Understanding z/VM 6.3 Through New Performance Toolkit Reports

Version 2.1

Session 17514

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

#SHAREorg

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

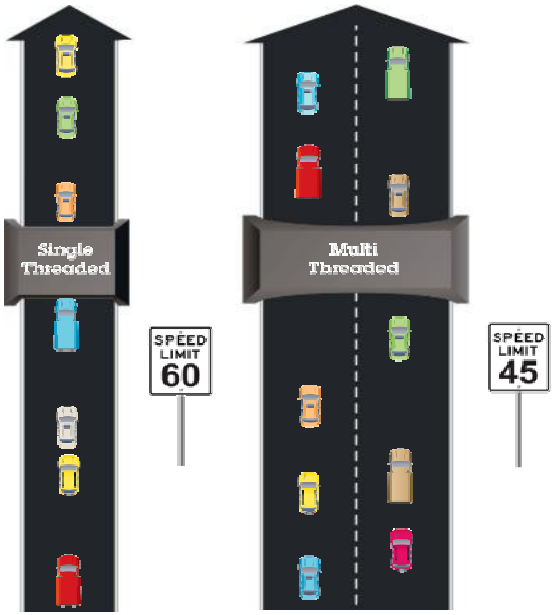
- Focus on two main areas changed during z/VM 6.3
  - SMT
  - Scalability
  - HiperDispatch
  - Memory Management

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

IBM z13 SMT Changes

Simultaneous Multithreading (SMT)

- Objective is to improve capacity, not performance.
- Allows z/VM to dispatch work on up to two threads of a z13 IFL
- VM65586 for z/VM 6.3 **only**
  - PTFs planned to be available March 13, 2015
- Transparent to virtual machine
  - Guest does not need to be SMT aware
  - SMT is not virtualized to the guest
- z13 SMT support limited to IFLs and zIIPs
  - z/VM support is only for IFLs
- SMT is disabled by default
  - Requires a System Configuration setting and re-IPL
  - When enabled, applies to the entire system
- Potential to increase the overall capacity of the system
  - Workload dependent

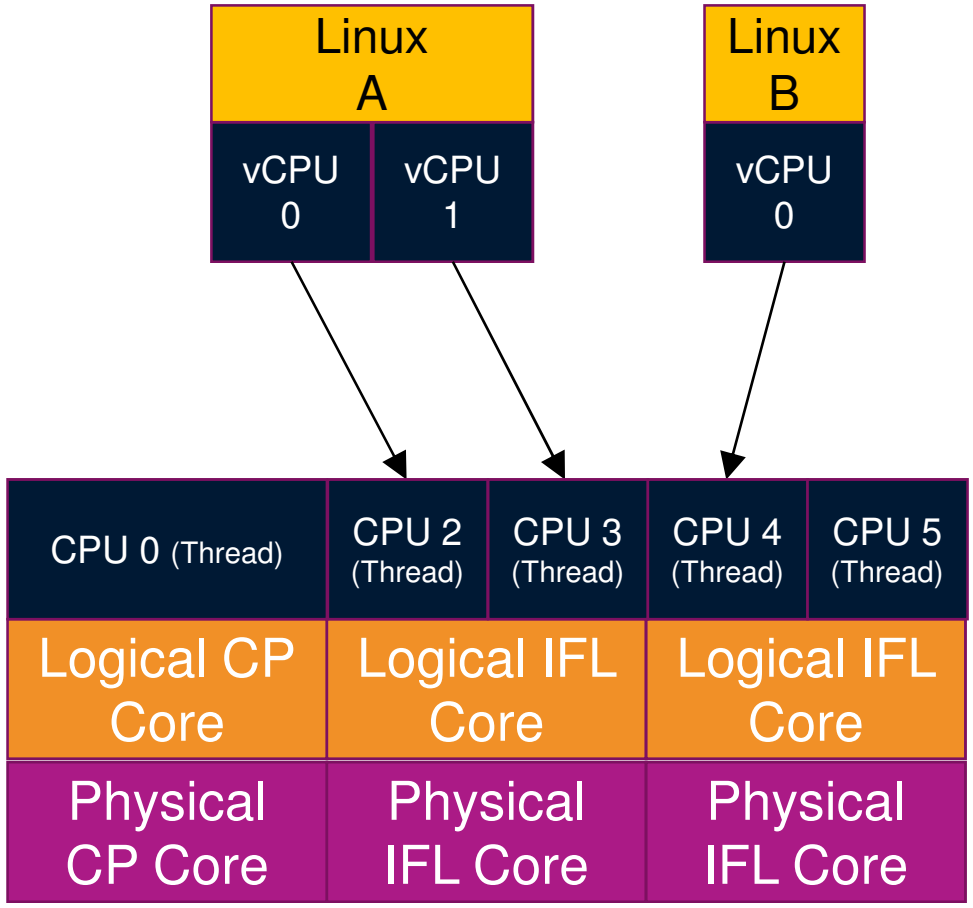


Which approach is designed for the higher volume of traffic? Which road is faster?

