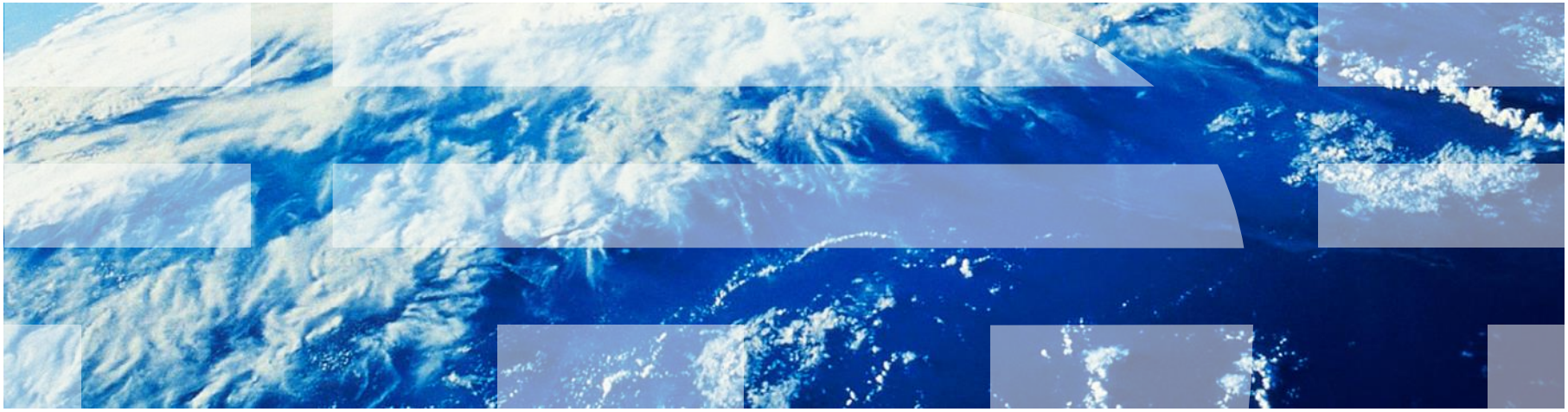
Is My z/VM System Full?

Revision 2021-08-18.1

Brian K. Wade, Ph.D.
IBM z/VM Development, Endicott, NY
bkw@us.ibm.com



Abstract

From time to time customers ask us whether their z/VM systems are "full." In other words, can the system absorb any more work? Does it have any capacity remaining? Could I do more with my system without adding hardware?

Answering this question requires looking at the capacity and utilization of each of a number of different physical resource types. For example, how much processor capacity is there, and how utilized is it? The same questions can be asked about memory, I/O, networking, and ISFC links.

In this presentation we use z/VM Performance Toolkit ("Perfkit") screens to explore systems' capacities and utilizations in a number of different categories of resource. Sample screen shots illustrate which Perfkit screens to consult, which columns are especially helpful, and what values in those columns represent "full."

Agenda

- What does “full” mean?
- Full on processors
- Full on memory
- Full on I/O
- Full on networking
- Full on CTCs and ISFC (SSI) links
- Summary

What Does “Full” Mean?

Some Various Meanings of “Full”

- Cores and processors
  - Is the physical machine full?
  - Is a physical type-pool full?
  - Is my partition full?
  - Is a guest full?
- Memory
  - Is central memory full?
  - Is the paging space full?
  - Are the paging devices keeping up?
  - Are we out of SXS space?
  - Is a particular guest full?
- I/O
  - Is an I/O device full?
  - Is a chpid full?
  - Is dump space full?
  - Is spool space full?
  - Is T-disk space full?
- Networking
  - Is a network adapter full?
  - Is the z/VM TCP/IP stack CPU-bound?
  - Are the SSL servers CPU-bound?
- ISFC
  - Is a CTC full?
  - Is an ISFC logical link full?

Are My Processors Full?

Vocabulary

A non-SMT partition has **logical processors**.

PR/SM dispatches them **alone** on physical cores.

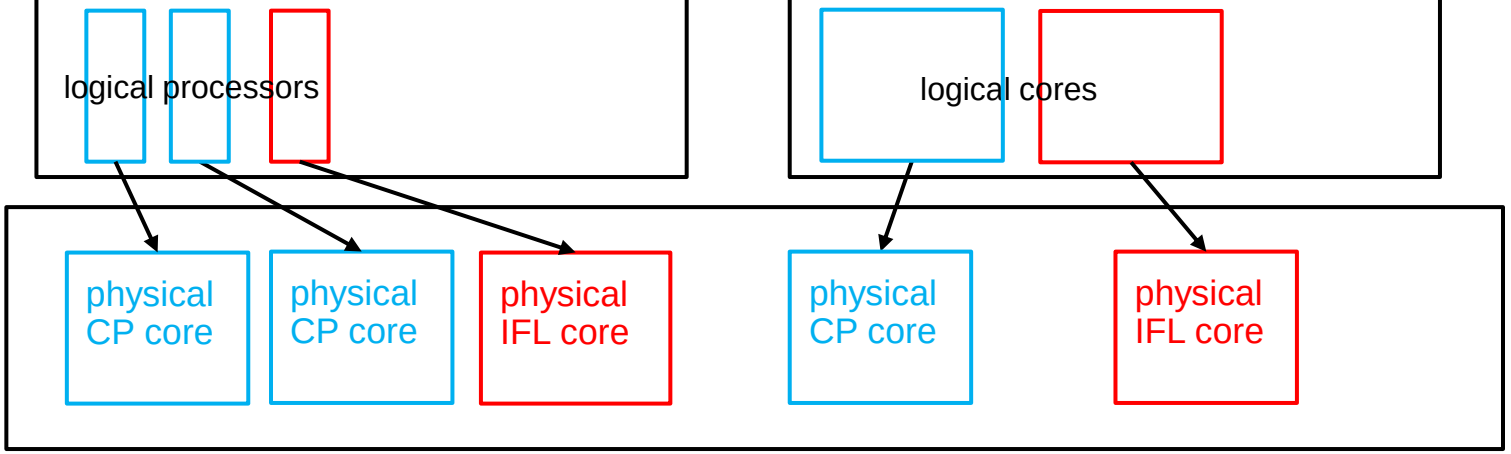
The other half of the physical core goes unused.

An SMT-enabled partition has **logical cores**.

At SMT-1 each logical core has one logical processor.

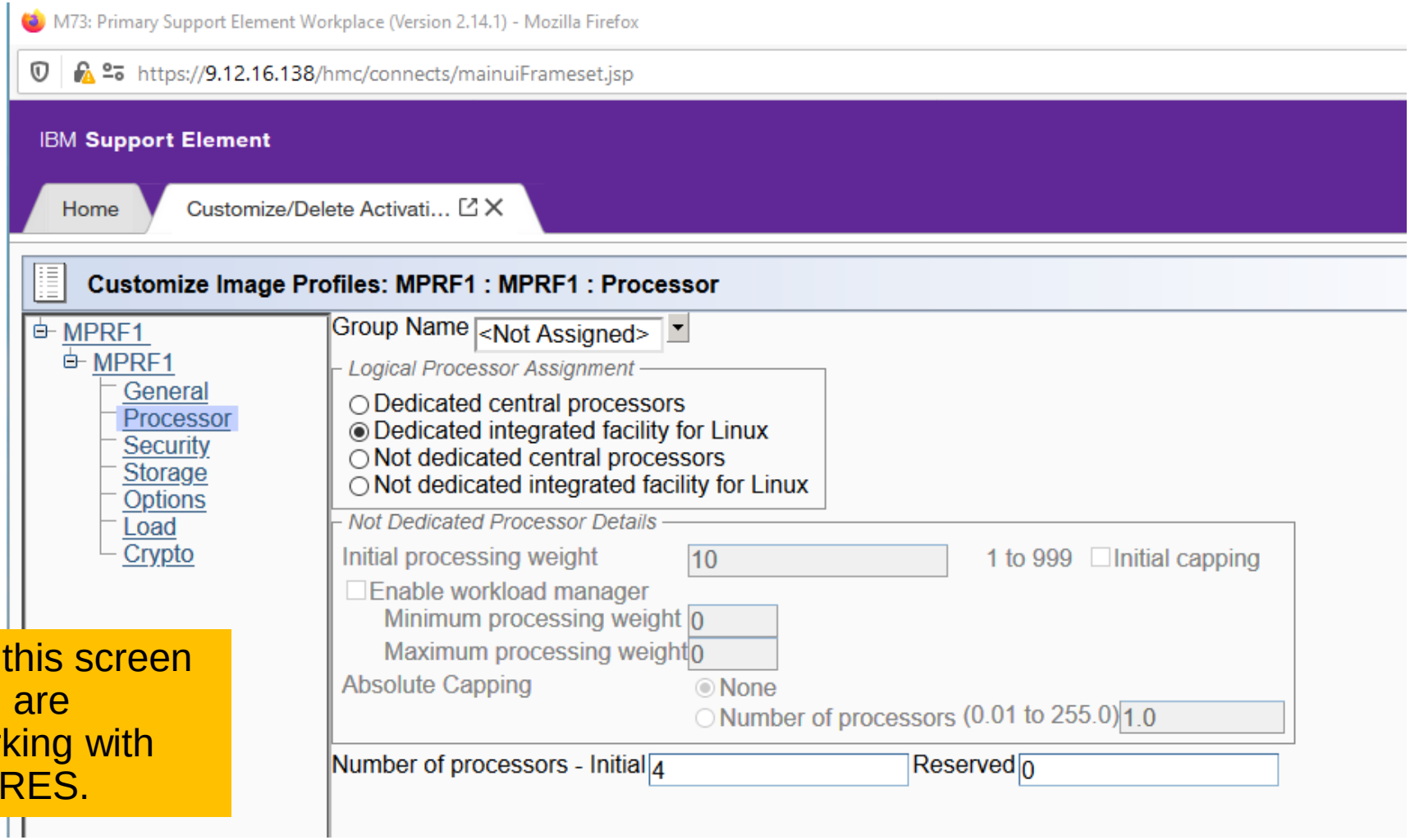
At SMT-2, in z/VM the logical IFL cores have two logical processors and all other types of logical cores have one logical processor.

