MEDIUM  ICF    100%
              ZIIP MEDIUM  ZIIP    100%

```

- Default Vertical Polarization
- Global Performance Data is On
- Using default values for EXCESSUSE and CPUPAD

Topology? – FCX298 - PUORGLOG

1FCX298 Run 2013/05/20 10:39:48

PUORGLOG

Processor Unit organization log

From 2013/05/19 03:39:31

To 2013/05/19 03:41:31

For 120 Secs 00:02:00

Result of GF003855 Run

Logical PU organization for Partition PPRF1 (GDLBOFVM)

Date	Time	CPU	Type	PPD	Ent.	Location
05/19	03:39:31	0	CP	VhD	100	1:6
05/19	03:39:31	1	CP	VhD	100	1:6
05/19	03:39:31	2	CP	VhD	100	1:5
05/19	03:39:31	3	CP	VhD	100	1:5
05/19	03:39:31	4	CP	VhD	100	1:5
05/19	03:39:31	5	CP	VhD	100	1:5
05/19	03:39:31	6	CP	VhD	100	1:5
05/19	03:39:31	7	CP	VhD	100	1:4
05/19	03:39:31	8	CP	VhD	100	1:4
05/19	03:39:31	9	CP	VhD	100	1:4
05/19	03:39:31	10	CP	VhD	100	1:4
05/19	03:39:31	11	CP	VhD	100	1:2
05/19	03:39:31	12	CP	VhD	100	1:2
05/19	03:39:31	13	CP	VhD	100	1:2
05/19	03:39:31	14	CP	VhD	100	1:2
... truncated ...						

Notes:

Vh: vertical high

Vm: vertical medium

VI: vertical low

VhD: vertical high, dedicated partition

Ent: entitlement with respect to a physical CPU

Location: book:chip (z10: book)

Is there Parking? – FCX304 - PRCLOG

```

FCX304          Data for 2013/10/15  Interval 09:28:00 - 10:05:
                                     <--- Percent Busy ---->
                                     Pct
Interval      C      Pct
End Time     U Type PPD Ent. DVID Time Total  User  Syst  Emul
>>Mean>>    0 IFL  vh  100 0000  0  24.9  21.4  3.5  20.6
>>Mean>>    1 IFL  vh  100 0001  0  36.8  34.7  2.1  33.7
>>Mean>>    2 IFL  vh  100 0002  0  37.6  36.0  1.6  35.1
>>Mean>>    3 IFL  vh  100 0003  0  38.4  36.9  1.4  36.1
>>Mean>>    4 IFL  vh  100 0004  0  43.7  42.4  1.3  41.6
>>Mean>>    5 IFL  vh  100 0005  0  39.1  37.4  1.8  36.3
>>Mean>>    6 IFL  vh  100 0006  0  37.9  36.4  1.4  35.5
>>Mean>>    7 IFL  vh  100 0007  0  40.5  39.3  1.3  38.4
>>Mean>>    8 IFL  vh  100 0008  0  40.6  39.4  1.2  38.5
>>Mean>>    9 IFL  vh  100 0009  0  42.0  40.7  1.3  39.9
>>Total>   10 IFL  vh  1000  MIX  0 381.5 364.6 16.9 355.6

```

- PRCLOG replaces PROCLOG
- Pct Park Time – remember processors may be parked and unparked multiple times in an interval.