**Illustrative numbers only*

SMT Dispatching

- Physical IFLs (or Cores) with SMT allow up to two threads to be used. You purchase these.
- Logical IFLs are presented to z/VM as in the past, defined in LPAR definition..
- z/VM creates a CPU associated with each thread for it to use.
- The virtual CPUs of guests can then be dispatched on different threads intelligently, based on topology information.
- In a mixed-engine environment, general purpose processors can not do threading, but a second CPU address is consumed (CPU 1 in example)



Monitor Changes

New Monitor	Record Name
Domain 5 Record 20	MT CPUMF counters

Change Monitor	Records Name
Domain 0 Record 2	Processor data (per processor)
Domain 0 Record 15	Logical CPU utilization (global)
Domain 0 Record 16	CPU utilization in a logical partition)
Domain 0 Record 17	Physical CPU utilization data for LPAR management
Domain 0 Record 19	System data (global)
Domain 0 Record 23	Formal spin lock data (global)
Domain 1 Record 4	System configuration data
Domain 1 Record 5	Processor configuration data (per processor)
Domain 1 Record 16	Scheduler settings
Domain 1 Record 18	CPU capability change
Domain 2 Record 4	Add user to dispatch list
Domain 2 Record 5	Drop user from dispatch list
Domain 2 Record 7	Set SRM changes
Domain 2 Record 13	Add VMDBK to limit list
Domain 2 Record 14	Drop VMDBK from limit list
Domain 4 Record 2	User logoff data
Domain 4 Record 3	User activity data
Domain 4 Record 9	User activity data at transaction end
Domain 5 Record 1	Vary on processor
Domain 5 Record 2	Vary off processor
Domain 5 Record 11	Instruction counts per processor
Domain 5 Record 13	CPU-measurement facility counters
Domain 5 Record 16	Park/unpark decision
Domain 5 Record 17	Real CPU data
Domain 5 Record 19	CPU pool utilization

Perfkit Screen SYSCONF (FCX180) – SMT Disable

```
FCX180  Run 2015/02/15 08:52:14      SYSCONF
                                     System Configuration, Initial and Changed
From 2015/02/14 16:04:29                                                    SYSTEMID
To   2015/02/14 16:14:59                                                  CPU 2964-704
For    630 Secs 00:10:30             "This is a performance report for SYSTEM XYZ"      z/VM   V.6.3.0

Multithreading          Disabled, No MULTITHREADING statement

 Server Time Protocol (STP) facility configuration
 XRC_TEST enabled            No        XRC_OPTIONAL enabled          No
 STP H/W feature installed   No        STP H/W feature enabled       No
 STP Timestamping enabled    No        STP Timezone usage enabled    No
 STP is active               No        STP is suspended              No
 STP susp. message issued    No
 STP TOD clock offset        +00:00:00.0000000000

 Initial Status on 2015/02/14 at 16:04, Processor 2964-704
                          Total  Conf  Stby  Resvd  Ded  Shrd
 Real Proc: Cap 492.0000    103     4     0     99
 Sec. Proc: Cap 492.0000     99    99     0      4
 Log. IFL : CAF       41      8     4     4      0    4     0

 <------- Processor --------->  Core/
 Num Serial-Nr Type Status    Thread
   0   012F17  IFL  Master      00/0
   1   012F17  IFL  Alternate   01/0
   2   012F17  IFL  Alternate   02/0
   3   012F17  IFL  Alternate   03/0
 Processor Configuration Mode: LINUX
```

Disabled as Config file does not contain a ‘multithreaded enabled’ statement.

Total of 4 cores and each core has a thread 0 associated with it.

Perfkit Screen SYSCONF (FCX180) – SMT Enabled

```
FCX180  Run 2015/02/15 08:52:10    SYSCONF
                                   System Configuration, Initial and Changed
From 2015/02/14 16:31:32
To   2015/02/14 16:42:02                                                  CPU 2964-704
For    630 Secs 00:10:30           "This is a performance report for SYSTEM XYZ"      z/VM   V.6.3.0
```

```
 Multithreading          Enabled

 Initial Status on 2015/02/14 at 16:31, Processor 2964-704
                          Total  Conf  Stby  Resvd  Ded  Shrd
 Real Proc: Cap 492.0000    103     4     0     99
 Sec. Proc: Cap 492.0000     99    99     0      4
 Log. IFL : CAF       41      8     4     0      0    4     0

 <------- Processor --------->  Core/
 Num Serial-Nr Type Status    Thread
   0   012F17  IFL  Master      00/0
   1   012F17  IFL  Alternate   00/1
   2   012F17  IFL  Alternate   01/0
   3   012F17  IFL  Alternate   01/1
   4   012F17  IFL  Alternate   02/0
   5   012F17  IFL  Alternate   02/1
   6   012F17  IFL  Alternate   03/0
   7   012F17  IFL  Alternate   03/1
 Processor Configuration Mode: LINUX
```

The z/VM system is enabled for SMT.

Total of 4 cores and each core has both a thread 0 and a thread 1 associated with it.

Perfkit Screen SYSSET (FCX154) – SMT Enabled

```
FCX154  Run 2015/02/15 08:52:10          SYSSET
                                         System Scheduler Settings, Initial and Changed
From 2015/02/14 16:31:32                                                        SYSTEMID
To   2015/02/14 16:42:02                                                    CPU 2964-704
For    630 Secs 00:10:30                 "This is a performance report for SYSTEM XYZ" z/VM   V.6.3.0
```

```
Initial Scheduler Settings: 2015/02/14 at 16:31:32
LIMITHARD algorithm    Consumption
DSPWD method           Reshuffle
Polarization           Vertical
Global Perf. Data      ON
EXCESSUSE: CP   ......  CPUPAD: CP     6400%
           ZAAP ......          ZAAP      0%
           IFL  ......          IFL       0%
           ICF  ......          ICF       0%
           ZIIP ......          ZIIP      0%
Multithreading          Enabled
              <--------- Threads ---------->
              H/W Requested System Activated
Max Threads             Max                2
CP core         1       Max      1         1
IFL core        2       Max      2         2
ICF core        2       Max      1         1
ZIIP core       2       Max      1         1

Changed Scheduler Settings
Date  Time      Changed
..... ........  No changes processed
```

For SMT to be enabled:
1. z/VM Dispatch Workload Algorithm must be at default of Reshuffle.
2. HiperDispatch polarization must be vertical.