PR/SM dispatches logical cores on physical cores.



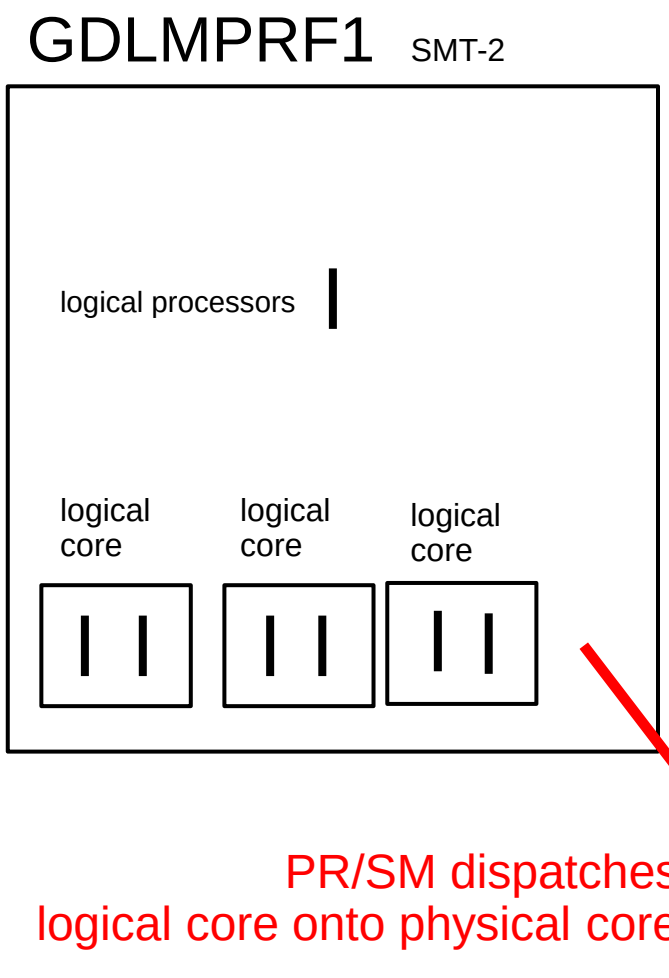
The machine has **physical cores.**
Each physical core contains two **physical processors.**

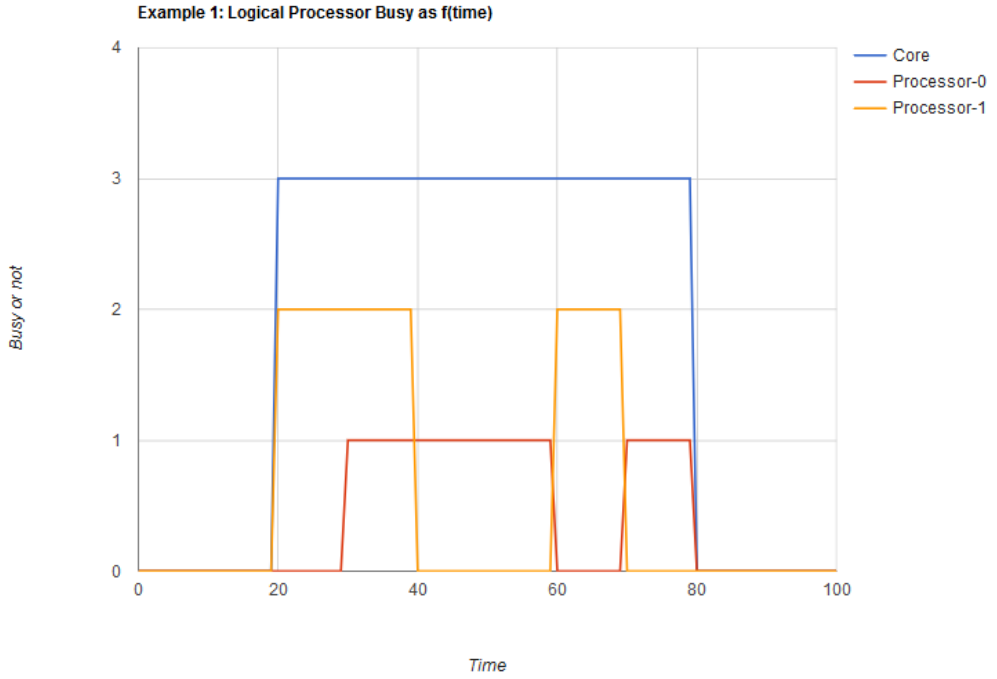
Do Not Let The Support Element (SE) Confuse You



On this screen you are working with CORES.

Processor: "Core Dispatch" and Core-Busy vs. Processor-Busy





(The heights are meaningless; it's just 0=idle or >0=busy)

During the time a logical core is dispatched, its logical processors can run and wait independently. This concept is very important in SMT-2 deployments. Check it out: http://www.vm.ibm.com/perf/tips/smtutil.html

Processor: So This Leads Us to Some Definitions

- **Core utilization**
  - For a physical core: percent of elapsed time a logical core is dispatched upon it
  - For a logical core: percent of elapsed time it is dispatched upon a physical core
- **Processor utilization**
  - Percent of elapsed time the processor is not in a wait PSW
- **You can see core utilization and processor utilization are two completely different phenomena**
- **Articles I very strongly suggest you read**
  - SMT vocabulary: http://www.vm.ibm.com/perf/tips/smtvocab.html
  - SMT utilization: http://www.vm.ibm.com/perf/tips/smtutil.html

Different Ways We Run Out of “Processor”

Situation	Core or Processor?	Primary indicators	Secondary indicators
The CPC has no more power to give.	core	FCX302 PHYSLOG	FCX239 PROCSUM Diag 9C FCX304 PRCLOG %Susp FCX299 PUCFGLOG
Our partition is being dispatched by PR/SM as much as it can be.	core	FCX126 LPAR	
All our logical processors are full.	processor	FCX304 PRCLOG	FCX104 PRIVOP Diag x’44’ FCX104 PRIVOP Diag x’9C’ FCX164 USTATLOG FCX114 USTAT FCX315 USTMPLOG FCX225 SYSSUMLG
A guest is running completely busy.	processor	FCX112 USER FCX162 USERLOG FCX288 USRMPLOG	

Processor: Is the CPC Full? (Core-busy)

```
1FCX302  Run 2019/03/20 13:15:32          PHYSLOG
                                          Real Core Utilization Log
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
```

Interval End Time	Type	<PhCore> Conf	Ded	Shrd Log.	Total Weight	%LgclC	%Ovrhd	LCoT/L	%LPmgt	%Total	TypeT/L
>>Mean>>	CP	80	0	467	210	1963.8	54.650	1.028	58.987	2077.4	1.058
>>Mean>>	IFL	81	0	417	350	7212.3	33.116	1.005	56.743	7302.2	1.012
>>Mean>>	ICF	4	0	1	10	.044	.010	1.235	2.696	2.750	62.935
>>Mean>>	ZIIP	4	0	1	10	.007	.006	1.889	.012	.024	3.629
>>Mean>>	>Sum	170	0	886	580	9176.2	87.782	1.010	118.44	9382.4	1.022

FCX302 PHYSLOG shows core-busy in each physical type pool.
"Full" = (%Total / 100) is close to Conf.
In this example, the IFL pool is closing in on full.

