VM Workshop 2012
Lexington, Kentucky
Avoidance of Self-Inflicted Wounds for Technicians

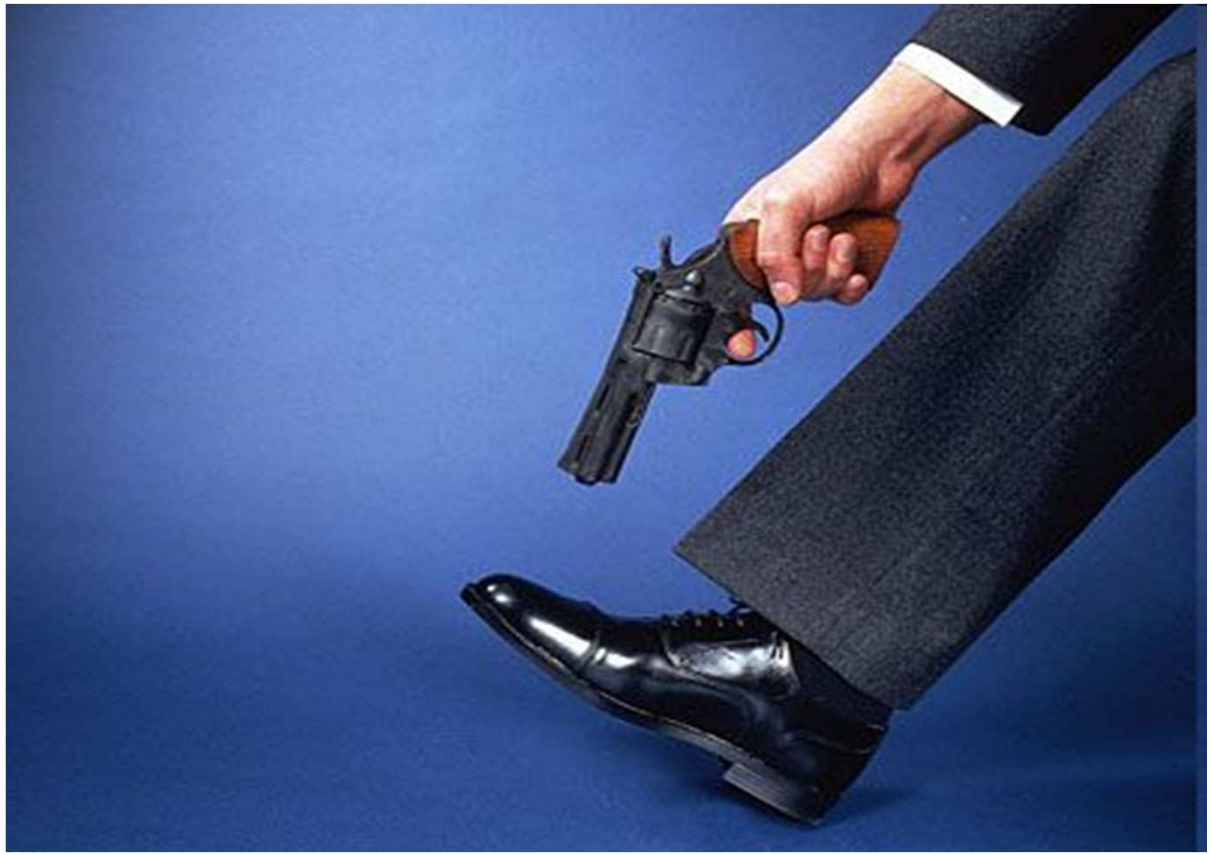
David Kreuter

Self Inflicted Wounds

- Some errors commonly made in z/VM and Linux on System z
- How to avoid SIWs
- What does YGYBYF mean?
- Holistic approaches to tasks and problems
- Heuristics, too.
- Why we self inflict pain.

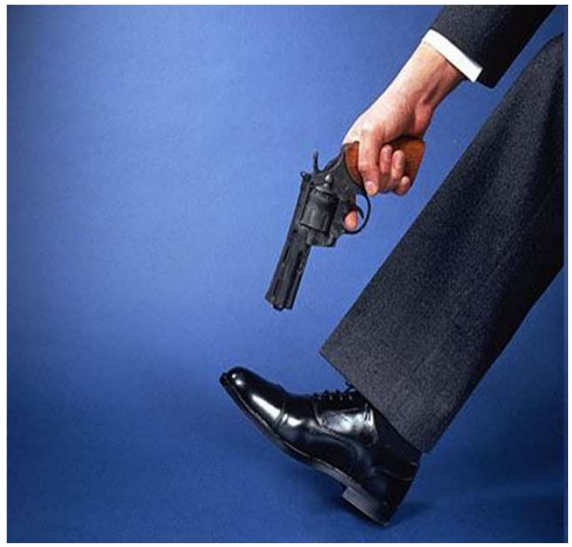
Acronym: YGYBYF
Your Gun Your Bullet Your Foot!

Not me I'm left handed!



Why do we hurt ourselves?

- **Always in a rush to get the task done**
  - **Cutting corners**
  - **Glibness**
- **Lack of skill**
  - **As system integrators responsible for a lot of software!**
    - **Guru level knowledge impossible**
- **Human nature**