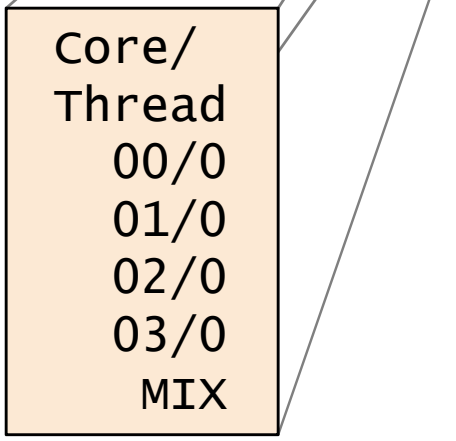
Maximum number of threads activated on this z/VM. Activated column = minimum(H/W, Requested, System)

Perfkit Screen PRCLOG (FCX304) – SMT Disabled

```
FCX304  Run 2015/02/15 08:52:14          PRCLOG                                                          Page   56
                                          Processor Activity, by Time
 From 2015/02/14 16:04:29                                                                                SYSTEMID
 To   2015/02/14 16:14:59                                                                 CPU 2964-704   SN 12F17
 For    630 Secs 00:10:30                "This is a performance report for SYSTEM XYZ"    z/VM   V.6.3.0 SLU 0000
```

							<--- Percent Busy ---->				<-- Rates per Sec. ---				<----- Paging ------->			<Co>	< Di>	
Interval End Time	C P U	Type	PPD	Ent.	DVID	Pct Park Time	Total	User	Syst	Emul	Inst Siml	DIAG	SIGP	SSCH	<2GB /s	PGIN /s	Fast Path %	Page Read /s	<mm> Msgs /s	< ag> X'9C' /s	Core/ Thread
>>Mean>>	0	IFL	VhD	100	0000	0	95.7	95.5	.2	88.2	38153	551.3	22.8	37.1	.0	.0	....	.0	.2	.0	00/0
>>Mean>>	1	IFL	VhD	100	0001	0	95.7	95.5	.2	88.2	37536	492.2	10.3	2.7	.0	.0	....	.0	.0	.0	01/0
>>Mean>>	2	IFL	VhD	100	0002	0	95.6	95.4	.2	88.0	38178	509.8	74.0	2.9	.0	.0	....	.1	.0	.0	02/0
>>Mean>>	3	IFL	VhD	100	0003	0	95.5	95.3	.2	87.8	38532	508.4	8.8	4.8	.0	.0	....	.1	.1	.0	03/0
>>Total>	4	IFL	VhD	400	MIX	0	382.5	381.6	.9	352.1	152k	2062	115.9	47.5	.0	.0	....	.2	.3	.0	MIX

Report remains similar to the past, especially with SMT disabled. You will again see the Core/Thread nomenclature.



Perfkit Screen PRCLOG (FCX304) – SMT Enabled

```
04  Run 2015/02/15 08:52:10          PRCLOG                                                    Page   56
                                      Processor Activity, by Time
m 2015/02/14 16:31:32                                                                        SYSTEMID
  2015/02/14 16:42:02                                                                  CPU 2964-704   SN 12F17
    630 Secs 00:10:30                 "This is a performance report for SYSTEM XYZ"          z/VM   V.6.3.0 SLU 0000
```

							<--- Percent Busy ---->			<-- Rates per Sec. --->			<----- Paging ------->			<Co>	< Di>	
terval	C P					Pct Park					Inst			<2GB	PGIN	Fast Path	Page Read	<mm> Msgs	< ag> X'9C'	Core/
d Time	U	Type	PPD	Ent.	DVID	Time	Total	User	Syst	Emul	Siml	DIAG	SIGP	SSCH	/s	/s	%	/s	/s	/s	Thread
Mean>>	0	IFL	VhD	100	0000	0	84.7	84.5	.2	77.0	30035	416.7	1124	34.6	.0	.0	....	.2	.2	.0	00/0
Mean>>	1	IFL	VhD	100	0000	0	84.3	84.1	.2	76.8	29845	447.8	1054	2.0	.0	.0	....	.0	.0	.0	00/1
Mean>>	2	IFL	VhD	100	0001	0	84.5	84.4	.2	76.8	31053	439.6	1098	1.4	.0	.0	....	.0	.0	.0	01/0
Mean>>	3	IFL	VhD	100	0001	0	84.6	84.4	.2	77.0	30648	491.9	1028	1.2	.0	.0	....	.0	.0	.0	01/1
Mean>>	4	IFL	VhD	100	0002	0	84.5	84.3	.2	77.0	29912	535.7	1106	1.7	.0	.0	....	.0	.0	.0	02/0
Mean>>	5	IFL	VhD	100	0002	0	84.9	84.7	.2	77.5	29667	526.1	1029	1.3	.0	.0	....	.0	.0	.0	02/1
Mean>>	6	IFL	VhD	100	0003	0	84.8	84.6	.2	77.3	29368	450.1	1062	2.1	.0	.0	....	.1	.0	.0	03/0
Mean>>	7	IFL	VhD	100	0003	0	84.7	84.5	.2	77.3	29026	566.8	1027	2.0	.0	.0	....	.0	.0	.0	03/1
Total>	8	IFL	VhD	800	MIX	0	677.0	675.5	1.5	616.6	240k	3875	8527	46.2	.0	.0	....	.2	.3	.0	MIX

With SMT enabled, you see each thread is shown as a “Logical CPU” on this report. The utilizations are of the thread, no longer the “core”.