Processor: Dispatch of Our Own Partition (Core-busy)

```
1FCX126  Run 2019/03/20 13:15:32          LPAR
                                          Logical Partition Activity
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
 ____________________________________________________________________

  LPAR Data, Collected in Partition MSTL1

  Processor type and model    : 3906-M05
  Nr. of configured partitions:      30
  Nr. of physical cores       :     170
  Dispatch interval (msec)    : dynamic

<I have removed a lot of stuff>

Partition Nr.  Upid #Core Weight Wait-C Cap %Load CID %Busy %Ovhd --- --- --- Type
MSTL1       7    29    80     60     NO  NO  37.1   0  47.6    .2 --- --- --- IFL
                              60         NO   ...   1  48.2    .2 --- --- --- IFL
                              60         NO   ...   2  48.5    .1 --- --- --- IFL
                              60         NO   ...   3  48.4    .1 --- --- --- IFL
                              60         NO   ...   4  48.4    .1 --- --- --- IFL
                              60         NO   ...   5  48.3    .1 --- --- --- IFL
                              60         NO   ...   6  97.1    .2 --- --- --- IFL
                              60         NO   ...   7  48.3    .1 --- --- --- IFL
```

FCX126 LPAR shows the core-busy of each of our logical cores.
"Full" = %Busy+%Ovhd is approaching 100% for each logical core.
Exception: dedicated LPARs: unparked cores 100%, parked cores < 100%.

Reports: Core-Busy Reports vs. Processor-Busy Reports

Core-Busy Reports

- FCX302 PHYSLOG
- FCX126 LPAR
- FCX202 LPARLOG
- FCX306 LSHARACT
- FCX299 PUCFGLOG

Processor-Busy Reports

- FCX144 PROCLOG
- FCX304 PRCLOG
- FCX225 SYSSUMLG
- FCX112 USER
- FCX162 USERLOG
- FCX288 USRMPLOG

Processor: Logical Processor Busy (Processor-busy)

```
1FCX304  Run 2019/03/20 13:15:32          PRCLOG
                                          Processor Activity, by Time
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
```

Interval End Time	CP U	Type	PPD	Ent.	DVID	Pct Park Time	%Susp	<--- Percent Busy ----> Total	User	Syst	Emul
>>Mean>>	0	IFL	Vh	100	0000	0	1.6	94.6	82.7	11.9	73.3
>>Mean>>	1	IFL	Vh	100	0001	0	1.0	96.0	87.0	9.0	79.7
>>Mean>>	2	IFL	Vh	100	0002	0	.9	96.7	88.0	8.7	82.2
>>Mean>>	3	IFL	Vh	100	0003	0	.9	96.5	87.8	8.7	81.6

FCX304 PRCLOG shows the processor-busy of each of our logical processors.
"Full" = "Total" approaching 100% for each logical processor.

This is a good time to explain the columns:

Emul	Percent pure guest time
User	Emul + percent induced time in the Control Program (e.g., simulation)
Syst	Percent non-induced Control Program time (e.g., cutting monitor records)
Total	User + Syst
Park	Percent spent parked
Susp	Percent spent suspended (PR/SM simulation or PR/SM dispatch contention)

Breaking Down High FCX304 PRCLOG Busy

Component	What a High Value Means
Emul	The guests are really busy, running their own work. Are they looping?
User – Emul	CP is doing a lot of simulation for the guests. Are the guests heavy on VSWITCH activity? I/O activity? Other stuff CP has to simulate for guests?
Syst	CP is busy doing overhead work it can't charge to specific guests. Is memory really tight? Is z/VM paging a lot?
Susp	Is CP doing stuff PR/SM has to simulate? Host Diag x'9C's, for example? OR, is PHYSLOG showing high physical core utilization? OR, is there a VH/VM/VL configuration problem?
Park	We didn't want to use the core on which that processor resides, so we parked the core.

Processor: Is a Guest Busy? (Processor-busy)

```
1FCX112  Run 2019/03/20 13:15:32          USER
                                          General User Resource Utilization
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
```

Userid	%CPU	TCPU	VCPU	T/V Ratio	CPU Load <-Seconds-> Total	Virtual IO/s DASD	Avoid	Diag 98	UR	Pg/s	User Status
WDB00002	175	2102	2095	1.00	.0	.0	.0	.0	.0	296	EME,CL3,DISP
WDB00001	88.4	1061	1056	1.00	.0	.0	.0	.0	.0	3.4	EME,CL3,DISP
LX00002	88.2	1058	1057	1.00	.1	.0	.0	.0	.0	1.0	EME,CL3,DISP
LX00004	88.1	1057	1044	1.01	.1	.0	.0	.0	.0	1.1	EME,CL3,DISP
LXI01629	87.9	1055	1053	1.00	.0	.0	.0	.0	.0	5.0	EME,CL0,DISP

FCX112 USER shows percent-busy by guest. A value of 100% = one processor's worth of power.

This is a good time to explain the columns:

%CPU	Total percent-busy for the guest (100 = one processor's worth of power)
TCPU	Total CPU seconds consumed (pure-guest plus induced-CP)
VCPU	Virtual CPU seconds consumed (pure-guest only)
T/V ratio	TCPU / VCPU

When could %CPU > 100?

Secondary Indicators We Might Be Full

Column	What It Means
FCX104 PRIVOP diag x'44' diag x'9C'	MP guests are sensing their virtual processors are having difficulty running. There could be lots of reasons: no CPU power available, etc.
FCX239 PROCSUM diag x'9C'	CP is sensing its logical processors are having difficulty running. There could be lots of reasons: no CPU power available, etc.
FCX299 PUCFGLOG LCei > UpCap	CP's workload ceiling estimate is higher than its available power estimate. You probably see parking. The CPC probably does not have enough physical core capacity. OR, there is something wrong with the weight-vs.-cores relationship.
FCX114 USTAT %CPU >> 0	Guests are ready to run but CP is not running them. Is PRCLOG reporting high total busy?
FCX304 PRCLOG %Susp >> 0	CP is sensing its logical processors are having a hard time running. Maybe the type-pool is overworked. OR, maybe CP is doing a lot of things PR/SM has to simulate.

Is My Memory Full?

Different Ways We Run Out of Memory

Situation	Primary indicators	Secondary indicators
We run low on storage < 2 GB.	FCX294 AVLB2GLG	FCX143 PAGELOG
We run low on storage > 2 GB.	FCX295 AVLA2GLG	FCX143 PAGELOG
We run out of System Execution Space (SXS).	FCX261 SXSAVAIL	
We run out of space for page management blocks.	FCX134 DSPACESH	
We run out of memory to use for MDC.	FCX138 MDCACHE FCX178 MDCSTOR	
We run out of paging space.	FCX109 DEVICE CPOWNED	
The paging DASD just can't do any more I/O.	FCX109 DEVICE CPOWNED FCX329 VOLUME FCX327 HPALIAS	
A guest fills its primary address space.	FCX292 UPGUTL FCX290 UPGACT	FCX147 VDISK FCX134 DSPACESH

Memory: Full Below 2 GB Bar?

```
1FCX294  Run 2019/03/20 13:15:32          AVLB2GLG
                                          Available List Data
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D
```

	<---------- Storage ---------->						<--Times-->	
Interval	<Available>		<Requests/s>		<Returns/s>		<-Empty/s->	
End Time	Sing	Cont	Sing	Cont	Sing	Cont	Sing	Cont
>>Mean>>	375K	601K	107K	3291	106K	17033	1.1	.0
06:53:16	240K	780K	116K	5052	121K	23142	.0	.0
06:54:16	188K	832K	122K	6281	126K	27648	.0	.0
06:55:16	272K	748K	86767	4642	97348	17545	.0	.0
06:56:16	304K	716K	93594	3686	98K	16862	.0	.0
06:57:16	412K	608K	115K	4164	124K	17749	.0	.0
06:58:16	376K	644K	109K	2662	108K	14882	.0	.0

These values are AMOUNTS. “601K” = 601 KB free. (aka 150 4-KB frames)

“Times Empty” = request rate at which CP wanted memory < 2 GB and there wasn’t any.

SXS pages sometimes need to be backed with < 2 GB frames.
If < 2 GB fills, operations needing SXS < 2 GB will not work.

Memory: Full Above 2 GB Bar?

```
1FCX295  Run 2019/03/20 13:15:32          AVLA2GLG
                                          Available List Data
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D
 ____________________________________________________________________

           <------------ Storage -------------> <--Times-->
 Interval  <Available> <Requests/s> <Returns/s> <-Empty/s->
 End Time   Sing  Cont  Sing   Cont  Sing  Cont  Sing  Cont
 >>Mean>>  2422K 4388K  102M   841K  103M 4502K  70.5    .0
 06:53:16  1936K 3992K  102M   872K  104M 3980K  91.2    .0
 06:54:16   896K 4440K  100M   953K  101M 5731K  68.9    .0
 06:55:16   320K 6836K   85M   918K   85M 7394K  46.6    .0
 06:56:16  1940K 4912K   91M  1085K   91M 8642K  39.4    .0
 06:57:16  1684K 5384K  107M   972K  108M 5173K  44.1    .0
 06:58:16  2824K 3760K   95M   807K   97M 3483K  76.4    .0
 06:59:16  2920K 4004K  106M   864K  107M 3702K  40.9    .0
```

These values are AMOUNTS. “4388K” = 4388 KB free. (aka 1097 4-KB frames)

“Times Empty” = request rate at which CP wanted memory > 2 GB and there wasn’t any.

NB: How much central memory is in the partition? FCX103 STORAGE, “Total available”.