Is there Parking? - FCX299 - PUCFGLOG

FCX299 Run 2013/06/24 09:36:54

PUCFGLOG
Processor Unit Configuration log

Page 6

From 2013/02/19 11:49:52
To 2013/02/19 11:56:10
For 378 Secs 00:06:18

Result of GFCM0107 Run

GFCM0107
CPU 2817-744 SN B6D85
z/VM V.6.3.0 SLU 0000

Date	Time	Type	OnL	Entitl	Type	Cap	CPUPAD	EX	Load	XP	XPF	T/V	LCEi	XPF	T/V	N	NotVh	UpCap	LPU	Unparked	mask
02/19	11:49:54	CP	24	1985.2	...	100.0	70	2.2	1159.4	892.8	3.519	3.9	885.9	200.5	2	.0	200.0	00300000	00000000		
02/19	11:49:56	CP	24	1985.2	...	100.0	70	.5	1153.3	888.1	256.0	1.7	883.4	201.3	2	.0	200.0	00300000	00000000		
02/19	11:49:58	CP	24	1985.2	...	100.0	70	.5	1159.7	893.1	122.3	1.7	885.2	204.2	2	.0	200.0	00300000	00000000		
02/19	11:50:00	CP	24	1985.2	...	100.0	70	.7	1136.7	875.4	53.45	1.7	857.7	172.5	2	.0	200.0	00300000	00000000		
02/19	11:50:02	CP	24	1985.2	...	100.0	70	.9	1128.6	869.2	4.531	1.7	863.0	172.5	2	.0	200.0	00300000	00000000		
02/19	11:50:04	CP	24	1985.2	...	100.0	70	1.3	1034.5	778.8	1.822	1.8	688.3	172.4	2	.0	200.0	00300000	00000000		
02/19	11:50:06	CP	24	1985.2	...	100.0	70	.6	1157.1	891.1	38.57	1.8	856.4	168.5	2	.0	200.0	00300000	00000000		
02/19	11:50:08	CP	24	1985.2	...	100.0	70	.5	1162.9	895.5	250.8	1.7	856.9	211.1	2	.0	200.0	00300000	00000000		
02/19	11:50:10	CP	24	1985.2	...	100.0	70	44.8	1161.8	894.7	2.214	89.1	858.9	211.1	2	.0	200.0	00300000	00000000		
02/19	11:50:12	* CPU		Park/Unpark	State																
02/19	11:50:12	CP	24	1985.2	...	100.0	70	199.7	1145.1	881.9	1.517	354.6	858.5	197.6	5	.0	500.0	00300000	00000000		
02/19	11:50:14	* CPU		Park/Unpark	State																
02/19	11:50:14	CP	24	1985.2	...	100.0	70	501.6	1155.6	890.0	1.009	803.5	858.3	197.5	10	.0	1000.0	013C0000	00000000		
02/19	11:50:16	* CPU		Park/Unpark	State																
02/19	11:50:16	CP	24	1985.2	...	100.0	70	999.6	1147.4	883.6	1.001	1497.6	857.9	146.5	16	.0	1600.0	0FFC0000	00000000		
02/19	11:50:18	* CPU		Park/Unpark	State																
02/19	11:50:18	CP	24	1985.2	...	100.0	70	1599.3	1155.1	889.6	1.001	2199.1	857.7	130.3	23	100.0	2300.0	FFFFF000	00000000		
02/19	11:50:20	* CPU		Park/Unpark	State																
02/19	11:50:20	CP	24	1985.2	...	100.0	70	2297.6	1179.7	908.5	1.001	2995.8	860.2	125.6	24	100.0	2400.0	FFFFFFE0	00000000		
02/19	11:50:22	* CPU		Park/Unpark	State																
02/19	11:50:22	CP	24	1985.2	...	100.0	70	2397.1	1144.5	881.4	1.005	2496.6	854.3	125.4	24	100.0	2400.0	FFFFFFF0	00000000		
02/19	11:50:24	CP	24	1985.2	...	100.0	70	2080.5	1181.8	910.1	1.002	2569.2	887.6	125.3	24	100.0	2400.0	FFFFFFF0	00000000		
02/19	11:50:26	CP	24	1985.2	...	100.0	70	1681.3	1140.0	878.0	1.002	2660.9	845.8	122.1	24	100.0	2400.0	FFFFFFF0	00000000		
02/19	11:50:28	CP	24	1985.2	...	100.0	70	1632.4	1169.6	900.7	1.002	2684.7	886.2	1.660	24	100.0	2400.0	FFFFFFF0	00000000		
02/19	11:50:30	CP	24	1985.2	...	100.0	70	1587.7	1149.4	885.2	1.002	2635.4	869.6	1.252	24	100.0	2400.0	FFFFFFF0	00000000		
02/19	11:50:32	CP	24	1985.2	...	100.0	70	1878.3	1129.6	869.9	1.011	2560.8	854.7	1.008	24	100.0	2400.0	FFFFFFF0	00000000		