Scalability

Areas Improved with Scalability Enhancements

- z/VM Scheduler Lock
  - Management of internal stacked work
  - Guests going into a wait state
- Locking for Memory Management
  - Most benefit during system initialization and when very constrained with memory
- Serialization and processing of VDisk I/Os
- Batching and processor-local queues for VSWITCH buffers

Perfkit Screen PRCLOG (FCX304)

FCX304 Run 2015/01/30 14:40:49 PRCLOG Page

Processor Activity, by Time

Interval End Time	CPU	Type	PPD	Ent.	DVID	Pct Park Time	<--- Percent Busy ----> Total	User	Syst	Emul	Inst Siml	<-- Ra						Core Threa
>>Mean>>	0	IFL	Vh	100	0000	0	47.1	35.3	11.8	32.0	3035							00/
>>Mean>>	1	IFL	Vh	100	0001	0	47.9	37.2	10.7	34.0	3031							01/
>>Mean>>	2	IFL	Vh	100	0002	0	49.8	39.5	10.3	36.5	2831							02/
>>Mean>>	3	IFL	Vh	100	0003	0	57.7	47.8	9.9	45.0	2646							03/
>>Mean>>	4	IFL	Vh	100	0004	0	48.0	36.0	12.0	32.7	3079							04/
>>Mean>>	5	IFL	Vh	100	0005	0	48.3	37.3	11.0	34.1	3096							05/
>>Mean>>	6	IFL	Vh	100	0006	0	50.5	40.1	10.4	37.1	2867							06/
>>Mean>>	7	IFL	Vh	100	0007	0	57.6	47.6	10.0	44.8	2678	347.4	4682	1.3	.0	.0	.0 15.1	2857 07/
>>Mean>>	8	IFL	Vh	100	0008	0	42.8	30.5	12.2	27.5	2637	257.1	9139	2.0	.0	.0	.0 25.3	3621 08/
>>Mean>>	9	IFL	Vh	100	0009	0	43.5	32.8	10.7	30.0	2610	287.3	6658	1.9	.0	.0	.0 24.3	3496 09/
>>Mean>>	10	IFL	Vh	100	000A	0	51.5	41.7	9.7	39.1	2403	357.2	5055	1.6	.0	.0	.0 20.4	3250 0A/
>>Mean>>	11	IFL	Vh	100	000B	0	47.6	36.1	11.5	32.9	3004	322.5	7643	1.9	.0	.0	.0 21.6	3912 0B/

<--- Percent Busy ----> <-- Ra

Total	User	Syst	Emul
47.1	35.3	11.8	32.0
47.9	37.2	10.7	34.0
49.8	39.5	10.3	36.5
57.7	47.8	9.9	45.0

"Syst" is pure z/VM System time, not associated with any one virtual machine, and a big part of what we worked to improve.
This system is impacted by higher System Time.

Perfkit Screen LOCKLOG (FCX265) – Scheduler Lock

```
FCX265   Run 2015/01/30 14:40:49          LOCKLOG

                                          Spin Lock Log, by Time
```

		<---------- Spin Lock Collision Activity ---------->											
		<----- Combined -------->				<----- Exclusive ------->				<------ Shared --------->			
Interval End Time	LockName	Colls /sec	AvgSpin usec	Spin %Busy	CADs /sec	Colls /sec	AvgSpin usec	Spin %Busy	CADs /sec	Colls /sec	AvgSpin usec	Spin %Busy	CADs /sec
>>Mean>>	SRMSLOCK	93695	12.23	114.6	.000	93684	12.23	114.6	.000	10.8	4.605	.005	.000
>>Mean>>	HCPTRQLK	5590	.532	.297	.000	5590	.532	.297	.000	.0	...	.000	.000
>>Mean>>	SRMALOCK	.0	...	.000	.000	.0	...	.000	.000	.0	...	.000	.000
>>Mean>>	HCPPGDPL	.0	...	.000	.000	.0	...	.000	.000	.0	...	.000	.000
>>Mean>>	HCPPGDSL	.0	...	.000	.000	.0	...	.000	.000	.0	...	.000	.000
>>Mean>>	HCPPGDTL	.0	...	.000	.000	.0	...	.000	.000	.0	...	.000	.000
>>Mean>>	HCPPGDAL	.0	...	.000	.000	.0	...	.000	.000	.0	...	.000	.000

	<----- Combined -------->			
LockName	Colls /sec	AvgSpin usec	Spin %Busy	CADs /sec
SRMSLOCK	93695	12.23	114.6	.000

- LOCKLOG gives details on Locks used. It's a log report which means it shows data over time, though for space, noly the averages are shown above. Based on wallclock time.
- SRMSLOCK is the scheduler lock, which is currently the only formal spin lock that can be used in shared mode, not just exclusive.