Memory: No Room for Page Management Blocks?

```
1FCX134  Run 2019/03/20 13:15:32          DSPACESH                                                    Page   17
                                          Shared Data Spaces Paging Activity
 From 2019/03/11 06:52:16                                                                  2U0C021D
 To   2019/03/11 07:12:16                                                                  CPU 3906-M05   SN DA1F7
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run                           z/VM   V.7.1.0 SLU 0000
```

Owning Userid	Data Space Name	Users Permt	Rate per Sec. Pgstl	Pgrds	Pgwrt	X-rds	X-wrt	X-mig	Number of Pages Total	<--Resid--> Resid	R<2GB	<-Locked--> Lock	L<2GB	<-Aliases-> Count	Lockd	XSTOR	DASD
>System<	139	0	.110	.860	.220	.000	.000	.000	1011k	27423	1	0	0	0	0	0	41568
SYSTEM	FULL$TRACK$CACHE$1	0	.000	.000	.000	.000	.000	.000	524k	0	0	0	0	0	0	0	0
SYSTEM	ISFCDATASPACE	0	.000	.000	.000	.000	.000	.000	524k	2	0	0	0	0	0	0	15
SYSTEM	PTRM000A	0	.097	6.678	.193	.000	.000	.000	1049k	110k	0	0	0	0	0	0	145k
SYSTEM	PTRM000B	0	.598	1.492	1.197	.000	.000	.000	1049k	94833	44	0	0	0	0	0	277k
SYSTEM	PTRM000C	0	.043	.392	.085	.000	.000	.000	1049k	10388	0	0	0	0	0	0	75720

A “page management block” (PGMBK) is a CP-owned data structure involved in DAT mappings.
PGMBKs reside in address spaces called “PTRM spaces”.
CP creates 128 PTRM address spaces for its use. Each one is 2 GB large (524288 pages).
You can see PTRM stats on FCX134 DSPACESH.
We did relief on this a while back. The probability of running out of PTRM address space is now very small.

Memory: System Execution Space (SXS) Full?

```
1FCX261  Run 2019/03/20 13:15:32          SXSAVAIL
                                          System Execution Space Page Queues Manag
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
```

Interval End Time	<-- Backed <2GB Page Queue --> Avail Pages	<-Pages/s--> Taken	Return	<Preferred> Used	Empty	<-- Backed >2GB Page Queue --> Avail Pages	<-Pages/s--> Taken	Return	<Preferred> Used	Empty	<---- Avail Pages
>>Mean>>	14	8.634	8.614	8.633	1.113	5	10.63	10.63	10.61	8.559	290391
06:53:16	29	6.367	6.400	6.367	.000	1	10.70	10.70	10.70	9.717	291009
06:54:16	24	6.517	6.433	6.517	.000	1	10.72	10.72	10.72	9.000	291298
06:55:16	26	6.267	6.300	6.267	.000	1	10.22	10.22	10.22	2.300	291085

There are 524,288 ADDRESSABLE PAGES in the SXS. (It's a 2 GB address space.)

These values are in PAGES.

CP backs the SXS with a mix of frames < 2 GB and frames > 2 GB.

What we are really worried about here is "Avail Pages". Those are the SXS pages not yet backed.

This system has plenty of SXS.

Memory: Minidisk Cache (MDC) Full?

```
1FCX178  Run 2019/03/20 13:15:32          MDCSTOR
                                          Minidisk Cache Storage Usage
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
 ____________________________________________________________________________

            <----------------- Main Storage Frames ----------------->
 Interval          <--Actual--->     Min     Max   Page     Steal
 End Time    Ideal    <2GB    >2GB     Set     Set Del/s Invokd/s    Bias
 >>Mean>>   130224       0       2       0  1049k    27    19.49    1.00
 06:53:16       12       0       0       0  1049k    29    19.00    1.00
 06:54:16       12       0      12       0  1049k    24    16.08    1.00
 06:55:16       12       0       0       0  1049k    27    17.15    1.00
 06:56:16       12       0       0       0  1049k    27    17.03    1.00
 06:57:16       12       0       0       0  1049k    23    15.37    1.00
```

FCX178 MDCSTOR discusses memory being used for MDC.

The book z/VM CP Planning and Admin tells how to plan the memory being used for MDC.

To control the storage used by MDC, use the command CP SET MDCACHE STORAGE.

FCX108 DEVICE shows, per-device, whether MDC is being effective.

FCX112 USER shows, per-user, whether MDC is being effective.

Memory: Is z/VM Paging?

```
1FCX225  Run 2019/03/20 13:15:32          SYSSUMLG
                                          System Performance Summary by Time
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
 ------------------------------------------------------------------------------

            <------- CPU --------> <--Users--> <---I/O---> <Stg> <-Paging-->
                  <--Ratio-->                    SSCH  DASD Users <-Rate/s-->
 Interval    Pct                 On-  Log-       +RSCH  Resp    in PGIN+ Read+
 End Time   Busy   T/V  User line   ged Activ     /s  msec Elist PGOUT Write
 >>Mean>>   87.1  1.17 .9266 72.1   297   275  3171   9.1    .0    .0 37061
 06:53:16   91.1  1.06 .9662 64.0   297   274  2788   9.0    .0    .0 38476
 06:54:16   91.4  1.07 .9656 64.0   297   276  2896   9.8    .0    .0 38149
 06:55:16   90.9  1.07 .9676 64.0   297   276  2527   9.0    .0    .0 31193
```

There are several places to see z/VM paging activity.
- FCX225 SYSSUMLG shows overall paging rate as f(time).
- FCX143 PAGELOG shows a breakdown of kinds of paging, as f(time).
- FCX109 DEVICE CPOWNED shows paging rate per-paging-volume.

There are also several places to see the activity of the memory manager.
- FCX296 STEALLOG shows percent-CPU-busy spent doing frame stealing.
- FCX297 AGELLOG shows activity done managing the global available list.

Memory: Paging Space Full?

```
1FCX109  Run 2019/03/20 13:15:32          DEVICE   CPOWNED
                                          Load and Performance of CP Owned Disks
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
 ------------------------------------------------------------------------------

  Page / SPOOL Allocation Summary
  PAGE slots available      935899k          SPOOL slots available       71140k
  PAGE slot utilization          57%         SPOOL slot utilization          14%
  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
  D227 3390-A TMD227 PAGE    11793600   57   247.4 220.8   ...   ... 468.2  35.5
  9756 3390-3 S31P03 PAGE      1-3338  100    17.6  16.3   ...   ...  33.9   3.0
  9762 3390-3 S31P04 PAGE      1-3338  100    18.4  17.5   ...   ...  35.9   3.5
  D210 3390-A TMD210 PAGE    11793600   57   248.5 220.1   ...   ... 468.6  35.4
  D20B 3390-A TMD20B PAGE    11793600   57   246.9 221.8   ...   ... 468.7  39.1
```

FCX109 DEVICE CPOWNED discusses how full the paging space is.

The book z/VM CP Planning and Admin tells how to plan how much paging space you need.

Memory: Paging DASD Just Can't Do Any More, 1 of 3

```
1FCX109  Run 2020/04/14 11:17:08         DEVICE   CPOWNED                                                    Page   35
                                         Load and Performance of CP Owned Disks
 From 2020/04/08 09:34:29                                                                                    2U0C913C
 To   2020/04/08 10:01:29                                                                                    CPU 8561-T01   SN DEE58
 For   1620 Secs 00:27:00                Result of 2U0C913C Run                                              z/VM   V.7.2.0 SLU 0000
```

```
 Page / SPOOL Allocation Summary
 PAGE slots available      747868k          SPOOL slots available       39847k
 PAGE slot utilization          10%         SPOOL slot utilization          53%
 T-Disk space avail. (MB)  .......         DUMP slots available       6659940
 T-Disk space utilization      ...%        DUMP slot utilization           66%
```

```
 .    .           .                        .     .     .     .     .     .     . _____    .     .     .     . .
 < Device Descr. ->                       <------------- Rate/s ------------>   I/O        Serv MLOAD Block %Used I
             Volume Area   Area      Used  <--Page---> <--Spool-->        SSCH Inter Queue  Time  Resp  Page   for O
 Addr Devtyp Serial Type   Extent       %  P-Rds P-Wrt S-Rds S-Wrt Total +RSCH feres Lngth /Page  Time  Size Alloc M
 AC24 3390-A JMAC24 PAGE    23239440    9   3633  2964   ...   ...  6596 965.4     0 99.96   1.7 110.5    14   100 T
 AC25 3390-A JMAC25 PAGE    23239440    9   3636  2961   ...   ...  6597 961.5     0 85.54   1.4  82.9    14   100 T
 AC1C 3390-A JMAC1C PAGE    23239440    9   3344  2968   ...   ...  6312 956.4     0 82.11   1.7  51.6    14   100 T
 AC1D 3390-A JMAC1D PAGE    23239440    9   3363  2962   ...   ...  6325 956.6     0 81.04   1.7  55.8    14   100 T
 AC35 3390-A JMAC35 PAGE    23239440    9   3240  2963   ...   ...  6203  1098     0 74.54   1.6  26.9    15   100 T
 AC28 3390-A JMAC28 PAGE    23239440    9   3629  2970   ...   ...  6598 980.0     0 73.32   1.5  42.2    14   100 T
```

Here we are looking at column "Queue Lngth".
This is the number of pages (PIOBKs) either waiting or in-flight, read or write.
When QL>0, paging I/O is standing in line.

