

VM/ESA Performance Case Studies Volume 2

VM & VSE Tech Conference
May 2000 - Orlando
Session M71

Bill Bitner

IBM Endicott

bitner@vnet.ibm.com

Last Updated: April 10, 2000

[RETURN TO INDEX](#)

Disclaimer

Legal Stuff

The information contained in this document has not been submitted to any formal IBM test and is distributed on an "as is" basis without any warranty either express or implied. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environment do so at their own risk.

In this document, any references made to an IBM licensed program are not intended to state or imply that only IBM's licensed program may be used; any functionally equivalent program may be used instead.

Any performance data contained in this document was determined in a controlled environment and, therefore, the results which may be obtained in other operating environments may vary significantly.

Users of this document should verify the applicable data for their specific environments.

It is possible that this material may contain references to, or information about, IBM products (machines and programs), programming, or services that are not announced in your country or not yet announced by IBM. Such references or information should not be construed to mean that IBM intends to announce such IBM products, programming, or services.

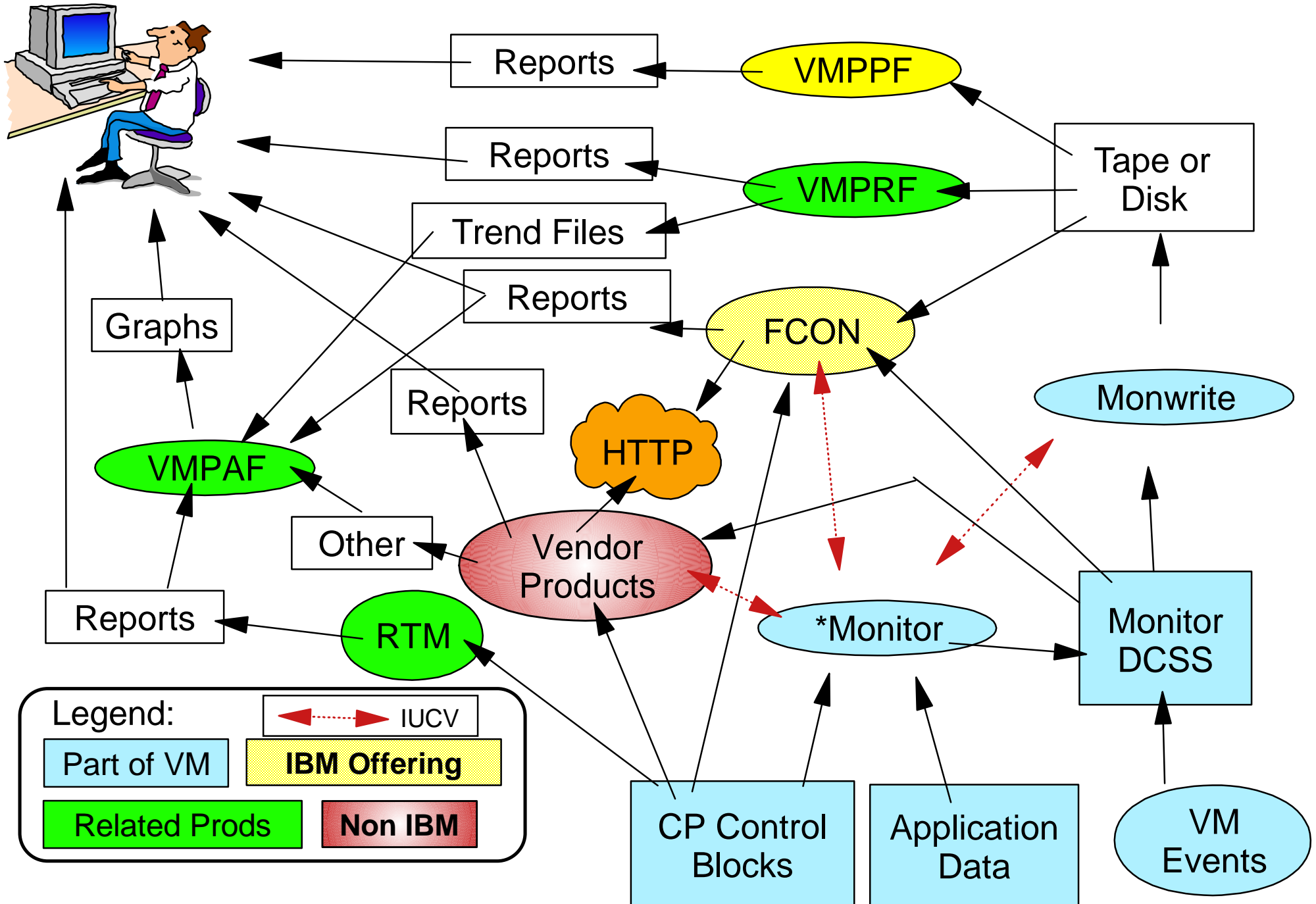
Should the speaker start getting too silly, IBM will deny any knowledge of his association with the corporation.