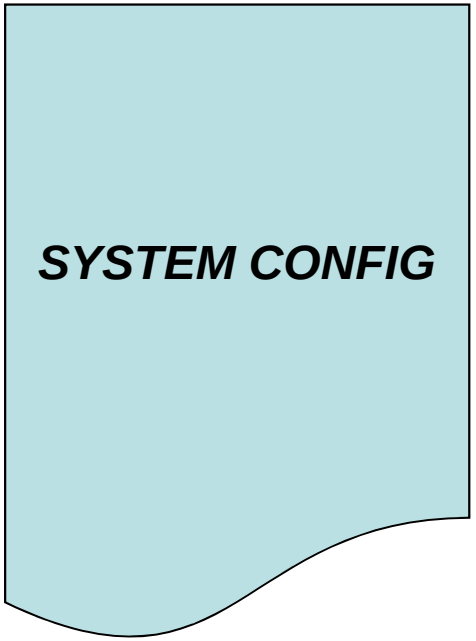
YGYBYF

Scenarios

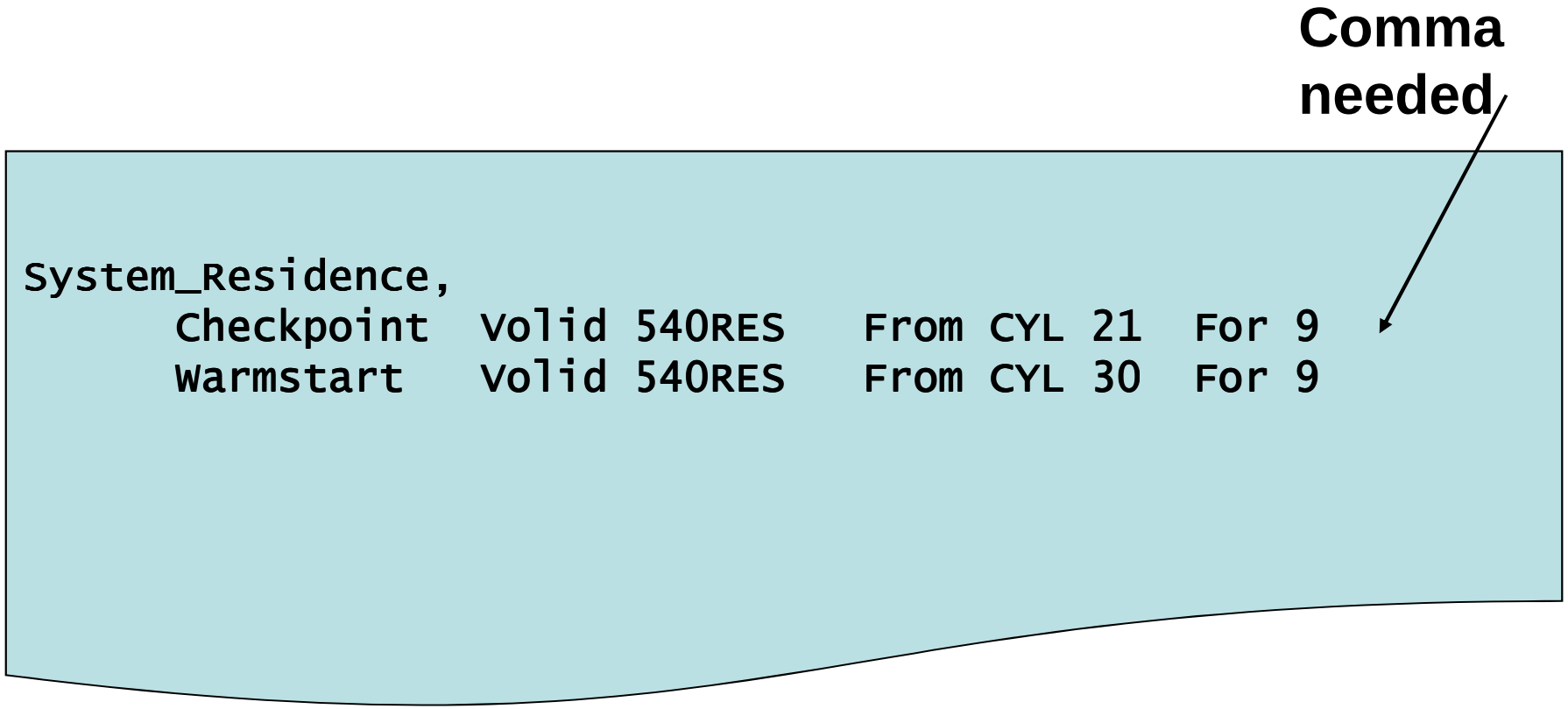
- **z/VM:**
  - **Sloppy practices in SYSTEM CONFIG**
  - **Inadvertent SHUTDOWN**
  - **MDISK overlap**
  - **Poor handling of return code**
- **Linux on System z:**
  - **Forgetting root password**
  - **Runaway writes to the file system**
  - **Package bloat**
  - **Absence of performance monitoring software**

Sloppy SYSTEM CONFIG Coding

- **Scenario: messed up SYSTEM CONFIG**
  - **Removed comma on SYSTEM_RESIDENCE volume**
- **Results: disabled wait 6774**
- **Solution: always run CPSYNTAX utility**
- **Forgiveness not always possible**
  - **Tolerance won't always help**
- **Issue: at 3270 messages roll too quickly**
  - **Use unknown 3270 address**



SYSTEM CONFIG: missing comma in the System_Residence statement



SAPL IPL Sequence

```
 STAND ALONE PROGRAM LOADER: z/VM VERSION 5 RELEASE 4.0

DEVICE NUMBER:   0123       MINIDISK OFFSET:   00000000   EXTENT:  1

MODULE NAME:     CPLOAD     LOAD ORIGIN:       2000

--------------------------------IPL PARAMETERS-----------------------------
   -

-----------------------------------COMMENTS--------------------------------
   -

--------------------------------------------------------------------------
   -


9= FILELIST  10=  OAD  11= TOGGLE EXTENT/OFFSET
```

IPL Fails, Screen rolls quickly; not in spooled console; resulting in a disabled wait PSW

```
HCPGIR450W CP entered; disabled wait PSW
 00020000 00000000 00000000 00006774
```

MSG HCP6774W All Help Information line 1 of 11
(c) Copyright IBM Corporation 1990, 2008

HELP

HCP6774W NO <WARMSTART|CHECKPOINT> AREA DEFINED.

Explanation: You have not specified a warm start or checkpoint area on the SYSTEM_RESIDENCE statement in the system configuration file.

System Action: CP loads a disabled wait state (wait state code = 6774).

User Response: Edit the system configuration file and IPL again.

To view the messages use an undefined console

Used a console address not coded in “operator_consoles” in SYSTEM CONFIG.

```
Operator_Consoles      0020 0021 0022 0023 0327 0328 0329 ,
                 0009      System_3270 System_Console
```

define 9 99
CONS 0099 DEFINED

TERM CONM 3270

Then IPL …

The IPL now displays the messages

```
10:50:56 z/VM  V5 R4.0  SERVICE LEVEL 1001 (64-BIT)
10:50:56 SYSTEM NUCLEUS CREATED ON 2010-05-21 AT 10:42:08, LOADED FROM 540RES
10:50:56
10:50:56 ************************************************************
10:50:56 * LICENSED MATERIALS - PROPERTY OF IBM*                      *
10:50:56 *                                                           *
10:50:56 * 5741-A05 (C) COPYRIGHT IBM CORP. 1983, 2008. ALL RIGHTS    *
10:50:56 * RESERVED. US GOVERNMENT USERS RESTRICTED RIGHTS - USE,     *
10:50:56 * DUPLICATION OR DISCLOSURE RESTRICTED BY GSA ADP SCHEDULE   *
10:50:56 * CONTRACT WITH IBM CORP.                                    *
10:50:56 *                                                           *
10:50:56 * * TRADEMARK OF INTERNATIONAL BUSINESS MACHINES.           *
10:50:56 ************************************************************
10:50:56
10:50:56 HCPZCO6718I Using parm disk 1 on volume 540RES (device 0123).
10:50:56 HCPZCO6718I Parm disk resides on cylinders 39 through 158.
10:50:56
10:50:56 HCPZPM6700E File SYSTEM CONFIG, record 34:
10:50:56 HCPZPM6701E Invalid system configuration file statement - Warmstart
10:50:56 HCPISA6774W NO WARMSTART  AREA DEFINED
HCPGIR450W CP entered; disabled wait PSW 00020000 00000000 00000000 00006774
```

To view the messages with influence of “tolerate_config_errors no” coded in SYSTEM CONFIG

Won’t self correct but will at least give decent documentation

```
 10:58:05 z/VM  V5 R4.0  SERVICE LEVEL 1001 (64-BIT)
 10:58:05 SYSTEM NUCLEUS CREATED ON 2010-05-21 AT 10:42:08, LOADED FROM 540RES
10:58:05
10:58:05 :
10:58:05
 10:58:05 HCPZCO6718I Using parm disk 1 on volume 540RES (device 0123).
 10:58:05 HCPZCO6718I Parm disk resides on cylinders 39 through 158.
 10:58:05
 10:58:05 HCPZPM6700E File SYSTEM CONFIG, record 37:
 10:58:05 HCPZPM6701E Invalid system configuration file statement - Warmstart
 10:58:05 HCPASK6717A
 10:58:05 HCPASK6717A One or more errors were encountered in processing
 10:58:05 HCPASK6717A sections of the system configuration file that were
10:58:05 HCPASK6717A marked not to tolerate errors.
10:58:05 HCPASK6717A
10:58:05 HCPASK6717A To ignore the errors and continue normally, enter GO.
10:58:05 HCPASK6717A To continue with the IPL without autologging any users,
10:58:05 HCPASK6717A enter NOAUTOLOG.
10:58:05 HCPASK6717A To abort the IPL, enter STOP.
:
HCPGIR450W CP entered; disabled wait PSW 00020000 00000000 00000000 00006774
```