Let's investigate.

Memory: Paging DASD Just Can't Do Any More, 2 of 3

```
1FCX329  Run 2020/04/21 07:57:36          VOLUME
                                          DASD Volume Performance Report
 From 2020/04/08 09:34:29
 To   2020/04/08 10:01:32
 For   1622 Secs 00:27:02                 Result of 2U0C913C Run
```

<---- Volume ----> Addr	Label	WA	LCU SSID	T	Mdisk Links	I/O	AIOR	<----- Volume Times (msec) ------> Pend	Disc	Conn	Serv	Wait	Resp	CUWt	Req. Qued	Pct Busy	<-Serv---> Base	Alias
AC1C	JMAC1C	CP	0803	H	0	957	1.77	1.01	.135	1.72	2.87	.260	3.13	1.23	.29	274	1.85	3.44
AC1D	JMAC1D	CP	0803	H	0	957	1.78	1.00	.136	1.70	2.84	.254	3.10	1.21	.30	272	1.86	3.39
AC1E	JMAC1E	CP	0803	H	0	944	1.78	.989	.139	1.75	2.88	.245	3.13	1.26	.33	272	1.87	3.45
AC1F	JMAC1F	CP	0803	H	0	944	1.82	1.00	.140	1.79	2.93	.252	3.19	1.29	.31	277	1.92	3.49
AC20	JMAC20	CP	0803	H	0	950	1.83	1.02	.144	1.80	2.96	.260	3.22	1.30	.29	281	1.93	3.52
AC21	JMAC21	CP	0803	H	0	948	1.87	1.01	.142	1.84	3.00	.256	3.25	1.34	.30	284	1.97	3.55
AC22	JMAC22	CP	0803	H	0	949	1.82	.992	.144	1.79	2.92	.245	3.17	1.29	.30	277	1.92	3.47

FCX329 VOLUME is a recently new Perfkit report.
It tells us how the *volume* is doing, accounting for contributions by aliases.

Here we are looking at column Pct Busy.
We see Pct Busy > 100%.
We are depending upon HyperPAV aliases to accomplish our paging I/O.

Let's check FCX327 HPALIAS for the affected LCU (here, SSID 0803).

Memory: Paging DASD Just Can't Do Any More, 3 of 3

```
1FCX327  Run 2020/04/21 07:57:36          HPALIAS
                                          HyperPAV Alias Activity Data
 From 2020/04/08 09:33:39
 To   2020/04/08 10:01:32
 For   1673 Secs 00:27:53                 Result of 2U0C913C Run
```

SSID	Pool	Base	Alias	SYSTEM	Tries	Fails	Low	WhenLow	High	WhenHigh
0205	2	16	6	0	0	0	...	.....	...	.....
062B	3	100	156	0	0	0	...	.....	...	.....
0633	4	19	156	0	0	0	...	.....	...	.....
063C	1	60	164	164	0	0	...	.....	...	.....
0803	0	61	195	195	32715.0	15551.6	0	2020/04/08,09:58:32	193	2020/04/08,09:51:30

FCX327 HPALIAS is a recently new Perfkit report.
It tells us how the HyperPAV alias pools, per-LCU resources, are doing.

Here we are looking at column Fails.
For the affected LCU we see Fails > 0.
We want more help from our HyperPAV aliases than we are getting.
Tip: check FCX108 DEVICE for the HyperPAV alias devices. 100% busy!

This explains why FCX109 DEVICE CPOWNED shows queueing.

Memory: Is z/VM Demand Scan Completely Busy?

```
1FCX296  Run 2019/03/20 13:15:32          STEALLOG
                                          Frame Steal Statistics, by Ti
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
 ------------------------------------------------------------------------

             Pct <-------------------- Storage/s --------------------->
  Interval  Time Total Write <--User---> <-Shared--> <Pvt Vdisk>  AgeL
  End Time  Actv Stoln OnDmd Inval Reval Inval Reval Inval Reval Reval
  >>Mean>>  81.9   76M 19316  596M  160M  275K  122K    .0    .0  420M
  06:53:16  77.1   80M 38025  602M  168M  290K  143K    .0    .0  419M
  06:54:16  79.7   78M  9284  615M  170M  285K  136K    .0    .0  430M
  06:55:16  76.3   64M    .0  560M  146M  234K  105K    .0    .0  404M
  06:56:16  79.3   69M 21436  568M  152M  201K 81784    .0    .0  404M
  06:57:16  89.0   86M 16248  611M  174M  288K  145K    .0    .0  416M
```

“Demand scan” is the process of running the CP real memory manager.
Demand scan can run on only one logical processor at a time.
Thus when it gets to 100% busy (one processor’s worth of power), that’s all it can do.

Memory: Is z/VM Paging a Specific Guest?

```
1FCX290  Run 2019/03/20 13:15:32          UPGACT
                                          User Page Activity
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
```

Userid	Stl Wt	Transition/s Inst	Relse	Inval	Reval	Steal/s Ready	NoRdy	Movement/s PGIN	PGOUT	Reads	Write	Migrate/s MWrit	Xrel	Nr of Users
>>Mean>>	1.0	25707	25327	2056K	2018K	263K	68.0	.0	.0	261K	236K	.0	.0	297
LXI01600	1.0	1601	1601	2935K	2867K	2911K	549.5	0	0	2999K	2832K	0	0	
LXI01601	1.0	4840	3826	3016K	2881K	3102K	617.8	0	0	3137K	3016K	0	0	
LXI01602	1.0	7615	3704	2990K	2884K	3064K	617.8	0	0	3172K	2944K	0	0	
LXI01603	1.0	7209	3017	3013K	2899K	3072K	768.0	0	0	3178K	2969K	0	0	
LXI01604	1.0	832.9	1106	3093K	3268K	3106K	474.5	0	0	3448K	3076K	0	0	

There are several places to see z/VM paging activity per-user. Be careful!
- FCX290 UPGACT – those columns are in units of BYTES per second
- FCX113 UPAGE – those columns are in units of PAGES per second

There are also places to see how z/VM's paging is affecting the guests.
- FCX114 USTAT %PGW – percent of samples where the guest is observed to be in page-wait
- FCX164 USTATLOG, FCX315 USTMPLOG – similar to FCX114 USTAT

Memory: Is a Guest Paging Itself?

```
1FCX147  Run 2019/03/22 06:50:22          VDISKS
                                          Virtual Disks in Storage: Activity and Paging
 From 2019/03/21 17:38:12
 To   2019/03/21 17:49:12
 For    660 Secs 00:11:00                 Result of AP1Z00D Run
```

Owning Userid	Virtual Disk Devno	Blocks	LINKs	IO/s	V	Data Space, Rate per Sec. Pgstl	Pgrds	Pgwrt	X-rds	X-wrt	X-mig	Nr of Pages Resid	IBR	Lockd	XSTOR	DASD
>System<	....	4144k	1	.000	-	56.45	.000	56.35	.000	.000	.000	16721	724	0	0	66315
LNSMASTR	2222	4144k	1	.000	N	56.45	.000	56.35	.000	.000	.000	16721	724	0	0	66315

If the guest uses a vdisk for its paging device, check FCX147 VDISKS.
If the guest uses a real disk for its paging device, check FCX108 DEVICE.

Is My I/O Full?

Different Ways We Run Out of I/O

Situation	Primary indicators	Secondary indicators
A device becomes completely busy.	FCX108 DEVICE FCX282 DEVICE HPF	
We run out of HyperPAV aliases in an LCU.	FCX327 HPALIAS	
A FICON or FCP chpid becomes completely busy.	FCX215 FCHANNEL	
An FCP device becomes completely busy.	FCX251 QDIO	
A PCI function becomes completely busy.	FCX381 PCIACT	
An EDEV becomes completely busy.	FCX249 SCSI FCX250 SCSILOG	
A DASD controller runs out of cache.	FCX177 CACHEXT	
A DASD controller becomes completely CPU-busy.	FCX108 DEVICE	
The CPC's I/O processors (SAPs) become completely busy.	FCX232 IOPROCLG	
We run out of SPOOL, DUMP, or T-DISK space.	FCX109 DEVICE CPOWNED	

I/O: Is Some Device Completely Busy?

```
1FCX108  Run 2019/03/20 13:15:32          DEVICE
                                          General I/O Device Load and Performance
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:19
 For   1203 Secs 00:20:03                 Result of 2U0C021D Run
```

<-- Device Descr. -->			Mdisk	Pa-	<-Rate/s->	<------- Time (msec) ------->					Req.	<Percent>	
Addr	Type	Label/ID	Links	ths	I/O	Avoid	Pend	Disc	Conn	Serv	Resp	CUWt	Qued	Busy	READ
>> All DASD <<			....	..	1.6	.0	.120	7.77	1.24	9.12	11.5	.532	.00	1	85
D117	3390-A	TMD117 CP	0	2	28.3	.0	.113	7.26	.242	7.62	7.62	.029	.00	22	89
D109	3390-A	TMD109 CP	0	2	28.2	.0	.112	7.19	.244	7.55	7.56	.026	.00	21	89
D103	3390-A	TMD103 CP	0	2	28.0	.0	.113	7.33	.240	7.68	7.69	.027	.00	22	89
D123	3390-A	TMD123 CP	0	2	27.8	.0	.121	7.40	.235	7.76	7.77	.022	.00	22	89

Device percent-busy: percent of elapsed time the device is busy actually doing I/O

When a device becomes 100% busy, it can't do any more I/Os per second.