02/19	11:50:34	CP	24	1985.2	...	100.0	70	1824.3	1176.2	905.8	1.002	2425.8	884.3	1.007	24	100.0	2400.0	FFFFFFF0	00000000		

- Shows what can happen to T/V when utilization is really low
- Shows parking because of high projected T/V
- Shows unpark when workload ramps up
- Shows how a varying U produces a high U'
- Shows XPF and XPF' in action
- Shows that the T/V projections level right out once the guest workload is well underway
- Shows what a non-Vh LPU will be "truly worth" during the next interval

Entitlement? – FCX306 - LSHARACT

```
1FCX306 Run 2013/02/19 12:10:57 LSHARACT
LPAR Data, Collected in Partition RPRF2
Physical PUs, shared: CP- 40 ZAAP- 2 IFL- 16 ICF- 1 ZIIP- 3
Dedicated: CP- 4 ZAAP- 0 IFL- 0 ICF- 0 ZIIP- 0
```

Proc Type	Partition Name	LPU Num	LPAR Weight	Entlment	<LPU Total, %> Busy	Excess	LPU Conf
CP	RCPX4	10	10	59.3	3.0	.0	o
CP	RCTS1	5	10	59.3	311.9	252.6	o
CP	RCTS2	5	30	177.8	1.0	.0	o
CP	RCT1	20	30	177.8	111.3	.0	o
CP	RCT2	10	10	59.3	11.2	.0	o
CP	REXT1	5	10	59.3	.0	.0	o
CP	REXT2	4	10	59.3	.0	.0	o
CP	RINS	10	10	59.3	.0	.0	o
CP	RPRF1	4	DED				
CP	RPRF2	24	335	1985.2	1548.4	.0	o
CP	RSPX1	6	40	237.0	481.3	244.3	o
CP	RSPX2	6	40	237.0	499.7	262.7	o
CP	RSPX5	6	40	237.0	126.5	.0	o
CP	RST1	10	10	59.3	16.2	.0	o
CP	RST1X	6	10	59.3	102.5	43.2	o
CP	RST2	6	50	296.3	.9	.0	o
CP	RST3	3	30	177.8	1.2	.0	o
ICF	RCTS2	1	10	25.0	.0	.0	-
ICF	RCT1	1	30	75.0	.0	.0	-
IFL	RCTS2	2	10	188.2	.0	.0	-
IFL	RCT1	2	30	564.7	.0	.0	u
IFL	RSTL1	16	45	847.1	449.2	.0	o
ZAAP	RCPX4	1	10	40.0	.1	.0	-
ZAAP	RCTS2	1	10	40.0	.0	.0	-
ZAAP	RCT1	1	30	120.0	.0	.0	u
ZIIP	RCPX4	1	10	60.0	.3	.0	-
ZIIP	RCTS2	1	10	60.0	.0	.0	-
ZIIP	RCT1	1	30	180.0	.0	.0	u

- Reports by partition & CPU type
- Reports entitlement in percent
- Reports percent-busy of the partition's CPUs of that type
- Reports whether the partition is consuming beyond its entitlement ("Excess")
- Reports LPU configuration with respect to entitlement:
 - "o" – overconfigured
 - "u" – underconfigured
 - "-" – apparently just right

Total CEC? – FCX302 - PHYSLOG

FCX302 Run 2013/06/24 09:36:54

From 2013/02/19 11:49:58
 To 2013/02/19 11:56:10
 For 372 Secs 00:06:12

PHYSLOG
 Real CPU Utilization Log

Result of GFCM0107 Run

Interval	<PU Num>	Total									
End Time	Type	Conf	Ded	weight	%Lgc1P	%Ovrhd	LpuT/L	%LPmgt	%Total	TypeT/L	
>>Mean>>	CP	44	4	675	3387.1	27.947	1.008	31.870	3446.9	1.018	
>>Mean>>	ZAAP	2	0	50	.093	.042	1.451	.424	.559	6.015	
>>Mean>>	IFL	16	0	85	448.16	1.017	1.002	2.108	451.28	1.007	
>>Mean>>	ICF	1	0	40	.004	.003	1.624	2.257	2.263	563.66	
>>Mean>>	ZIIP	3	0	50	.193	.090	1.465	1.204	1.487	7.694	
>>Mean>>	>Sum	66	4	900	3835.5	29.099	1.008	37.864	3902.5	1.017	
11:50:04	CP	44	4	675	1963.9	33.262	1.017	36.226	2033.4	1.035	
11:50:04	ZAAP	2	0	50	.004	.001	1.306	.037	.042	10.107	
11:50:04	IFL	16	0	85	501.44	1.087	1.002	2.372	504.90	1.007	
11:50:04	ICF	1	0	40	.007	.004	1.566	2.277	2.289	312.13	
11:50:04	ZIIP	3	0	50	.005	.002	1.334	.093	.100	19.003	
11:50:04	>Sum	66	4	900	2465.4	34.356	1.014	41.006	2540.7	1.031	

You now have an easy way to see how busy your CEC is.

Features:

- Talled by CPU type (CP, IFL, ...)