How to avoid the Self Inflicted Wound (SIW)

- *Always run the CPSYNTAX utility*
- **See above point**

```
cpsyntax system config z
HCPSYK6700E FILE SYSTEM CONFIG Z, RECORD 37:
HCPSYK6701E INVALID SYSTEM CONFIGURATION FILE STATEMENT -
WARMSTART
Ready(00008); T=0.30/0.31 11:06:05
```

```
/*tolerate_config_errors no*/

 System_Residence,
    Checkpoint  Volid 540RES   From CYL 21  For 9
    Warmstart   Volid 540RES   From CYL 30  For 9
```

How to avoid the Self Inflicted Wound (SIW): “tolerate_config_errors no”

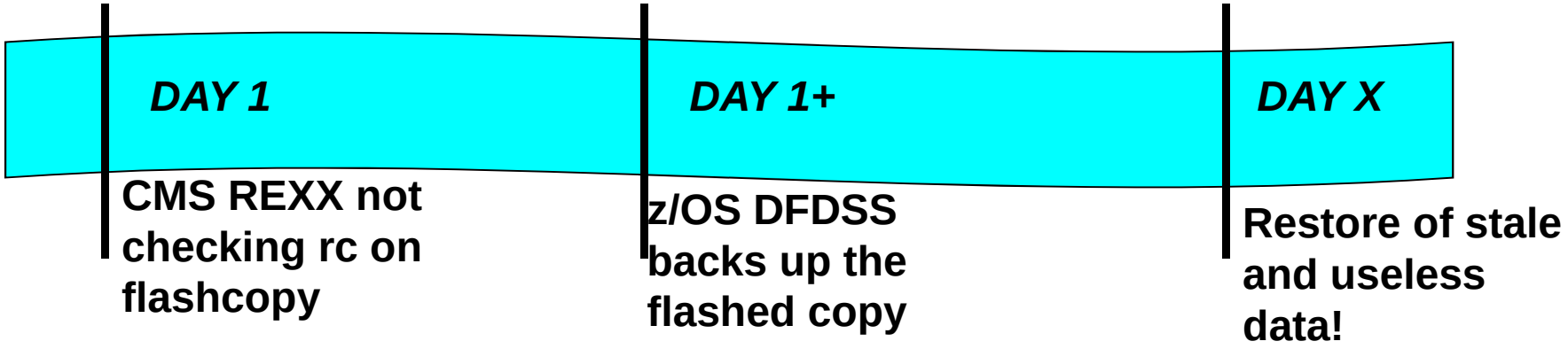
- **Gives more information when error encountered**

```
cpsyntax system config z
HCPSYK6700E FILE SYSTEM CONFIG Z, RECORD 37:
HCPSYK6701E INVALID SYSTEM CONFIGURATION FILE STATEMENT - WARMSTART
HCPSYK6826E ONE OR MORE ERRORS WERE ENCOUNTERED IN PROCESSING SECTIONS OF
THE SYSTEM CONFIGURATION FILE THAT WERE MARKED NOT TO TOLERATE ERRORS.
Ready(00008); T=0.30/0.31 11:07:13
```

tolerate_config_errors no

System_Residence,
Checkpoint Volid 540RES From CYL 21 For 9
Warmstart Volid 540RES From CYL 30 For 9

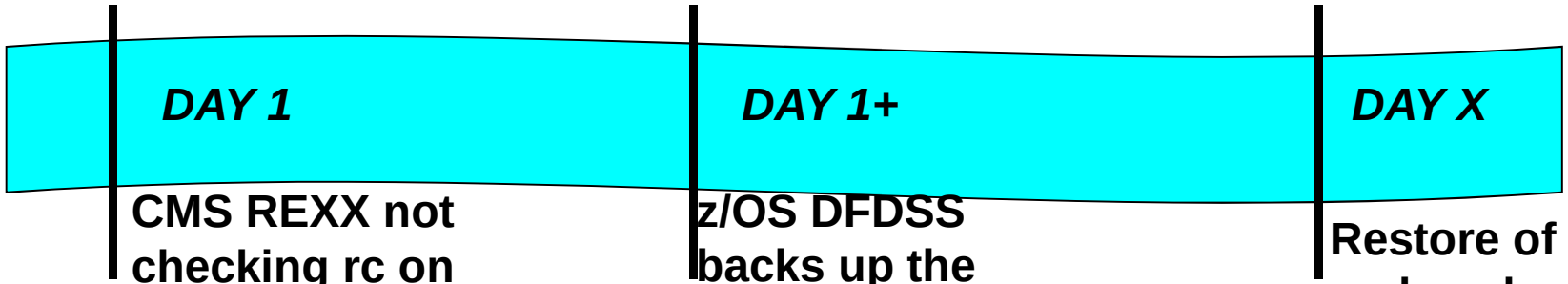
Poor coding practice: return code ignorance



- Scenario: return code not checked in EXEC that does FLASHCOPY commands
  - Used for system level backup
- Results: later stage of backup process is backing up useless stale data
- Solution: check and handle return codes appropriately.

DAY 1	*DAY 1+*	*DAY X*
CMS REXX not checking rc on flashcopy	**z/OS DFDSS backs up the flashed copy**	**Restore of stale and useless data!**

```
 VOLSM2 = substr(VOL,1,2)'@'substr(VOL,4,1)
 IF SUBSTR(VOL,1,2) = 'VJ' |,
    SUBSTR(VOLSM2,1,4) = 'VM@V' THEN DO
    CMD = 'FLASHCOPY' UNIT '0 END TO' FLASH '0 END'
    CYL0OUT.Y = VOL ':' CMD
    Y = Y + 1
 END
 ELSE DO
    CMD = 'FLASHCOPY' UNIT '1 END TO' FLASH '1 END'
 END
INTERPRET CMD
 :
??????? RC  ????????
```



CMS REXX not checking rc on flashcopy

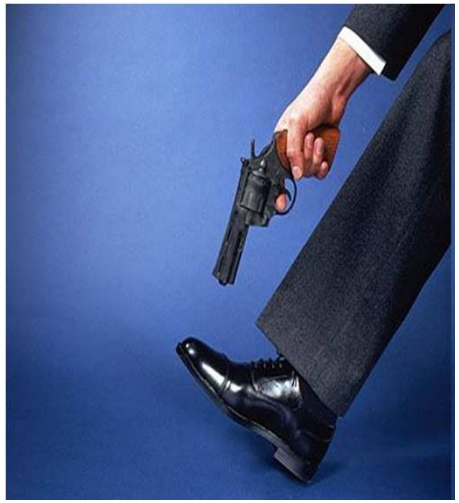
z/OS DFDSS backs up the flashed copy

Restore of stale and useless data!

```
 VOLSM2 = substr(VOL,1,2)'@'substr(VOL,4,1)
 IF SUBSTR(VOL,1,2) = 'VJ' |,
    SUBSTR(VOLSM2,1,4) = 'VM@V' THEN DO
    CMD = 'FLASHCOPY' UNIT '0 END TO' FLASH '0 END'
    CYL0OUT.Y = VOL ':' CMD
    Y = Y + 1
 END
 ELSE DO
    CMD = 'FLASHCOPY' UNIT '1 END TO' FLASH '1 END'
 END
INTERPRET CMD
 :
IF RC <> ‘0’ THEN DO/* TAKE APPROPRIATE ERROR ACTIONS */
   … handle
   … notify
   … complain some more
END
```