Trademarks

The following are trademarks of the IBM Corporation:

IBM, VM/ESA

Performance Data Food Chain



Case 1: The Case of Crowded Storage

- I got an e-mail from Erich Amrehn about a customer...

I have a customer question about CMS performance looks like the 16MB (line) is a problem for him and he is looking for some help to identify the problem and possible ways to fix it. Can you help ??

- Erich and I exchanged a couple notes with my last response being...

I would recommend looking at the CMS Storage Utilities: STORMAP, SUBPMAP, and STDEBUG. I'm not familiar with stairs, is it a long running application? Server like? If so, they might want to use the EXTSET option to allow the utilities to collect data while the program is running.

Case 1: No News is Good News?

- No response for a month.
- Then a thank you note from the customer
- Problem solved through Storage Utilities and web page hints.
- The LE segment and Pipelines segments somehow were overlaying each other.
- LE would try to load segment, but could not.
- LE then proceeded to load run time below 16MB storage. (LE is not small)

Case 1: Conclusion

- Bit is not really needed any more, just check the VM Home Page.
- The Storage Utilities can be helpful:
 - ▶ STORMAP - map out storage in CMS
 - ▶ SUBPMAP - map out subpool storage
 - ▶ STDEBUG - track storage obtains/releases
- It is worth double checking segments, especially when LE is involved.

Case 2: The Case of the Rotten RSU

- Customer concern: After applying RSU 9904 to VM/ESA 2.3.0 response time is worse.
- Integrated Server (P390) 256MB/64MB
- Development house:
 - ▶ various 2nd level systems, VM and MVS
 - ▶ 1st level CMS work, batch, ~~SQL/DS~~ DB/2
 - ▶ TCP/IP
- Discussions narrowed it down to almost everyone is affected
- Sent in monitor data

Case 2: Device Config

PRF084 Run 12/08/1999 14:38:10

DEVICE_CONFIGURATION

Configuration Report

VMPRF 1.2.1

From 12/02/1999 10:04:36

To 12/02/1999 10:04:36

For 0 Secs 00:00:00

Bill Bitner Analysis

<-----Ranges----->

<---Channel Path Ids-->

Device Number	Device Sid	Number Of Devices	Device Type	1	2	3	4	5	6	7	8	Control Unit	Status
000C	0000	1	Unit Rec	01		Online
000E-000F	0001-0002	2	Unit Rec	01	2821.01	Online
001E	0003	1	Unit Rec	01	2821.01	Online
0100	0004	1	Unknown		Offline
0101-0104	0005-0008	4	3370 Disk	01	3880-01	Online
0123-0124	0009-000A	2	Unknown		Offline
0126-012F	000B-0011	7	3380 Disk	01	3880-23	Online
0140-0141	0012-0013	2	9336 Disk	01	6310-1	Online
0181-0184	0014-0017	4	3370 Disk	01	3880-01	Online
0200-0204	0018-001C	5	3270	01	3274.1D	Online
0222-0226	001D-001F	3	3380 Disk	01	3880-23	Online
0240	0020	1	Special	01	3745.D1	Online
0280-0284	0021-0025	5	3480 Tape	01	3480.22	Online
0285	0026	1	3480 Tape	01	3480.22	Offline
0290	0027	1	tape	01	3490.51	Online
0300-030E	0028-0035	14	9336 Disk	01	6310-1	Online
0310-031E	0036-0043	14	9336 Disk	01	6310-1	Online
0320-032B	0044-004F	12	3370 Disk	01	3880-01	Online