- One group of rows every sample interval

Reports all three ways CPU gets used:

- By logical CPUs
- By PR/SM, chargeable
- By PR/SM, unchargeable

New concepts:

- LPU T/L: like "guest T/V"
- Type T/L: like "system T/V"

Dispatch Vector Activity? – FCX303 - DSVSLOG

1FCX303 Run 2013/05/20 10:32:38

DSVSLOG

DSVBK Steals per logical CPU Log, by Time

From 2013/05/19 02:03:25
To 2013/05/19 02:05:19
For 114 Secs 00:01:54

Result of GF003820 Run

Interval	C	P	Type	PPD	Ent.	DVID	Pct Park Time	DSVBK Steal /s					
End Time	U							Lvl-00	Lvl-01	Lvl-02	Lvl-03	Lvl-04	Lvl-05
>>Mean>>	0	CP	vh	100	0000	0	0	4.404	4.088	.000
>>Mean>>	1	CP	vh	100	0001	0	0	2.456	2.561	.000
>>Mean>>	2	CP	vh	100	0002	0	0	6.877	.921	.000
>>Mean>>	3	CP	vh	100	0003	0	0	7.596	.930	.000
>>Mean>>	4	CP	vh	100	0004	0	0	4.500	.482	.000
>>Mean>>	5	CP	vh	100	0005	0	0	3.614	.228	.000
>>Mean>>	6	CP	vh	100	0006	0	0	4.518	.482	.000
>>Mean>>	7	CP	vh	100	0007	0	0	2.912	.386	.000
>>Mean>>	8	CP	vh	100	0008	0	0	1.412	.421	.000
>>Mean>>	9	CP	vh	100	0009	0	0	1.386	.184	.000
>>Mean>>	10	CP	vh	100	000A	0	0	2.070	.544	.000
>>Mean>>	11	CP	vh	100	000B	0	0	2.114	.149	.000
>>Mean>>	12	CP	vh	100	000C	0	0	5.886	1.623	.000
>>Mean>>	13	CP	vh	100	000D	0	0	3.772	.702	.000
>>Mean>>	14	CP	vh	100	000E	0	0	3.026	.675	.000
>>Mean>>	15	CP	vh	100	000F	0	0	2.658	.360	.000
>>Total>	16	CP	vh	1600	MIX	0	0	59.202	14.737	.000

Reports VCPU steal behavior by the distance the steal dragged the VCPU.

- Lvl-00: you stole it from a CPU in your chip (z10: ... in your book)
- Lvl-01: you stole it from a CPU in your book (z10: ... in another book)
- Lvl-02: you stole it from a CPU on another book (z10: ... not applicable)

Caution: LPAR Suspend Time

FCX126 CPU 2827 SER 15D37 Interval 17:49:26 - 17:54:26 GDLVM7

LPAR Data, Collected in Partition GDLVM7

Processor type and model : 2827-793
 Nr. of configured partitions: 11
 Nr. of physical processors : 101
 Dispatch interval (msec) : dynamic

Partition	Nr.	Upid	#Proc	weight	wait-C	Cap	%Load	CPU	%Busy	%Ovhd	%Susp	%VMld	%Logld
GDLVM7	11	49	4	32	NO	NO	.3	0	6.3	.3	.4	6.0	6.0
				32		NO	...	1	7.0	.1	.2	6.8	6.8
				32		NO	...	2	6.6	.1	.2	6.4	6.5
				32		NO	...	2	6.6	.1	.2	6.4	6.5
				32		NO	...	3	6.8	.1	.1	6.6	6.6

- %Susp = %Suspend = an estimate of how much time the logical processor was suspended from running either due to contention with other LPARs, or in some cases z/VM voluntarily giving up control due to spin locks.
- In z/VM 6.3, this also can include “parked” time. We’re looking at how we might want to change the calculation.
- Be careful on systems where you have set an alert for higher suspend times.
- Appears on FCX126 LPAR and FCX202 LPARLOG reports

Other Changes

Monitor Record Changes

- All the HiperDispatch changes
- All the Large Memory changes
- For FCP Data Router: D1 R19, D6 R25
- HiperSockets changes: D1 R19, D6 R25, D6 R26, D6 R27
- For Large Memory Dump: D1 R7, D3 R1
- VSWITCH Edge Port Aggregator: D6 R21, D6 R35
- VSWITCH Recovery Stall Prevention: D6 R22
- Additional debug: D0 R17, D0 R20, D3 R4, D3 R11, D5 R8, D5 R10, D6 R3, D6 R4, D6 R7, D6 R8, D6 R14, D6 R31, D9 R3

z/VM Performance Toolkit

- High Performance FICON changes
 - SYSLOG, SYSTEM, DEVICE HPF, HPFLOG, SYSCONF, IOCHANGE, LCHANNEL all updated
- VSWITCH HiperSockets Bridge changes
 - GVNIC, VNIC, GVSWITCH, VSWITCH, QDIO, IOCHANGE all updated
- LGR changes
 - New reports LGRELOG and LGRDATA
- Large Memory Changes
 - 6 changed, 2 deleted, 8 new
- HiperDispatch Changes
 - 7 changed, 1 obsolete, 8 new

Questions?

Please remember to do an evaluation.

Session 15741

Thanks!