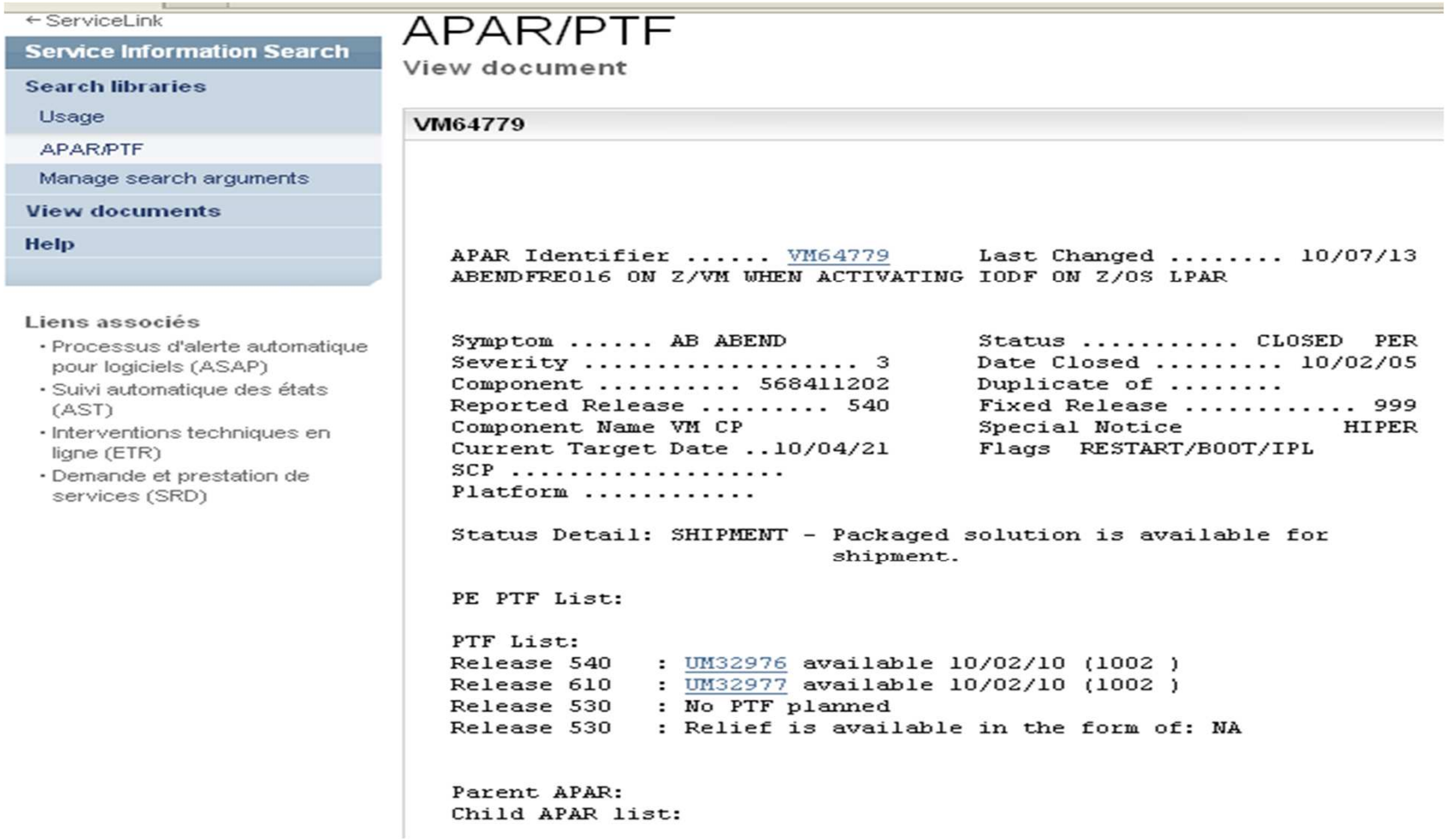
Running with Back Level Code

- **Scenario: Back level operating system releases, application packages, etc.**
  - **True for CPU models as well**
- **Results: exposure to abends, lack of tools, potential for slow performance, etc.**
- **Solution: regular and timely service and maintenance windows.**
  - **Always be researching fixes, RSUs, service pack availability, etc.**
- **Issues are tough to solve:**
  - **Site policies.**
  - **Management directives.**
- **Bane of our existence.**



YGYBYF

Back levelling: APAR/PTF that prevents a potential abend in z/VM CP



← ServiceLink

Service Information Search

Search libraries

Usage

APAR/PTF

Manage search arguments

View documents

Help

Liens associés

- Processus d'alerte automatique pour logiciels (ASAP)
- Suivi automatique des états (AST)
- Interventions techniques en ligne (ETR)
- Demande et prestation de services (SRD)