Case 2: System Summary

PRF002 Run 12/08/1999 14:38:08

SYSTEM_SUMMARY_BY_TIME

System Performance Summa

From 12/02/1999 10:04:36

VMPRF 1.2.1

To 12/02/1999 17:01:36

For 25020 Secs 06:56:59

Bill Bitner Analysis

<-----CPU-----> <Vec> <--Users--> <---I/O--->
 <--Ratio-->

From Time	To Time	Pct Busy	T/V	Cap- ture	On- line	Pct Busy	Log- ged	Activ	Rate	DASD Resp Time
10:04	10:09	27.2	1.22	.9397	1.0	0	48	19	18	0
10:09	10:14	54.1	1.15	.9584	1.0	0	48	19	35	0
10:14	10:19	24.2	1.19	.9334	1.0	0	48	19	14	0
10:19	10:24	48.5	1.20	.9561	1.0	0	48	20	34	0

Case 2: Dasd I/O ?

PRF016 Run 12/08/1999 14:38:11

CACHE_DASD_BY_ACTIVITY

Page 3

Cache DASD Activity Ordered by Activity

From 12/02/1999 10:04:36

VMPRF 1.2.1

To 12/02/1999 17:01:36

CPU 7490

SN 2086

For 25020 Secs 06:56:59

Bill Bitner Analysis

VM/ESA

2.3.0 SLU 990

<-----Device----->					<-SSCH+RSCH->			<-----Time----->					<-----Percent----->		
Num- ber	Volume Serial Type	Control Unit	Cache Size Avail		Count	Rate	Pct Busy	Pend	Disc	Conn	Serv	Resp	Read	Cache Hits	Read Miss
012C	SDSMV4 3380-K	3880-23	0MB		84184	3.4	0	0	0	0	0	0.0	0	0	0
0B22	SCPMV5 3380-E	3880-23	0MB		55190	2.2	0	0	0	0	0	0.1	0	0	0
062D	V81001 3380-E	3880-23	0MB		18111	0.7	0	0	0	0	0	0	0	0	0
062A	230CP0 3380-K	3880-23	0MB		17610	0.7	0	0	0	0	0	0.9	0	0	0
0413	E22W02 3380-J	3880-23	0MB		14206	0.6	0	0	0	0	0	0	0	0	0
065A	BLS35A 3380-J	3880-23	0MB		11731	0.5	0	0	0	0	0	0	0	0	0
0627	BLS627 3380-E	3880-23	0MB		11600	0.5	0	0	0	0	0	0.9	0	0	0

Case 2: Any Knobs turned?

PRF072 Run 12/08/1999 14:38:08

SYSTEM_CONFIGURATION

<---Initial Scheduler Settings----->

IABIAS Intensity	95	Percent
IABIAS Duration	3	Minor Timeslices
DSPSLICE Minor Tslice	10.000	Milliseconds
Hotshot Timeslice	3.999	Milliseconds
STORBUF Q1 Q2 Q3	125	Percent of Main Storage
STORBUF Q2 Q3	105	Percent of Main Storage
STORBUF Q3	95	Percent of Main Storage
LDUBUF Q1 Q2 Q3	100	Percent of DASD Paging Exposures
LDUBUF Q2 Q3	75	Percent of DASD Paging Exposures
LDUBUF Q3	60	Percent of DASD Paging Exposures
Loading User	2	DASD Page Reads per Minor Tslice
Loading Capacity	3	DASD Paging Exposures
MAXWSS	9999	Percent of Main Storage
DSPBUF Q1	70	Openings in Q1 Dispatch List
DSPBUF Q2	20	Openings in Q2 Dispatch List
DSPBUF Q3	10	Openings in Q3 Dispatch List
XSTOR	0	Percent of XSTORE