Perfkit Screen SYSTEM (FCX102)

```
FCX102  Run 2015/01/30 14:40:49          SYSTEM
```

Operation	Count	Rate/s
Real SSCH instructions	1411k	198
Real HSCH instructions	0	.0
SVC instr. simulated	0	.0
SVC 76 reflected	0	.0
FP external call simul.	0	.0
Fast-path SIGP simulat.	0	.0
Diag.X'9C', Not Ready	6	.0
Diag.X'9C', Soft Wait	6	.0
FP successful x-lates	188032	26.3
Fast-path aborts	8	.0
Nr. of SIE executions	1135m	158926
Entries to enabled wait	450m	63071
HPF write	50612	7.1

Operation	Count	Rate/s
Real CSCH instructions	0	.0
El. time slice drops	21499	3.0
SVC interrupts reflectd	0	.0
Diagnose I/O requests	106778	15.0
FP partial executions	1811k	254

- SYSTEM report shows various system level counters. There is a count, the absolute number over the interval, and a rate per second.
- One of the enhancements for z/VM 6.3 was more efficiently handling when guests load a wait state. Which in this data is pretty high at 63071/second. Note also the SIE executions at 158296/second. About 40% of the SIE breaks are due to wait states.

Perfkit Screen LOCKLOG (FCX265) – Memory Management

```
FCX265   Run 2015/01/30 14:40:49          LOCKLOG

                                          Spin Lock Log, by Time
```

		<---------- Spin Lock Collision Activity ---------->											
		<----- Combined -------->				<----- Exclusive ------->				<------ Shared --------->			
Interval End Time	LockName	Colls /sec	AvgSpin usec	Spin %Busy	CADs /sec	Colls /sec	AvgSpin usec	Spin %Busy	CADs /sec	Colls /sec	AvgSpin usec	Spin %Busy	CADs /sec
>>Mean>>	SRMSLOCK	93695	12.23	114.6	.000	93684	12.23	114.6	.000	10.8	4.605	.005	.000
>>Mean>>	RSA2GLCK	2.6	1.863	.000	.000	2.6	1.863	.000	.000	.0	...	.000	.000
>>Mean>>	HCPTRQLK	5590	.532	.297	.000	5590	.532	.297	.000	.0	...	.000	.000
>>Mean>>	SRMALOCK	.0	...	.000	.000	.0	...	.000	.000	.0	...	.000	.000
>>Mean>>	HCPPGDPL	.0	.										
>>Mean>>	HCPPGDSL	.0	.										
>>Mean>>	HCPPGDTL	.0	.										
>>Mean>>	HCPPGDAL	.0	.										

	<----- Combined -------->			
LockName	Colls /sec	AvgSpin usec	Spin %Busy	CADs /sec
RSA2GLCK	2.6	1.863	.000	.000

- RSA2GLCK is one of many memory management locks (be careful many of them are named similarly). Again this is in wallclock time.

Perfkit Screen VDISKS (FCX147) – VDisk Activity

```
FCX147   Run 2015/01/30 14:40:49          VDISKS
```

Owning	<---- Virtual Disk ---->			P	<--- Data Space, Rate --->				<--- Nr of Pages --->				
Userid	Devno	Blocks	LINKs	IO/s	V	Pgstl	Pgrds	Pgwrt	Resid	IBR	Lockd	XSTOR	DASD
>System<	....	1167k	1	.005	-	.000	.000	.000	22727	17	0	0	0 70
E420301B	0160	200000	1	.000	Y	.000	.000	.000	237	0	0	0	0
E420301B	0161	2133k	1	.000	Y	.000	.000	.000	2180	0	0	0	0
E52AP058	0160	200000	1	.000	Y	.000	.000	.000	1662	0	0	0	0

Owning	<---- Virtual Disk ---->			
Userid	Devno	Blocks	LINKs	IO/s
>System<	....	1167k	1	.005

- VDISKS Report shows information on each Virtual Disk in Storage. It also shows a summary of the all the VDisks. At the far right of he >Sysem< average line, you see the total number of VDisks. In our example, it is 70.
- Taking IO/s x Number = total system VDisk I/Os per second. In our example .005 x 70 = 0.35 /second.
- Benefits from scaling seen with 1000s /second.