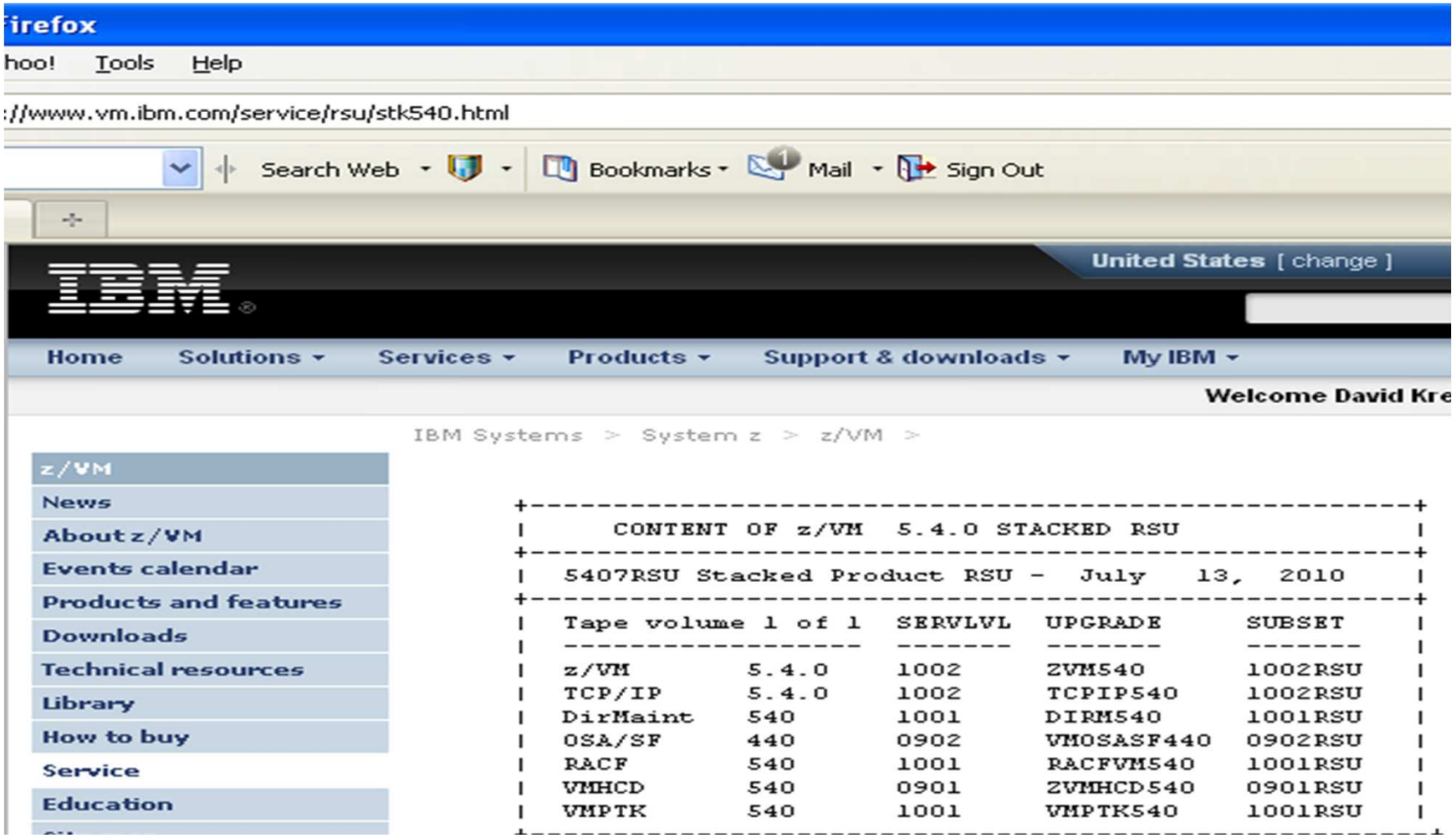
APAR/PTF

View document

VM64779

```
APAR Identifier ..... VM64779       Last Changed ........ 10/07/13
ABENDFRE016 ON Z/VM WHEN ACTIVATING IODF ON Z/OS LPAR


Symptom ..... AB ABEND              Status ............ CLOSED  PER
Severity ..................... 3    Date Closed ......... 10/02/05
Component ........... 568411202     Duplicate of ........
Reported Release ......... 540      Fixed Release ............. 999
Component Name VM CP                Special Notice            HIPER
Current Target Date ..10/04/21      Flags  RESTART/BOOT/IPL
SCP ....................
Platform .............

Status Detail: SHIPMENT - Packaged solution is available for
                          shipment.

PE PTF List:

PTF List:
Release 540    : UM32976 available 10/02/10 (1002 )
Release 610    : UM32977 available 10/02/10 (1002 )
Release 530    : No PTF planned
Release 530    : Relief is available in the form of: NA


Parent APAR:
Child APAR list:
```

Back leveling avoidance: list of latest z/VM RSUs



CONTENT OF z/VM 5.4.0 STACKED RSU

5407RSU Stacked Product RSU - July 13, 2010

Tape volume 1 of 1	SERVLVL	UPGRADE	SUBSET
z/VM 5.4.0	1002	ZVM540	1002RSU
TCP/IP 5.4.0	1002	TCPIP540	1002RSU
DirMaint 540	1001	DIRM540	1001RSU
OSA/SF 440	0902	VMOSASF440	0902RSU
RACF 540	1001	RACFVM540	1001RSU
VMHCD 540	0901	ZVMHCD540	0901RSU
VMPTK 540	1001	VMPTK540	1001RSU

PTF Available in RSU 1002 that can improve performance

PTF List:
Release 530 : UM33014 available 10/04/07 (1000)
Release 540 : UM33015 available 10/04/07 (1002)
Release 610 : UM33016 available 10/04/07 (1002)
ERROR DESCRIPTION: PEVM64225 PGMBKs are being skipped during steal processing, causing some sub-optimal performance on the system.

```
service cp status um33015
VMFSRV2760I SERVICE processing started
VMFSRV1227I UM33015 is not received or applied to CP
   (5VMCPR40%CP)
VMFSRV2760I SERVICE processing completed successfully
```

Adverse impact remaining back leveled!

Checking CP Level

Logged on as MAINT

```
service cp status
VMFSRV2760I SERVICE processing started
VMFSRV1225I CP (5VMCPR40%CP) status:
VMFSRV1225I    Service Level     RSU-1001
VMFSRV1225I    Production Level  RSU-1001
VMFSRV2760I SERVICE processing completed successfully
Ready; T=0.16/0.17 10:29:10
query cplevel
z/VM Version 5 Release 4.0, service level 1001 (64-bit)
Generated at 05/21/10 10:42:08 EDT
IPL at 05/25/10 11:17:57 EDT
Ready; T=0.01/0.01 10:29:28
```

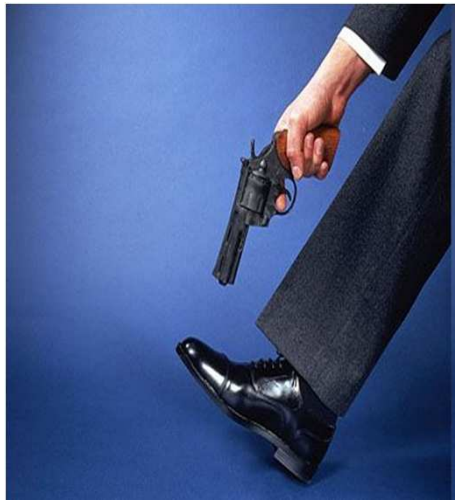
Checking Linux Kernel and Distro Levels

Logged in as root

```
cat /proc/version /etc/SuSE-release
cat /proc/version /etc/SuSE-release
Linux version 2.6.16.60-0.54.5-default (geeko@buildhost)
(gcc version 4.1.2 2007
0115 (SUSE Linux)) #1 SMP Fri Sep 4 01:28:03 UTC 2009
SUSE Linux Enterprise Server 10 (s390x)
VERSION = 10
PATCHLEVEL = 3
:~ #
```