If the device is not an ECKD disk, you're stuck.

If the device is an ECKD disk, maybe you can use HyperPAV aliases to do more I/Os per second to the disk *volume.*

I/O: Are We Running Out of HyperPAV Aliases?

```
FCX327  Run 2018/03/14 11:50:50          HPALIAS
                                         HyperPAV Alias Activity Data
From 2016/09/15 08:19:55
To   2016/09/15 08:37:25
For   1050 Secs 00:17:30                 "This is a performance report"
```

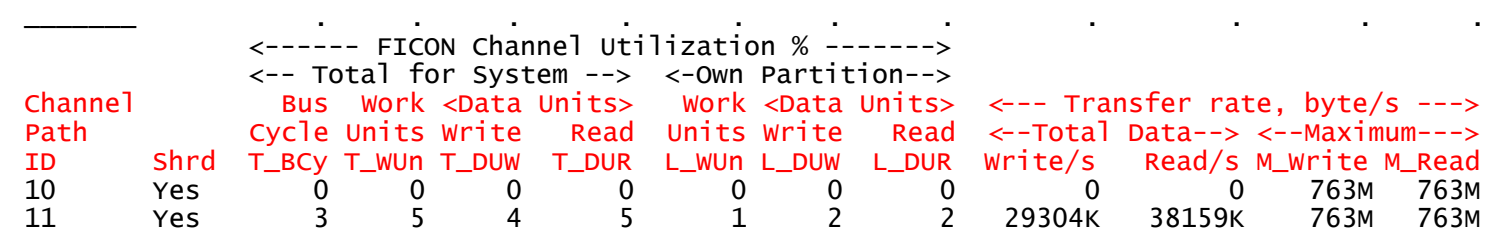
		<------Counts------>			<---Rate/s---->		<--------------Available Alias Extremes--------------->			
SSID	Pool	Base	Alias	SYSTEM	Tries	Fails	Low	WhenLow	High	WhenHigh
0600	0	16	128	4	875.4	763.6	0	2016/09/15,08:30:25	4	2016/09/15,08:30:25

When Fails >> 0 we might want to correct it somehow:
- Add more aliases?
- Reorganize our data into different LCUs?
- Change how our application behaves?
- Adjust MDC?

There might be other choices.

I/O: QDIO, FICON, or FCP Chpid Completely Busy?

```
1FCX215  Run 2019/03/20 13:15:32          FCHANNEL
                                          FICON Channel Load
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
```



Channel Path ID	Shrd	Bus Cycle T_BCy	Work Units T_WUn	<Data Units> Write T_DUW	Read T_DUR	Work Units L_WUn	<Data Units> Write L_DUW	Read L_DUR	<--- Transfer rate, byte/s ---> <--Total Data--> Write/s	Read/s	<--Maximum---> M_Write	M_Read
10	Yes	0	0	0	0	0	0	0	0	0	763M	763M
11	Yes	3	5	4	5	1	2	2	29304K	38159K	763M	763M

Ways a FICON or FCP chpid can fill up:

1. Bus cycles becomes 100%.
2. Work units, whole CEC, becomes 100%.
3. Data units read or written, whole CEC, becomes 100%
4. Work units, the reporting partition, becomes 100%.
5. Data units read or written, the reporting partition, becomes 100%.
6. Data transfer rate, read or written, is close to what the chpid claims it can do.

Cautions.

1. Certain families of FICON/FCP ExpressXX cards have buggy whole-CEC counters.
2. Total Data Write/s and Read/s are based on the whole-CEC counters.

I/O: QDIO or FCP Device Completely Busy?

```
1FCX251  Run 2019/03/20 13:15:32          QDIO
                                          QDIO Device Activity
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
```

Device Number	<-Attached--> Userid	Vdev	QDIO Fmt	<Queues> In	Out	<----- SIGA Instructions/sec ------> <-- By Guest ---> Gst-r	Gst-w	Gst-s	<---- By CP ----> CP-r	CP-w	CP-s	<---- Data Transfer/sec ----> <-- Written -> WBuffer	WBytes	<-- Read ----> RBuffer	RBytes
>> All	QDIO <<	----	----	---	---	62.61	.000	254.4	.000	70.21	.000	77.96	354915	76.00	616
1D0D	WDB00002	0253	FCP	1	1	.000	.000	.000	.000	.000	.000	1.393	17091	1.322	0
1D3C	WDB00001	0253	FCP	1	1	.000	.000	.000	.000	.000	.000	.412	3353	.396	0

Look for WBytes or RBytes to approach the card's capacity.
Usually I estimate the card's bytes/sec capacity to be its bit capacity / 10.
For example, a 16 Gb/sec (gigabit) FICON card can move about 1.6 GB (gigabytes) / sec.

I/O: EDEV Gets Too Busy

```
1FCX249  Run 2019/03/20 13:15:32          SCSI
                                          SCSI device performance data
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
```

Device Number	Bytes per block	<--- Activity/sec ---> KByte	Trans fers	Seeks	<---Blocks---> Read	Written	Percent Connect	EDEV Attr
01FD	512	3	.769	.000	.020	6.133	0	XIV
01FE	512	3	.824	.000	.027	6.567	0	XIV
01FF	512	4	1.049	.000	.007	8.387	0	XIV

Look for percent-connect close to 100. This is the EDEV analogue to FCX108 DEVICE percent-busy.

I/O: DASD Controller Runs Out of Cache

```
1FCX177  Run 2019/03/20 13:15:32          CACHEXT                                                    Page   31
                                          Cache Extended Functions Performance
 From 2019/03/11 06:52:16                                                                    2U0C021D
 To   2019/03/11 07:12:19                                                                    CPU 3906-M05   SN DA1F7
 For   1203 Secs 00:20:03                 Result of 2U0C021D Run                             z/VM   V.7.1.0 SLU 0000
```

<--Device Descr.->			Stg	C	D	D	<--------- Rate/s ---------->			<------------------------ Percent ------------------------>												
			Ctlr	A	F	U	Total	Total	Read	Read	Write	<------ Hits ------>					</TotWrt>	<--Staging--->		DFW	BYP	Time
Addr	Type	VOLSER	ID	C	W	L ST	Cache	SCMBK	N-Seq	Seq	FW	Read	Tot	RdHt	Wrt	DFW	CFW	%DFW	%CFW	Norm	%Seq	DeSt	Byps	+ICL	Outs
D218	3390-A	TMD218	0633	A	A	- 00	55.9	40.4	44.4	1.9	9.6	83	55	46	100	100	..	100	0	65	14	38	0	0	
D109	3390-A	TMD109	062B	A	A	- 00	55.7	1046	44.2	1.9	9.6	83	55	46	100	100	..	100	0	65	14	38	0	0	

```
1FCX176  Run 2019/03/20 13:15:32          CTLUNIT                                                    Page   33
                                          Cached Control Units Overall Performance
 From 2019/03/11 06:52:16                                                                    2U0C021D
 To   2019/03/11 07:12:19                                                                    CPU 3906-M05   SN DA1F7
 For   1203 Secs 00:20:03                 Result of 2U0C021D Run                             z/VM   V.7.1.0 SLU 0000
```

Sub-		<-----Cache Size------>		<---------- DASD Load Data ---------->								<---------- Cache Utilization Data ---------->								
sys-		<Volatile->	<Non-Volat>	<--Total-->								<-Total Rates/s->		<-------- Percentages -------->						
tem	Control	<--(MB)--->	<--(kB)--->	<I/O rates>	Pct	<----- Time (ms) ------>						Read	Read	Write	</TotWrt>	<-----Hits----->				
ID	Unit	Conf	Avail	NV-Cf	NV-Av	Cache	SCMBK	Busy	Pend	Disc	Conn	Serv	Resp	N-Seq	Seq	FW	Read	%DFW	%CFW	RdHt	Wrt	DFW	CFW
062B	2107-E8	123	109	4096	4096	2113	5541	46	.138	6.99	1.12	8.25	8.25	1669	70.9	373.1	82	100	0	45	100	100	..
0633	2107-E8	123	109	4096	4096	2173	1554	15	.120	8.52	1.06	9.70	9.72	1718	71.8	383.3	82	100	0	45	100	100	..

FCX177 CACHEXT is by-volume.
FCX176 CTLUNIT is by-LCU (control unit).

What we are mostly looking for here is hit rates << 100.