Perfkit Screen VSWITCH (FCX240) – Virtual Switch Activity

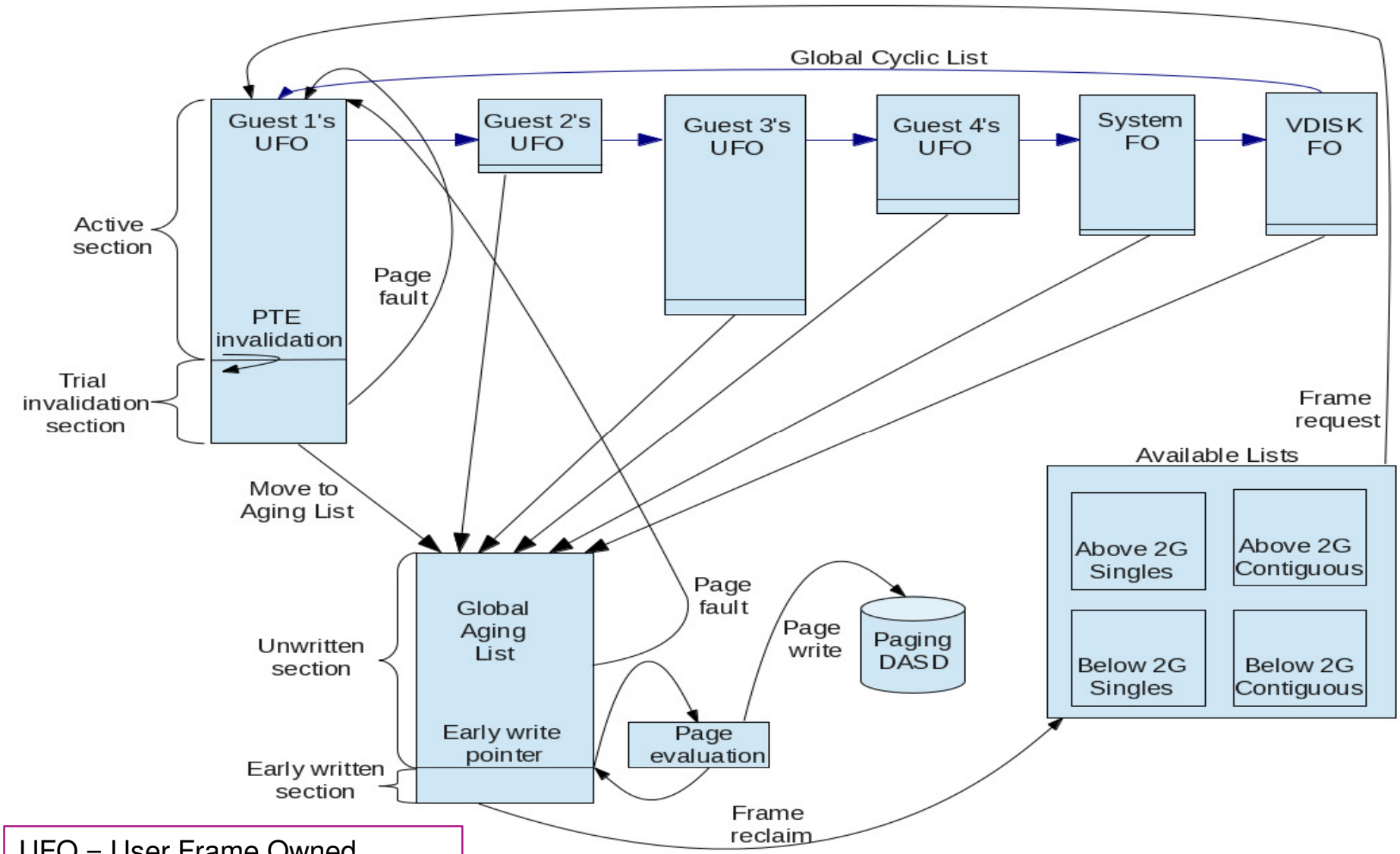
FCX240 Run 2015/01/30 14:40:49 VSWITCH

VSWITCH Activity

			Q	Time	<--- Outbound/s --->			<--- Inbound/s ---->		
			S	Out	Bytes	<--Packets-->		Bytes	<--Packets-->	
Addr	Name	Controlr	V	Sec	T_Byte	T_Pack	T_Disc	R_Byte	R_Pack	R_Disc
>>	System (	26) <<	8	300	606496	1586	0	447004	2321	0
FD80	PRODVZR1	DTCVSW1	8	300	5490k	18624	0	5451k	28846	0
FD9C	PDBVZR2	DTCVSW1	8	300	95201	174	0	139532	203	0
FDB0	PRODVZR1	DTCVSW2	8	300	0	0	0	0	0	0

- There are several VSwitch reports, the one shown is the basic information that gives activity for each individual VSwitch perspective. The amount of data moved (Outbound + Inbound) is what is interesting for one aspect of the improvements, the other is the number of packets.

Large Memory



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	MRSCLADL	event	Add User to Dispatch List	D C
D2	R5	MRSCLDDL	event	Drop User from Dispatch List	D C
D2	R6	MRSCLAEL	event	Add User to Eligible List	C
D2	R8	MRSCLSTP	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 — *page state transition rates*
  - FCX291 UPGACTLG, User Page Activity (benchmarks a user)
  - FCX292 UPGUTL, User Page Utilization Data — *page residency counts*
  - FCX293 UPGUTLLG, User Page Utilization Data (benchmarks a user)
  - FCX294 AVLB2GLG, Available List Data Below 2G, by Time — *available list counts*
  - FCX295 AVLA2GLG, Available List Data Above 2G, by Time
  - FCX296 STEALLOG, Steal Statistics, by Time — *steal algorithm activity*
  - FCX297 AGELLOG, Age List Log, by Time — *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	Instantiations: the rate at which valid memory is being created Instantiated: the amount of valid memory
Relse	Releases: the rate at which memory is being released
Inval	Invalidations: the rate at which demand scan is marking memory invalid as a way to determine whether it is being touched
Reval	Revalidations: the rate at which invalid pages are being made valid because somebody touched them
Ready	Ready reclaims or ready steals: the frame was found and selected for reclaim and had already been prewritten to auxiliary storage
Not Ready	Notready reclaims or notready steals: 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	Private, not referenced: 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	Below 2 GB or Above 2 GB: tells where the real backing frames are in real central
Sing	Singles: free frames surrounded by in-use frames (cannot coalesce)
Cont	Contigs: free frames in strings of two or more
Prot	Protect threshold: 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 ------------->
            Volume Area   Area      Used  <--Page---> <--Spool-->         SSCH
Addr Devtyp Serial Type   Extent       %  P-Rds P-Wrt S-Rds S-Wrt Total  +RSCH
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

           <Page Slots>  <Spool Slots>  <Dump Slots>  <----- Spool Files ----->
Interval     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 Systm 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 x 47 = 39.95

Virtual Machine Memory Usage – FCX292 - UPGUTL

FCX292 Data for 2013/10/15 Interval 10:04:00 - 10:05:00 Monitor Scan

Userid	Data Spaces Owned	WSS	Inst	Resvd	Storage / Resident / Total T_All	Total T<2G	Total T>2G	Locked L<2G	Locked L>2G	Invalid But Resident UFO U<2G	UFO U>2G	PNR P<2G	PNR P>2G	AgeList A<2G	AgeList A>2G	XSTOR	AUX	Base Space Size
>>Mean>>	.9	1807M	2669M	86780	1529M	7588K	1522M	7567	504K	2378	550K	76557	11M	168K	33M	.0	2222M	3315M
DJSLA101	0	5120M	5113M	0	4404M	19M	4384M	0	208K	0	960K	16K	11M	280K	55M	0	3434M	5120M

Userid	Data Spaces Owned	WSS	Inst	Resvd
>>Mean>>	.9	1807M	2669M	86780
DJSLA305	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													
										Invalid But Resident								
	Data Spaces				Total			Locked		UFO		PNR		AgeList				Base 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

Userid	Data Spaces 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.