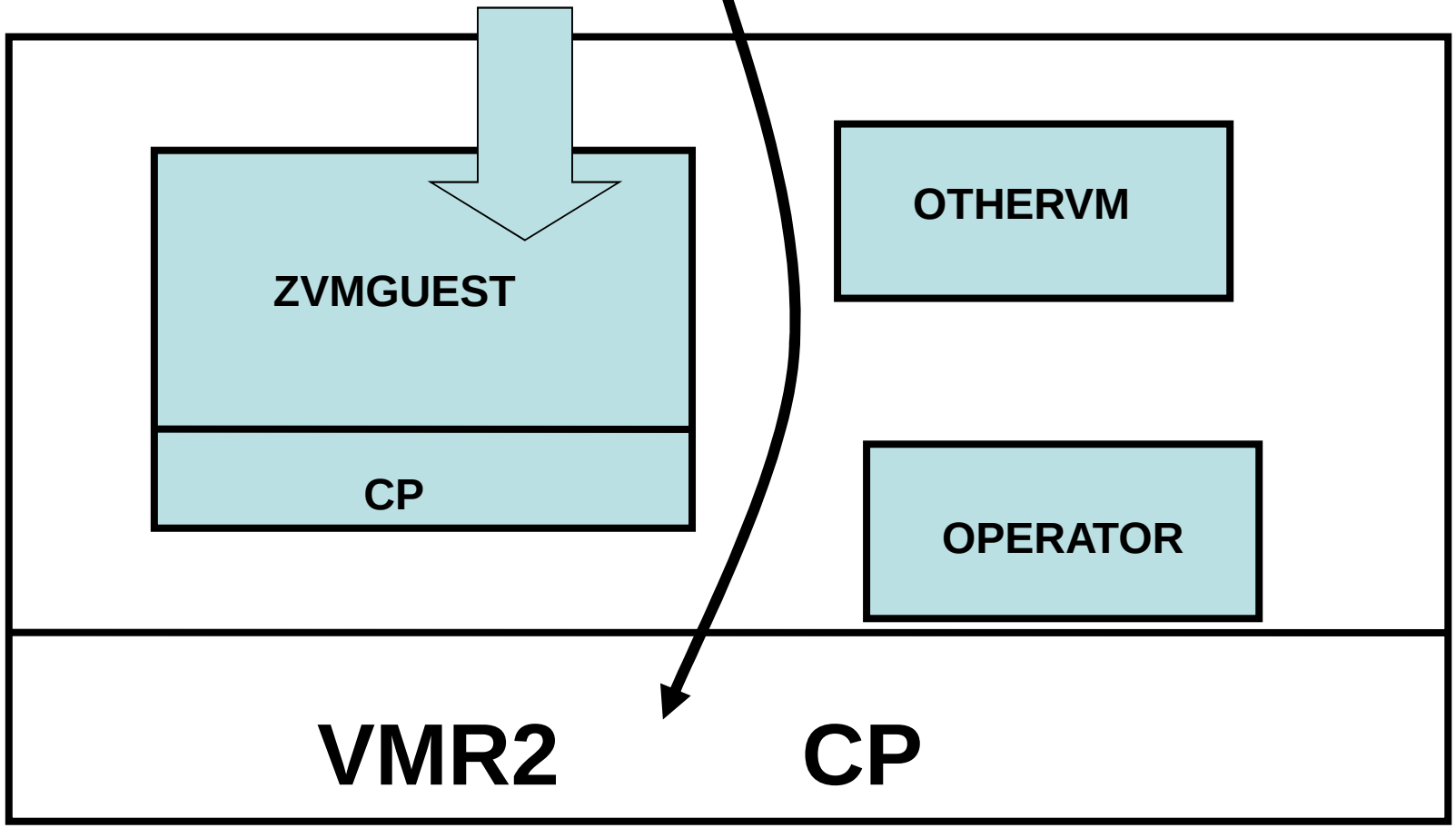
Shutting down an LPAR inadvertently from a 2nd Level System!

- Scenario: Attempting to shutdown 2nd level z/VM system mistakenly shutdown 1st level system
  - Many of us have done this!
- Results: service outage
- Solution: stingy and judicious use of CP privileged commands
  - Never give class “A” or override class (“S”?) to 2nd level systems
- Recommendations:
  - self audit virtual machines that have privileged commands
  - Use logonby for privileged few rather than bolstering personal userids.



YGYBYF

MAINT logged in to ZVMGUEST, does SHUTDOWN at CP READ ka-booming VMR2 (“LPAR”)



Shutting down an LPAR inadvertently from a 2nd Level System! From the production level

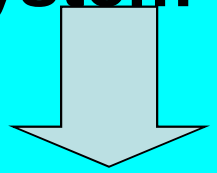
At the console of OPERATOR of production system the 2nd level guest ZVMGUEST is queried. ZVMGUEST has IPLed from its 123 device.

```
q userid
10:33:34 OPERATOR AT VMR2
Ready; T=0.01/0.01 10:33:34
10:35:26 GRAF  0200 LOGON  AS  ZVMGUEST USERS = 5
ind user zvmguest
10:36:00 USERID=ZVMGUEST MACH=ESA STOR=64M VIRT=V XSTORE=NONE
10:36:00 IPLSYS=DEV 0123 DEVNUM=00223
10:36:00 PAGES: RES=00016192 WS=00016185 LOCKEDREAL=00000007 RESVD=00000000
10:36:00 NPREF=00000000 PREF=00000000 READS=00000011 WRITES=00000003
10:36:00 CPU 00: CTIME=00:01 VTIME=000:01 TTIME=000:01 IO=006914
10:36:00         RDR=000000 PRT=000003 PCH=000000 TYPE=IFL  CPUAFFIN=ON
Ready; T=0.01/0.01 10:36:00
query names
10:36:03 ZVMGUEST - 0200, OPERSYMP - DSC , DISKACNT - DSC , EREP  - DSC
10:36:03 OPERATOR - 0009
Ready; T=0.01/0.01 10:36:03
```

Shutting down an LPAR inadvertently from a 2nd Level System!
2nd level OPERATOR console (in the 3270 logon of ZVMGUEST)

```
query names
10:38:00 TCPIP     - DSC , OPERSYMP - DSC , DISKACNT
   - DSC , EREP      - DSC
10:38:00 OPERATOR - 0009
10:38:00 VSM      - TCPIP
Ready; T=0.01/0.01 10:38:00
query userid
10:38:39 OPERATOR AT ZVMGUEST
Ready; T=0.01/0.01 10:38:39
```

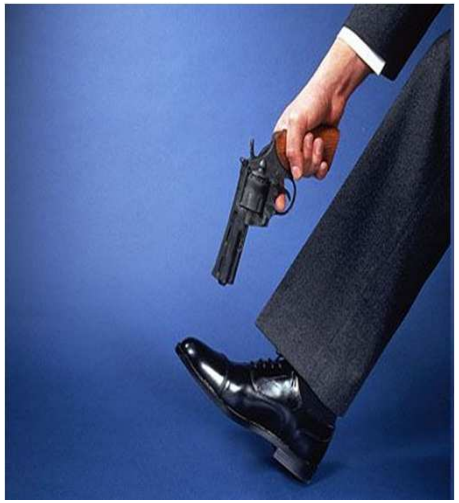
Using the pa1 key so talking to production system VMR2:



CP Read VMR2

Shutting down an LPAR inadvertently from a 2nd Level System!
OPERATOR in ZVMGUEST issues SHUTDOWN to production system

```
shutdown

                                                    CP Read   VMR2
```