Case 2: Conclusion

- Not clear who changed various scheduler settings.
- Went back to defaults.
- Things seem much better.
- The third value for DSPBUF was limiting only 10 users to be dispatchable at one time.

SET SRM DSPBUF:
Turn with Care!!

Case 3: The case of a Needle in the Haystack

- Customer with complex processing involving
 - ▶ IBM Products (Callup)
 - ▶ Other Vendor Products
 - ▶ SFS, Spool
 - ▶ Nightly processing
- 9672-R86 partition with 5 logical processors
- 2GB/4GB
- User data transformation
 - ▶ using SFS
 - ▶ using Callup Product for directory services

Case 3: SFS still on my mind

PRF083 Run 08/20/1999 04:20:22

SFS_BY_TIME

SFS Activity by time

From 08/19/1999 19:00:05

VMPRF 1.2.1

To 08/19/1999 23:45:05

For 17100 Secs 04:44:59

<-----Time Per File Pool Request---->

From Time	To Time	Userid	FPR Count	FPR Rate	Total	CPU	Lock	Block I/O	ESM	Other
19:00	23:45	SFS	556891	32.567	0.001	0.000	0.000	0.001	0	0.000

<-----Server Utilization-----> <----Agents-->

Total	CPU	Page Read	Check-point	QSAM	Active	Held	Deadlocks w/ RB
1.1	1.2	0	0.0	0	0.0	1.7	0

SFS is looking good.

Case 3: Silly User loves CPU

PRF008 Run 08/20/1999 04:20:12
 From 08/19/1999 19:00:05
 To 08/19/1999 23:45:05
 For 17100 Secs 04:45:00

USER_RESOURCE_UTIL
 VMPRF 1.2.1
 CASE STUDY 3

Userid	<-----CPU----->				<Vec>	<-User Time->		<-DASD->
	Pct	Total	Virt	T/V Ratio	Secs	<--Minutes-->		Rate
						Logged	Active	While
								Logged
SILLY	12.7	10894	10834	1.0	0	191	191	15.52
AWAYR	0.8	716	699	1.0	0	285	266	27.10
CHANGE1	0.7	624	609	1.0	0	23	23	120.86
VMBACKUP	0.4	382	343	1.1	0	285	285	14.27
VMSPPOOL	0.4	375	336	1.1	0	285	122	13.54
TRANSFOR	0.4	350	281	1.2	0	285	121	8.14
RSCS2	0.4	301	130	2.3	0	285	285	0.00
SFS0005	0.2	199	103	1.9	0	285	282	40.04

Case 3: Why so much Silly User?

- Processing logs showed approximately 283 transformations during that time frame
 - ▶ 10834 seconds virtual CPU
 - ▶ 60 seconds CP CPU
 - ▶ $10834 / 283 = 38.3$ seconds CPU
 - ▶ Roughly 1.9 Billion instructions per transaction!
 - Little CP activity
 - Little activity from the other users
- Something wrong in the SILLY user!

Case 3: Tracking Silly User

- Silly user runs large Execs
- TRACEEXEC tool (from Kent Fiala of SAS)
 - See workshop tool tape pages
- A few **CP Q TIME** inserted in the key exec
- Narrowed it down to a routine named **CALLDBI** which is interface to Callup to get directory record layouts.
 - Note: we could have used the profiling capabilities of TRACEEXEC to narrow down the problem to this level.