Reserved? – FCX290 - UPGACT

```
FCX290      Data for 2013/10/15  Interval 10:04:00 - 10:05:00    Monitor Scan
.            . _____     .     .     .     .     .     .     .     .     .     .
               <------------------------------ Storage ------------------------------>
                                                  <------- Movement/s -------->
           Stl <--- Transition/s ----> <-Steal/s->                          <Migrate/s>
Userid      Wt  Inst Relse Inval Reval Ready NoRdy  PGIN PGOUT Reads Write MWrit  Xrel
DJSLA329     1 64853 74069 38571 18978 15292     0     0     0  4506     0     0     0
```

- PGIN/PGOUT – zero due to not using expanded storage
- Reads would be what would be most important in relationship to Reserved.

- Also note rate of Invaliding and Revalidating
  - Reval / Inval = percentage of times trial invalidation leads to page moving back to top of user frame owned list.

- Note: FCX113 UPAGE still produced, but UPGACT is improved

New Toolkit Report FCX295 AVLA2GLG

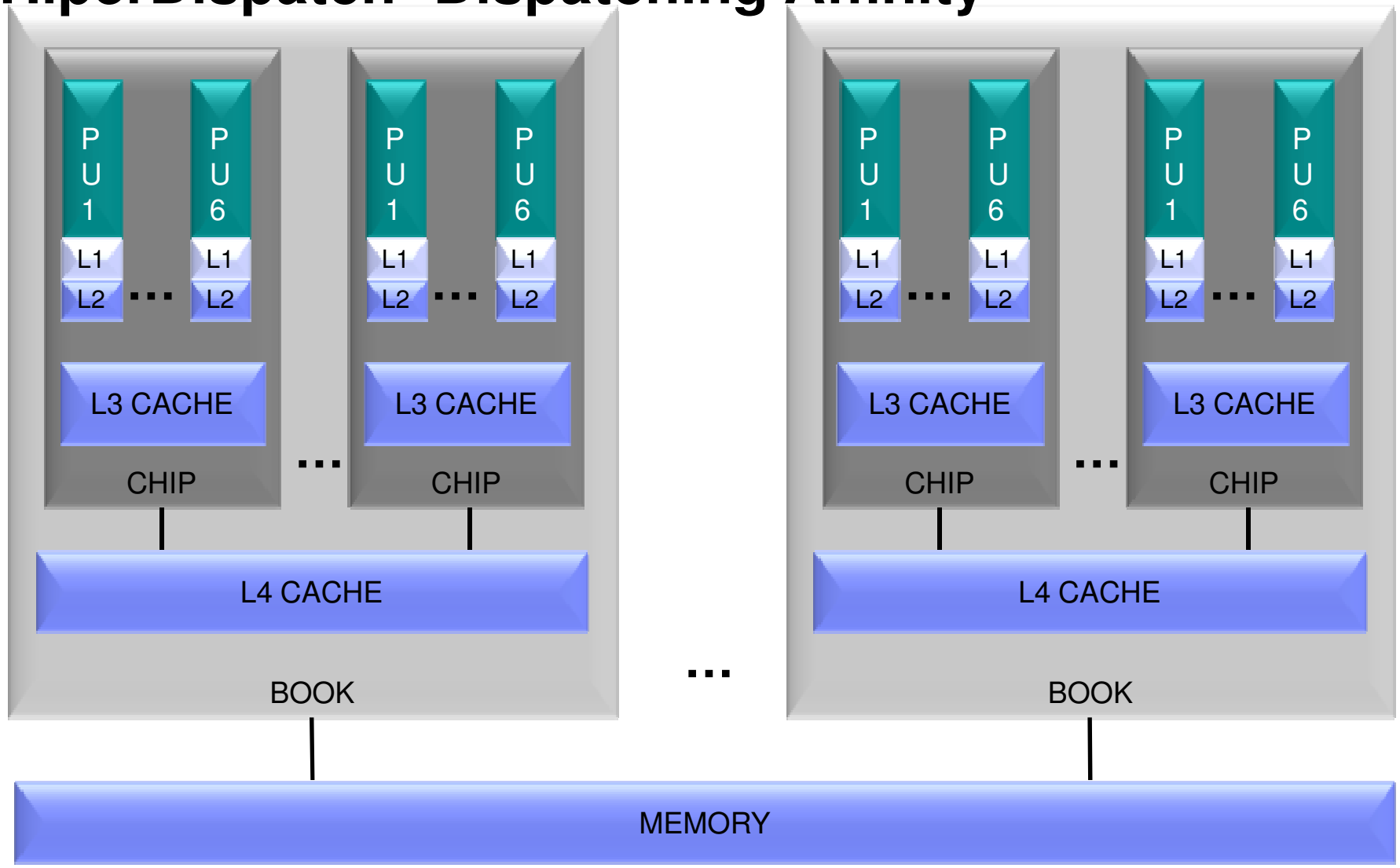
```
FCX295  Run 2013/04/10 07:38:36          AVLA2GLG                                                    Page   25
                                         Available List Data Above 2G, by Time
From 2013/04/09 16:02:10                                                              SYSTEMID
To   2013/04/09 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 End Time	Storage <Available> Sing	Storage <Available> Cont	Storage <Requests/s> Sing	Storage <Requests/s> Cont	Storage <Returns/s> Sing	Storage <Returns/s> Cont	Times <-Empty/s-> Sing	Times <-Empty/s-> Cont	Frame Thresh Sing Low	Frame Thresh <-Contigs-> Low	Frame Thresh <-Contigs-> 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

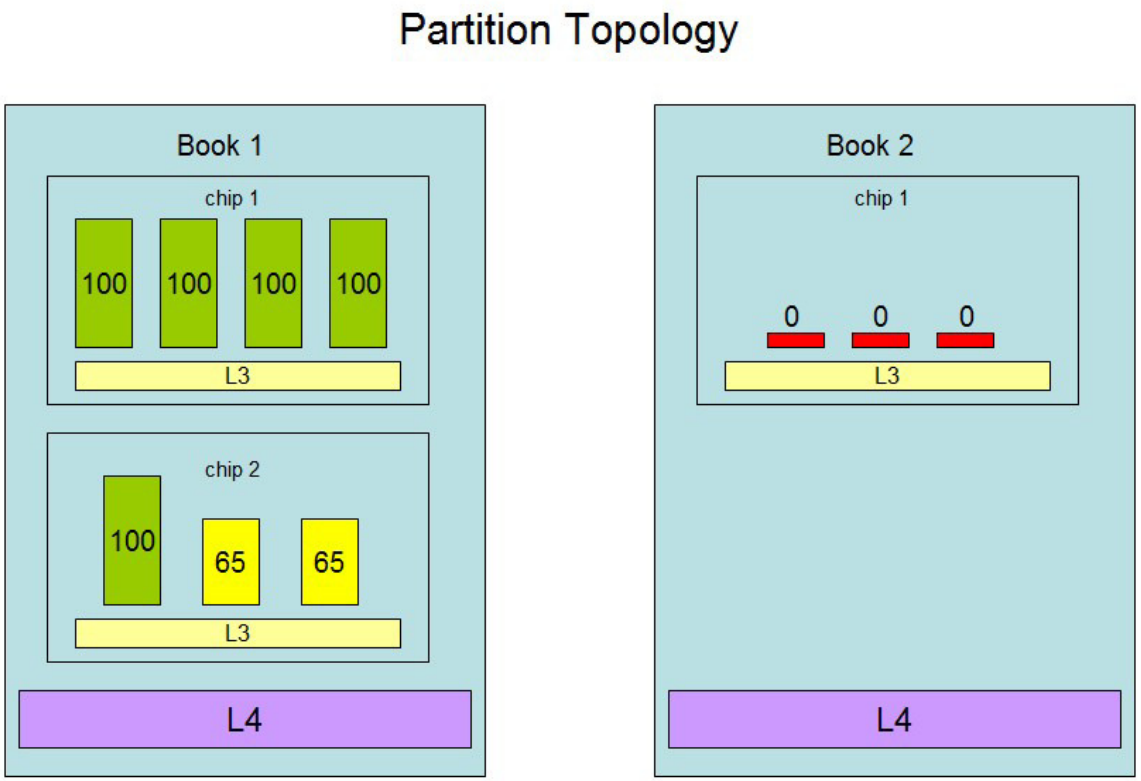
- Look for the new concepts: Singles Contigs Prot
- Amounts are in bytes, suffixed. Not page counts!
- FCX254 AVAILLOG is no longer produced.

HiperDispatch

HiperDispatch- Dispatching Affinity



IBM z Systems: The Partition Knows Its Placement



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?

HiperDispatch 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

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
Vl: 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:

Interval End Time	CPU	Type	PPD	Ent.	DVID	Pct Park Time	<--- Percent Busy ----> 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                                                                  Page    6