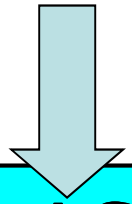
YGYBYF

Shutting down an LPAR inadvertently from a 2nd Level System!
On the production system


```
10:41:07 HCPWRP963I SHUTDOWN STEP USOAC - JOURNAL USER TERMINATION
10:41:07 HCPWRP963I SHUTDOWN STEP MFRSD - TERMINATE HARDWARE LOADER
10:41:08 HCPWRP963I SHUTDOWN STEP APISD - TERMINATE OTHER PROCESSORS
10:41:09 HCPWRP963I SHUTDOWN STEP ENASD - DISABLE TERMINAL DEVICES
10:41:10 HCPWRP963I SHUTDOWN STEP ISHDN - SHUT DOWN I/O SUBSYSTEM
10:41:10 HCPWRP963I SHUTDOWN STEP TTRAL - TERMINATE CONCURRENT COPY
   SESSIONS
10:41:11 HCPWRP963I SHUTDOWN STEP SVACV - ACTIVATE TERMINATION SAVE AREAS
10:41:12 HCPWRP963I SHUTDOWN STEP CHMOF - DISABLE CHANNEL MEASUREMENT
10:41:12 HCPWRP963I SHUTDOWN STEP ISHDA - DISABLE ALL DEVICES
10:41:13 HCPWRP963I SHUTDOWN STEP CKPSH - TAKE A CHECKPOINT
10:41:13 HCPWRP963I SHUTDOWN STEP OPRCK - SAVE OPERATOR CONSOLE LIST
10:41:14 HCPWRP963I SHUTDOWN STEP MCWMD - DETERMINE MACHINE CHECK STATUS
10:41:15 HCPWRP963I SHUTDOWN STEP SDVRS - RESET IBM DASD CU CHARACTERISTICS
HCPWRP962I VM SHUTDOWN COMPLETED IN 9 SEC
10:41:16 HCPWRP963I SHUTDOWN STEP SVADV - DEACTIVATE TERMINATION SAVE AREAS
10:41:16 HCPWRP961W SYSTEM SHUTDOWN COMPLETE
HCPGIR450W CP entered; disabled wait PSW 00020000 00000000 00000000
   00000961
```

Repair the directory statement to remove class A

Broken as implemented!



```
USER ZVMGUEST NOPASS 64M 64M AG
```

Fixed after SIW!

```
USER ZVMGUEST NOPASS 64M 64M G
```

Shutting down an LPAR inadvertently from a 2nd Level System!
After directory changed – all is good!

```
shutdown
HCPCMD001E Unknown CP command: SHUTDOWN
```

CP Read VMR2

Sometimes an error message is a beautiful thing!

VM administration: destructive overlaps

- **Scenario: loading a directory with DIRECTXA doesn't flag overlaps**
  - **Destructive overlaps are poisonous**
- **Results:**
  - **Corrupted file systems**
- **Solution:**
  - **Use a directory manager (DIRMAINT)**
  - **Or execute and examine output from DISKMAP after making directory changes, before putting the directory online.**
  - **Same poison with simultaneous LINK MW**

USR1

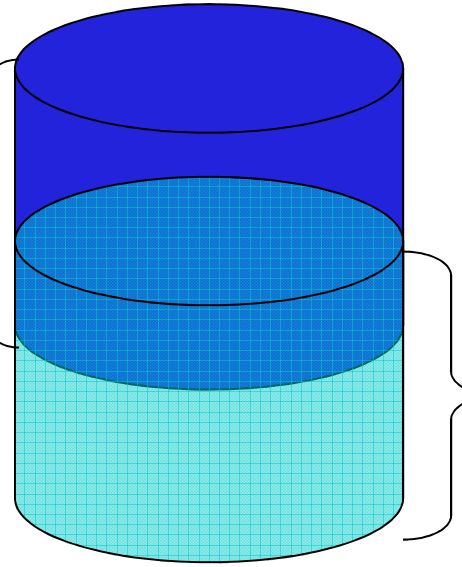
USR2

Destructive Overlap

```
USER USR1 WHATEVER 64M 999M G
   INCLUDE IBMDFLT
   IPL CMS
   MACH XA
   LINK MAINT 019F 019F RR
MDISK 191 3390 6001 20 USR00F MR
```

Minidisk definitions overlap each other

USR1



USR00F

```
USER USR2 WHATEVER 64M 999M G
   INCLUDE IBMDFLT
   IPL CMS
   MACH XA
   LINK MAINT 019F 019F RR
MDISK 191 3390 6011 20 USR00F MR
```

USR2

Destructive Overlap: USR1 does CMS file activities

```
format 191 a
DMSFOR603R FORMAT will erase all files on disk A(191). Do you wish to continue?
Enter 1 (YES) or 0 (NO).
1
DMSFOR605R Enter disk label:
us1191
DMSFOR733I Formatting disk A
DMSFOR732I 20 cylinders formatted on A(191)
Ready; T=0.02/0.17 12:42:24
cp q  mdisk 191 loc
TargetID Tdev OwnerID  Odev Dtype Vol-ID Rdev   StartLoc       Size
USR1     0191 USR1     0191 3390  USR00F 100E       6001         20
```

Format the disk; query its location

Destructive Overlap: USR2 does CMS file activity

```
format 191 a
DMSFOR603R FORMAT will erase all files on disk A(191). Do you wish to continue?
Enter 1 (YES) or 0 (NO).
1
DMSFOR605R Enter disk label:
us1191
DMSFOR733I Formatting disk A
DMSFOR732I 20 cylinders formatted on A(191)
Ready; T=0.02/0.17 12:42:24

cp q mdisk 191 loc
TargetID Tdev OwnerID  Odev Dtype Vol-ID Rdev   StartLoc       Size
USR2     0191 USR2     0191 3390  USR00F 100E       6011         20
Ready; T=0.01/0.01 12:50:06
```

Format the disk;

Query its location

USR2

Destructive Overlap: USR1 does CMS file activity

```
pipe literal $ | dup 88888| specs recno 1 | pad 80 | > $some$ $data$ A f 80
DMSDKD1307T File system error detected by DMSRCM at address 01065892 (offset 000
0273A):
DMSDKD1307T WRTK request failed with a permanent I/O error (sense bytes = 000800
00 C6FFFF00 00000021 00001704 02008B4D 43160F04 000000E2 00177B00) while process
ing file $SOME$ $DATA$ A1
HCPGIR450W CP entered; disabled wait PSW 000A0000 80F082CA
```

USR1

create a file – the cms record manager finds a damaged block!

Solution with DIRMAINT: prevent destructive overlaps

With extent checking on overlaps cannot be formed:

```
dirm extnchk ?
DVHXMT1191I Your EXTNCHK request has been sent for processing.
Ready; T=0.30/0.35 13:08:34
DVHREQ2288I Your EXTNCHK request for MAINT at * has been accepted.
DVHXTN338OI Extent checking is currently ON.
DVHREQ2289I Your EXTNCHK request for MAINT at * has completed; with RC =
DVHREQ2289I 0.
```

USR1

USR1 has a mdisk on USR00F starting on cylinder 6031:

```
cp q mdisk userid usr1 191 drct loc
TargetID Tdev OwnerID  Odev Dtype Vol-ID Rdev   StartLoc       Size
USR1     0191 USR1     0191 3390  USR00F 100E       6031         20
Ready; T=0.01/0.01 13:11:54
```