Case 3: Recreate the Crime

- Initial attempts to recreate unsuccessful
- Customer found that CALLDBI seemed sensitive to the number of Rexx variables that exist when you invoke CALLDBI
- At this point I was able to recreate
 - ▶ TRACEEXEC
 - ▶ STARS (System Trace Analysis Reports) - currently internal use only

Case 3: RXCALLV is Ugly

Stem Variables	1000	10000	100000
Front matter	0.03	0.03	0.03
calldr	0.07	0.06	0.06
middle 1	0.00	0.00	0.00
rxcallv get	0.06	0.56	5.63
middle 2	0.01	0.00	0.00
rxcallv set	0.00	0.00	0.01
globalv	0.01	0.01	0.01
parse select	0.02	0.02	0.02
about to exit	0.04	0.04	0.04

RXCALLV GET basically passes REXX variables between routines.

Case 3: STARS narrows in

	1 Var	10 Vars	100 Vars	1000 Vars
TOTAL	1.00	1.29	4.23	33.55
DMSITS	1.00	1.33	4.64	37.69
RXCALLV	1.00	1.35	4.90	40.37
DMSITSX	1.00	1.43	5.70	48.45
DMSREX	1.00	1.41	5.48	46.22
IXXRVA	1.00	1.39	5.24	43.80
DMSFRG	1.00	1.00	1.00	1.00
DMSFRR	1.00	1.00	1.00	1.00
DMSGU	1.00	1.36	4.96	40.96
DMSFRE	1.00	1.00	1.00	1.00

Number of instructions normalized to 1 variable

Case 3: Conclusion

- Short Term: in transformation processing, only use CALLDBI to get the directory layout once when processing starts instead of for each transformation.
- Long Term: Bit needs to come up with alternative for RXCALLV and get the Callup Product owners to accept it.

Case 4: The Case of "You think you got enough storage there?"

- 9672-Z17 G6 Turbo
- 1994MB cstore / 29GB xstore
- 1150 CMS Users
- A few server machines
- Wanted to know why they saw paging activity with so much storage?
 - ▶ 18 dedicate page volumes
 - ▶ Paging not on any of them

Case 4: Paging Reports

```

PRF088  Run 02/01/2000 13:44:02  DASD_SYSTEM_AREAS
                                           DASD System Areas by Type: Paging and Spooling Activity
From 02/01/2000 13:26:26  VMPRF 1.2.1
To   02/01/2000 13:40:26
For   840 Secs 00:14:00      9672 G6 Turbo CMS Run
  
```

<--Device-->		<-----Slots----->				<-----Rate----->						Serv	PctTim
Num-	Volume		Avail-	Pct	Pct	Page	Page	Spool	Spool		SSCH	Time	PctTim
ber	Serial	Type	able	InUse	InUse	Read	Write	Read	Write	Total	+RSCH	/Page	Alloca
0E0A	LSP3VM	Spool	40680	0	15.1	190.9	0	22.2	19.0	232.1	53.3	0	100
F5C0	SPOL01	Spool	400500	0	0.1	0	0	2.9	3.0	5.9	5.9	1.1	25
F5C1	SPOL02	Spool	400500	0	4.3	0	0	5.6	5.5	11.1	11.1	1.2	45
F5C2	SPOL03	Spool	400500	0	0.0	0	0	1.7	1.9	3.6	3.7	1.0	20
Sum/Mean		Spool	310545	0	1.9	47.7	0	8.1	7.3	63.1	18.5	0.8	47
0E08	LSP1VM	-PgSp	35820	1.5	17.7	321.5	8.4	28.6	13.5	372.1	81.4	0.4	60
Sum/Mean		-PgSp	35820	1.5	17.7	321.5	8.4	28.6	13.5	372.1	81.4	0.4	60

Why page to spool areas?

Case 4: Which Segments?

- We page the initial read of a segment in from spool.