                                         Processor Unit Configuration log
From 2013/02/19 11:49:52                                                                                GFCM0107
To   2013/02/19 11:56:10                                                                                CPU 2817-744   SN B6D85
For    378 Secs 00:06:18                 Result of GFCM0107 Run                                         z/VM   V.6.3.0 SLU 0000
```

					Type			<--- Last --->			<--- Next --->						
Date	Time	Type	OnL	Entitl	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 changed																
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 changed																
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 changed																
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 changed																
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	FFFF0000_00000000
02/19	11:50:20	* CPU Park/Unpark State changed																
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	FFFFFE00_00000000
02/19	11:50:22	* CPU Park/Unpark State changed																
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	FFFFFF00_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	FFFFFF00_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	FFFFFF00_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	FFFFFF00_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	FFFFFF00_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	FFFFFF00_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	FFFFFF00_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 Partition   LPU   LPAR          <LPU Total,%> LPU
  Type Name        Num Weight Entlment   Busy Excess 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 -
  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          PHYSLOG
                                         Real CPU Utilization Log
From 2013/02/19 11:49:58
To   2013/02/19 11:56:10
For    372 Secs 00:06:12                 Result of GFCM0107 Run
```

Interval End Time	Type	<PU Num> Conf	Ded	Total Weight	%LgclP	%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

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

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

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

Please remember to fill out an evaluation. Thanks!