I/O: DASD Controller Becomes Too CPU-Busy

```
1FCX108  Run 2020/04/21 07:57:36          DEVICE                                                    Page   34
                                          General I/O Device Load and Performance
 From 2020/04/08 09:34:29                                                                  2U0C913C
 To   2020/04/08 10:01:32                                                                  CPU 8561-T01   SN DEE58
 For   1622 Secs 00:27:02                 Result of 2U0C913C Run                           z/VM   V.7.2.0 SLU 0000
```

<-- Device Descr. --> Addr	Type	Label/ID	Mdisk Links	Pa- ths	<-Rate/s-> I/O	Avoid	<------- Time (msec) -------> Pend	Disc	Conn	Serv	Resp	CUWt	Req. Qued	<Percent> Busy	READ	SEEK Cyls	Recov SSCH	<-Throttle-> Set/s	Dly/s	Devs/ SSID
AC1A	3390-A	JMAC1A ..	19	3	126	.0	3.27	.214	2.19	5.68	47.8	.000	5.33	71	100	15267	0	...	...	0803
AC15	3390-A	JMAC15 ..	19	3	125	.0	3.28	.214	2.19	5.68	47.7	.000	5.26	71	100	15211	0	...	...	0803
AC17	3390-A	JMAC17 ..	19	3	124	.0	3.31	.216	2.20	5.73	48.5	.000	5.36	71	100	15268	0	...	...	0803
AC1B	3390-A	JMAC1B ..	19	3	124	.0	3.32	.216	2.22	5.75	48.4	.000	5.29	71	100	15143	0	...	...	0803
AC18	3390-A	JMAC18 ..	19	3	123	.0	3.33	.214	2.21	5.75	48.9	.000	5.42	71	100	15172	0	...	...	0803
AC16	3390-A	JMAC16 ..	19	3	123	.0	3.34	.213	2.21	5.76	48.6	.000	5.34	71	100	15199	0	...	...	0803
AC19	3390-A	JMAC19 ..	19	3	122	.0	3.35	.218	2.21	5.78	48.7	.000	5.27	71	100	15311	0	...	...	0803
AC0F	3390-A	JMAC0F ..	19	3	120	.0	3.43	.216	2.19	5.84	48.8	.000	5.25	70	100	15094	0	...	...	0803

No field in monitor or Perfkit directly reports the CPU-busy of the DASD control unit.

The only indication or clue we have is when Pend >> 0.

Pend is the time from when the host sends the I/O to when the host receives the control unit's initial response (IR).

When the control unit is experiencing high CPU-busy, it can take a while for the control unit to send IR.

There are other reasons Pend will elevate.

In general, any Pend >>0 is bad.

I/O: System Assist Processors (SAPs) Busy?

```
1FCX232  Run 2019/03/20 13:15:32          IOPROCLG
                                          I/O Processor Activity by Time
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
```

Interval End Time	Proc Number	<-Activity/Sec--> Beg_SSCH	I/O_Int	Proc %Busy	<- Busy conditions per SSCH -> Channel	Switch	CU	Device
>>Mean>>	0	325.1	653.8	1.965	.000	.000	.000	.000
>>Mean>>	1	325.0	3.2	1.958	.000	.000	.000	.000
>>Mean>>	2	325.1	3.5	2.075	.000	.000	.000	.000
>>Mean>>	3	325.0	.0	5.280	.000	.000	.000	.000
>>Mean>>	4	325.0	.0	3.711	.000	.000	.000	.000
>>Mean>>	5	325.1	.1	2.524	.000	.000	.000	.000

The SAPs are the processors where the channel subsystem runs.
Very rarely will you see a SAP more than a few percent CPU-busy.
Channel-busy > 0 probably means not enough chpids (paths) to the control unit.
I have never seen the other three > 0.

I/O: Spool, Dump, T-Disk Full?

```
1FCX109  Run 2019/03/20 13:15:32          DEVICE   CPOWNED
                                          Load and Performance of CP Owned Disks
 From 2019/03/11 06:52:16
 To   2019/03/11 07:12:16
 For   1200 Secs 00:20:00                 Result of 2U0C021D Run
 ------------------------------------------------------------------------------

  Page / SPOOL Allocation Summary
  PAGE slots available      935899k          SPOOL slots available       71140k
  PAGE slot utilization          57%         SPOOL slot utilization         14%
  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
  D227 3390-A TMD227 PAGE    11793600   57   247.4 220.8   ...   ... 468.2  35.5
  9756 3390-3 S31P03 PAGE      1-3338  100    17.6  16.3   ...   ...  33.9   3.0
  9762 3390-3 S31P04 PAGE      1-3338  100    18.4  17.5   ...   ...  35.9   3.5
  D210 3390-A TMD210 PAGE    11793600   57   248.5 220.1   ...   ... 468.6  35.4
  D20B 3390-A TMD20B PAGE    11793600   57   246.9 221.8   ...   ... 468.7  39.1
```

FCX109 DEVICE CPOWNED discusses how full these spaces are.

The book z/VM CP Planning and Admin tells how to plan how much you need.

Is My Networking Full?

Different Ways We Run Out of Networking

Situation	Primary indicators	Secondary indicators
We run out of capacity on an OSA chpid	FCX215 FCHANNEL	
We run out of capacity on an OSA device	FCX251 QDIO	
We run out of capacity on a VSWITCH	FCX240 VSWITCH	
We run out of capacity on a HiperSocket	FCX231 HIPSOCK	
We run out of capacity on some PCI function	FCX312 PCIACT format 0 FCX319 PCIACT format 1 ROCE FCX320 PCIACT format 2 ZEDC FCX322 PCIACT format 3 ISM and their PCILOG correspondents	
Critical guests (e.g., SSL workers) become CPU-bound	FCX112 USER	

Networking: OSA Chpid Full?

```
1FCX215  Run 2019/01/11 16:17:52          FCHANNEL
                                          FICON Channel Load
 From 2019/01/11 14:57:27
 To   2019/01/11 15:02:27
 For    300 Secs 00:05:00                  Result of SRB0111 Run
```

Channel Path		Bus Cycle	Work Units	<Data Units> Write	Read	Work Units	<Data Units> Write	Read	<--- Transfer rate, byte/s ---> <--Total Data-->	<--Maximum--->	
		<------ FICON Channel Utilization % ------->				<-Own Partition-->					
		<-- Total for System -->									
52	Yes	61	34	93	0	32	93	0	1164M	6001794	1250M 1250M

The card’s data bus is 61% busy.
The card's microprocessor is 34% busy.
The card is writing 93% of its stated capacity.

“Total for System” means the whole CPC. (Remember, chpids can be shared across partitions.)
“Own Partition” means traffic of the reporting partition.

In my experience, max byte rate achievable = (card’s stated bit rate / 10) * 0.85 or so.

Networking: OSA Device Full?

```
1FCX251  Run 2019/01/11 16:17:52          QDIO
                                          QDIO Device Activity
 From 2019/01/11 14:57:27
 To   2019/01/11 15:02:27
 For    300 Secs 00:05:00                 Result of SRB0111 Run
```

Device Number	<-Attached--> Userid	Vdev	QDIO Fmt	<Queues> In	Out	<-- SIGA Instructions/sec ------> <-- By Guest ---> Gst-r	Gst-w	Gst-s	<---- By CP ----> CP-r	CP-w	CP-s	<---- Data Transfer/sec ----> <-- Written -> WBuffer	WBytes	<-- Read ----> RBuffer	Rbytes
6760	TCPCB1	6760	QDIO	1	1	.000	.000	.000	.000	61058	.000	35858	1176M	9227	5898K

This device is an OSA Express16S, so I would expect about 1.6 GB/sec * 0.85 = 1.36 GB/sec from it.

It’s almost full.

Handy tip: Perfkit suffixes: k, m, g are powers of 10; K, M, G are powers of 2.

Networking: Vswitch Full?

```
1FCX240   Run 2019/01/11 16:17:52         VSWITCH
                                          VSWITCH Activity
 From 2019/01/11 14:57:27
 To   2019/01/11 15:02:27
 For     300 Secs 00:05:00                Result of SRB0111 Run
 ____________________________________________________________________________

  ____ .         .              .       .      .      .      .      .      .
                         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 (   5)   << 8  300   246m  27306      3 820980  12089      0
  4003 VSWPRF1  DTCVSW1  8  300      0      0      0      0      0      0
  4006 VSWPRF2  DTCVSW2  8  300      0      0      0      0      0      0
  4009 VSWPRF3  DTCVSW1  8  300      0      0      0      0      0      0
  400C VSWPRF4  DTCVSW2  8  300      0      0      0      0      0      0
  6760 CCBVSW1  TCPCB1   8  300  1228m   137k     14  4105k  60443      0
```

This is pretty much an alternate expression of FCX251 QDIO. Observe:

FCX240 VSWITCH: T_Byte = 1228m = 1228 * 1000000 = 1 228 000 000
FCX251 QDIO: Wbytes = 1176M = 1176 * 1048576 = 1 233 125 376
IIRC the difference is due to what exactly is counted, such as framing bytes.