PRF089 Run 02/01/2000 13:44:02

NSS_DCSS

Name of NSS or DCSS	Spool File Number	Creation-Date	<---Users--->		<-----Pages----->		
			Shared Mode / ImgLib	Non- Shared Mode	Shared	No Data Saved	Privat Resid
ASMAPSEG	101	08/06/1996 09:19:20	1	0	256	0	128
CMS	124	09/29/1999 09:10:13	1146	0	1298	0	1060
CMSINST	127	09/02/1999 14:08:14	1153	0	512	0	496
CMSPIPES	120	09/02/1999 13:02:40	1151	0	256	0	255
CMSVMLIB	121	09/02/1999 13:18:38	1150	0	512	0	344
FORTTRAN	100	12/18/1991 09:00:37	0	0	0	0	381
GCSXA	123	09/29/1999 09:03:04	2	0	120	1173	1293
GOODHELP	125	09/29/1999 09:54:09	0	0	0	0	192
MONDCSS	95	01/18/1999 12:18:06	2	0	0	1280	11
VTAMXA	97	08/31/1993 08:34:01	2	0	256	0	128

Difficult to tell due to packed and logical segments.

Case 4: The answer is in there...

- MONITOR LIST1403 describes layout for monitor records.
- Monitor Domain 3 Record 16 (MRSTOSHHD):
NSS/DCSS/SSP Removed From Storage
 - Storage Domain Event Record
 - Saved Segment Name at offset 20 for 8
 - Spool id at offset 40 for 2
- Read raw monitor data with:
 - Using Monview (from samples disk)
 - Pipelines STARMON stage

Case 4: Getting the data out...

```
/* Example with Pipelines Starmon to get D3/R16 spool name and spid */  
'Segment Load MONDCSS'  
'PIPE (endchar ?) STARMON MONDCSS SHARE',  
'| locate 5 x03 | locate 8 x10',  
'| spec 21.8 2 41.2 c2x 11',  
'| a: fanout ',  
'| sort count',  
'| cons',  
'? a: | drop 100 | pipestop'
```



```
MONVIEW raw_monitor out_file (DR 3 16
```

```
Pipe < out_file | monvu t20.8 40.2 (nohdr | sort count | cons
```

```
632  ASMAPSEG  0065  
834  FORTRAN   0064
```

Case 4: Conclusion

- Preload the Assembler and Fortran segments.
- Get rid of the mixed page/spool volume.
- Get a bigger workload to use that storage.

Case 5: The Case of Tired Iron

- Eleventh hour type of migration for Year 2000.
- HPO5 workload running second level on VM/ESA 2.3.0.
- Migrate HPO5 system to VM/ESA, again second level.
 - ▶ V=R Guest
 - ▶ CMS level stays at CMS 5.
- Performance becomes horrible!

Case 5: E before S before J

- Customer mentions it is a 3090
- VM/ESA uses SIE to dispatch guests
- HPO uses LPSW to dispatch guests
- On new configuration we have SIE on top of SIE
- To run efficiently it needs Interpretted SIE assist.
- What model of 3090?

Case 5: If we had RTM

From D GENERAL or D USER or D ULOG:

<USERID>	%CPU	%CP	%EM	ISEC	PAG	WSS	RES	UR	PGES	SHARE	VMSIZE	TYP,CHR,STAT
BITMAN	98	34	64	.00	.00	9995	8198	.0	0	100	256M	VUS,VSI,SIMW
KARLAC	40	2.0	38	27	.00	1824	1829	.0	0	100	64M	VUX,---,IDLE
HOLDER	2.0	.28	1.7	15	.00	2814	2814	.0	523	100	64M	VUC,IAB,IDLE

From D PRIVOPS (last page):

CNTRNAME	INTLVCNT	NSEC	TOTALCNT	NSEC
STOSM	0	0	0	0
TB	0	0	133037	7
V/SIE	82521	2750	2.50E6	145

Case 5: Conclusion

- It was a model "E"
 - ▶ Assist is on the "J" models
 - ▶ Available as RPQ for "S" models
 - ▶ Not available for "E" models
- Remove duplicate layers of VM/ESA
- Running better after down to 1 level of SIE
- All current 9672s and Multiprises support the Assist.