FCX240 VSWITCH does not count the intra-vswitch traffic. It counts only the uplink port traffic.

Remember intra-vswitch traffic is moved by CP on logical processors, so "vswitch full" is also a CPU thing.

Networking: HiperSocket Full?

```
1FCX231  Run 2021/01/14 08:39:50          HIPSOCK
                                          Hipersocket Channel Activity
 From 2020/04/20 11:19:00
 To   2020/04/20 11:25:00
 For    360 Secs 00:06:00                 Result of GONG-01 Run
 ____________________________________________________________________________

  _______          .         .         .         .         .         .
                  <----------------- Hipersocket Activity/Sec. ------------------>
  Channel         <--- Total for System --->  <--------- Own Partition --------->
  Path            <-Transferred-->    Failed  <-Transferred-->  <--- Failed ---->
  ID      Shrd    T_Msgs  T_DUnits  T_NoBuff  L_Msgs  L_DUnits  L_NoBuff  L_Other
  20      Yes        .00       ...       .00     .00       .00       .00      .00
  E0      Yes        .00       ...       .00     .00       .00       .00      .00
```

I put an empty report here on purpose. Here's why.

This report does not tell you whether your HiperSocket chpid is full. It tells you only what its activity is.

HiperSockets data transfers are done by firmware running right on the logical processor that is doing the transfer.

So whether a HiperSockets device is full is entirely about whether there is enough CPU power to run it.

Factors:
1. Is there enough CPU available in the source partition to send the data?
2. Is there enough CPU available in the target partition to receive the data?

When CPU is pretty busy, HiperSockets data transfers will start to become erratic.

Do not use HiperSockets where CPU utilization is already high or where erratic transfer performance will compromise the application. Instead, use an OSA. (Good example: Oracle RAC)

Networking: PCI Functions

- All of these are PCIACT:
  - FCX312 (format 0)
  - FCX318 (format 1, RoCE, an Ethernet adapter)
  - FCX320 (format 2, ZEDC, a data compression card)
  - FCX323 (format 3, ISM, "internal shared memory", an intra-CEC networking card)
- These are very rarely seen
- Most of these variants have Bytes/s column groups on them
  - The exception is FCX320 which reports a load percent (yay!)
- To know whether the PCI function is full, just look up what kind of physical PCI function it is and compare its capacity to the bytes/sec columns on the Perfkit report

Networking: Critical Guests CPU-bound?

```
1FCX112  Run 2019/02/04 11:31:29          USER
                                          General User Resource Utilization
 From 2018/11/12 15:28:36
 To   2018/11/12 15:33:06
 For    270 Secs 00:04:30                 "This is a performance report for SYSTE
```

Userid	%CPU	CPU Load <-Seconds-> TCPU	VCPU	T/V Ratio	Virtual IO/s Total	DASD	Avoid	Diag 98	UR	Pg/s	User Status
TCPIP	.62	1.687	1.045	1.61	6.2	.1	.0	.0	.0	.0	ESA,CL0,DISP
FTPSERVE	.00	.000	.000	.00	.0	.0	.0	.0	.0	.0	XC, ---,DORM
SSLDCSSM	.00	.005	.005	1.00	.0	.0	.0	.0	.0	.0	ESA,---,DORM
SSL00001	10.1	27.36	27.18	1.01	8.0	.0	.0	.0	.0	.0	ESA,CL0,DISP
SSL00002	.00	.006	.005	1.20	.1	.0	.0	.0	.0	.0	ESA,---,DORM
SSL00003	.00	.006	.005	1.20	.1	.0	.0	.0	.0	.0	ESA,---,DORM
SSL00004	.00	.006	.005	1.20	.1	.0	.0	.0	.0	.0	ESA,---,DORM
SSL00005	.00	.006	.005	1.20	.1	.0	.0	.0	.0	.0	ESA,---,DORM

FCX112 USER illustrates the TCP/IP stack guest, the SSL guests, the FTP guest, etc.

Are My CTCs or ISFC Links Full?

Different Ways We Run Out of CTC

Situation	Primary indicators	Secondary indicators
We run out of bus capacity on a FICON adapter	FCX215 FCHANNEL	
We run out of fiber capacity on a FICON adapter	FCX215 FCHANNEL	
A CTC rdev becomes 100% busy	FCX108 DEVICE	

We have already discussed FCX215 FCHANNEL.

Let's look at FCX108 DEVICE and FCX275 ISFLCONF.

CTCs or ISFC: CTC Full?

```
1FCX108  Run 2019/03/26 09:58:52          DEVICE
                                          General I/O Device Load and Performance
 From 2011/10/14 04:51:43
 To   2011/10/14 04:55:43
 For    240 Secs 00:04:00                 Result of H001709C Run
```

<-- Device Descr. --> Addr	Type	Label/ID	Mdisk Links	Pa- ths	<-Rate/s-> I/O	Avoid	<------- Time (msec) -------> Pend	Disc	Conn	Serv	Resp	CUWt	Req. Qued	<Per Busy
6000	CTCA	....	...	1	61.8	...	.507	1.86	13.4	15.8	15.8	.000	.00	98
6001	CTCA	....	...	1	61.7	...	.509	1.85	13.5	15.8	15.8	.000	.00	98
6002	CTCA	....	...	1	61.7	...	.513	1.85	13.5	15.8	15.8	.000	.00	98
6003	CTCA	....	...	1	61.7	...	.516	1.88	13.4	15.8	15.8	.000	.00	98
6020	CTCA	....	...	1	61.5	...	.488	1.85	13.5	15.9	15.9	.000	.00	98
6021	CTCA	....	...	1	61.5	...	.489	1.85	13.5	15.9	15.9	.000	.00	98
6022	CTCA	....	...	1	61.4	...	.494	1.89	13.5	15.9	15.9	.000	.00	98
6023	CTCA	....	...	1	61.3	...	.499	1.87	13.5	15.9	15.9	.000	.00	98
6040	CTCA	....	...	1	171	...	.364	1.99	3.11	5.46	5.46	.000	.00	93
6041	CTCA	....	...	1	170	...	.369	2.01	3.10	5.48	5.48	.000	.00	93
6042	CTCA	....	...	1	.8	...	.310	.142	1.16	1.61	1.61	.000	.00	0
6043	CTCA	....	...	1	472	...	.243	.075	.888	1.21	1.21	.000	.00	57

Use FCX108 DEVICE, just as you would for a DASD. Look for %Busy approaching 100.

What functional areas use CTC links?
- ISFC (SSI) and TSAF
- z/VM TCP/IP point-to-point links
- PVM point-to-point links
- RSCS point-to-point links

Networking: ISFC Logical Link Full?

```
1FCX275  Run 2019/03/26 10:03:57          ISFLCONF
                                          ISFC Logical Link Configuration
 From 2011/10/14 04:50:43
 To   2011/10/14 04:55:43
 For    300 Secs 00:05:00                 Result of H001709C Run
 ------------------------------------------------------------------------

  Initial Status on 2011/10/14 at 04:50:43

  Partner  Devs Rdevs
  GDLBOFVM   12  6000 6001 6002 6003 6020 6021 6022 6023 6040 6041 6042 6043
```

Use FCX275 ISFLCONF to find out which CTC RDEVs belong to your logical link.

Then use FCX108 DEVICE to inspect each of those CTC RDEVs. (FC DEVGROUP?)

Feel free to add more CTC RDEVs, up to the limit of 16.

Configuration notes:

- The minimum recommended configuration is two CTC RDEVs. This is so as to avoid write collisions.

- Best performance happens when there are two CTC RDEVs per CTC chpid.

- Best utilization of CTC chpid resource happens when there are four CTC RDEVs per CTC chpid.

Networking: ISFC Logical Link Activity

- There are some reports that tell you about ISFC link *activity*: data rates, etc.
  - FCX274 ISFLACT ISFC Link Activity
  - FCX273 ISFEACT ISFC Endpoint Activity
- Knowing *activity* is not the same as knowing *how full* something is
- Whether the ISFC link is *full* is really about whether its CTCs are individually busy from an I/O perspective, as we discussed a couple of charts back
- I have not used any of these ISFC reports since 2010
- You might use them if you have to debug a Single System Image data transfer problem

Summary

Summary

- “Is my system full?” is a really difficult question
  - Which resource are we talking about?

- Different resource types require different reports and techniques
  - Processors
  - Memory
  - I/O
  - Networking
  - ISFC

- Go forth and measure

References

- z/VM Version 7 Release 1 Performance Toolkit Reference, IBM file number SC24-6303.
- z/VM Version 7 Release 1 Performance Toolkit Guide, IBM file number SC24-6302.
- SMT vocabulary: http://www.vm.ibm.com/perf/tips/smtvocab.html
- SMT utilization in an SMT world: http://www.vm.ibm.com/perf/tips/smtutil.html
- Understanding z/VM CPU utilization: http://www.vm.ibm.com/perf/tips/lparinfo.html
- z/VM Version 7 Release 1 CP Commands and Utilities, IBM file number SC24-6268.