Prevention is worth a pound of cure

The amdisk for usr2 specifying explicit cylinders fails because dirmaint determines cylinders are already allocated

```
dirm for usr2 amd 191 3390 6041 20 usr00f mr
DVHXMT1191I Your AMDISK request has been sent for processing.
Ready; T=0.33/0.37 13:18:38
 DVHREQ2288I Your AMDISK request for usr2 at * has been accepted.
 DVHSCU3541I Work unit 29131839 has been built and queued for processing.
 DVHSHN3541I Processing work unit 29131839 as MAINT from VSFVM540,
 DVHSHN3541I notifying MAINT at VSFVM540, request 170 for usr2 sysaffin
 DVHSHN3541I *; to: AMDISK 0191 3390 6041 20 USR00F MR
 DVHALC3610E No gap of sufficient size found in candidate area(s).
 DVHAMD3298E DASD allocation attempt failed.
 DVHAMD3298E For: usr2 0191 Request: 3390 6041 20 USR00F MR
 DVHSRL3414I Workunit 29131839 has failed.  It is being saved as 29131839
 DVHSRL3414I WUCFFAIL.
 DVHSRL3414I Workunit 29131839 is being rolled back by DIRMAINT.
 DVHSMA3212E Unexpected RC= 3298, from: EXEC DVHSSHND 29131839
 DVHREQ2289E Your AMDISK request for usr2 at * has failed; with RC =
 DVHREQ2289E 3212.
 DVHREQ2288I Your UNLOCK request for usr2 at * has been accepted.
 DVHREQ2289I Your UNLOCK request for usr2 at * has completed; with RC =
 DVHREQ2289I 0.
```

USR2

Linux: forgetting the root password

- **Scenario: Linux v.m. logged onto 3270 but you have forgotten root password**
  - **Come on, really? A little embarassing, n'est-ce pas?**
- **Results:**
  - **limited use Linux machine**
  - **Um, sysadmin work a little difficult**
- **Solution:**
  - **Use a rescue system**
  - **Or pass in a parameter**

Forgotten root password!

vmrlnx03 login: root
root
Password: rootpass
Login incorrect
:
Red Hat Enterprise Linux Server
release 5.4 (Tikanga)
Kernel 2.6.18-164.el5 on an s390x
vmrlnx03 login:

Forgotten root password: at least determine IPL device

```
CP INDICATE USER *
USERID=VMRLNX03 MACH=ESA STOR=512M VIRT=V XSTORE=NONE
IPLSYS=DEV 0200 DEVNUM=00017
PAGES: RES=00101875 WS=00101546 LOCKEDREAL=00000014
   RESVD=000[illegible]000
NPREF=0001[illegible] PREF=00000000 READS=00000521 WRITES=00014545
XSTORE=01[illegible] ADS=000430 WRITES=025054 MIGRATES=014050
CPU 00: C[illegible].15 VTIME=005:32 TTIME=005:42 IO=052898
        RD[illegible]000 PRT=002464 PCH=000000 TYPE=IFL  CPUAFFIN=
```

Forgotten root password: Use SIGNAL SHUTDOWN CP command to terminate VMRLNX03 (Class G)

```
CP SIGNAL SHUTDOWN WITHIN 30
 INIT: Switching to runlevel: 6
 INIT: Sending processes the TERM signal
Stopping setroubleshootd: [  OK  ]
:
Stopping cups: [  OK  ]
Shutting down xfs: [  OK  ]
:
Starting killall:  [  OK  ]
Sending all processes the TERM signal...
Sending all processes the KILL signal...
:
Unmounting pipe file systems:
Please stand by while rebooting the system...
md: stopping all md devices.
Restarting system.
.
HCPGIR450W CP entered; disabled wait PSW 00020001 80000000 00000000 00000FFF
```

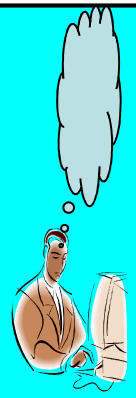
The rescue system: VMRLNX01

The IBM tools vmcp, chccwdev and lsdasd.

```
vmcp link vmrlnx03 200 f200 w write
[root@vmrlnx01 ~]#chccwdev -online 0.0.f200
Setting device 0.0.f200 online
dasd(eckd): 0.0.f200: 3390/0C(CU:3990/01) Cyl:3333 Head:15 Sec:224
dasd(eckd): 0.0.f200: (4kB blks): 2399760kB at 48kB/trk compatible disk
   layout
 dasdf:VOL1/  0X0200: dasdb1
Done

[root@vmrlnx01 ~]# lsdasd 0.0.f200
lsdasd 0.0.f200
Bus-ID     Status      Name      Device  Type  BlkSz  Size      Blocks
==============================================================================
0.0.f200   active      dasdb     94:4    ECKD  4096   2343MB    599940
```

The rescue system: mounting

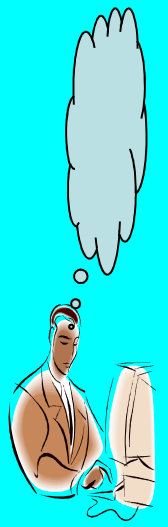
```
[root@vmrlnx01 ~]# mount /dev/dasdb1 /mnt
mount /dev/dasdb1 /mnt

[root@vmrlnx01 ~]#
mount
/dev/dasda1 on / type ext3 (rw)
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
devpts on /dev/pts type devpts (rw,gid=5,mode=620)
tmpfs on /dev/shm type tmpfs (rw)
none on /proc/sys/fs/binfmt_misc type binfmt_misc (rw)
sunrpc on /var/lib/nfs/rpc_pipefs type rpc_pipefs (rw)
/dev/dasdb1 on /mnt type ext3 (rw)
```



VMRLNX03 root partition

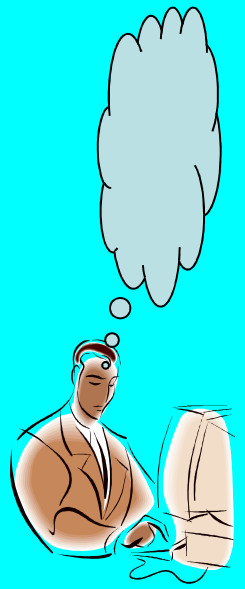
VMRLNX01: chroot, passwd command and rescue termination

```
[root@vmrlnx01 /]#
[root@vmrlnx01 ~]# chroot /mnt
chroot /mnt
[root@vmrlnx01 /]# passwd
   passwd
Changing password for user root.
New UNIX password: test9ng

BAD PASSWORD: it is based on a dictionary word
Retype new UNIX password: test9ng

passwd: all authentication tokens updated successfully.
[root@vmrlnx01 /]#
#cp signal shutdown within 060
```



VMRLNX03 root login works

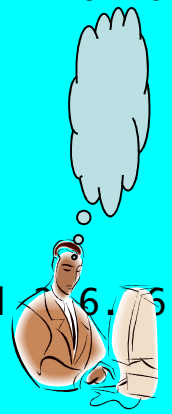
```
Red Hat Enterprise Linux Server release 5.4
   (Tikanga)
Kernel 2.6.18-164.el5 on an s390x

vmrlnx03 login: root
root
Password: test9ng
Last login: Fri Jul 16 13:43:54 on console
[root@vmrlnx03 ~]#
```



Alternate approach using a boot parameter

Forgotten root password again this time using SuSE Linux!

```
Welcome to SUSE Linux Enterprise Server 10 SP2 (s390x) - Kernel 2.6.16.60-
   0.21-default (ttyS0).
svlvdc21 login: root
root
Password: whatever

Login incorrect
Welcome to SUSE Linux Enterprise Server 10 SP2 (s390x) - Kernel 2.6.16.60-
   0.21-d
efault (ttyS0).

svlvdc21 login:
```



Alternate approach with init=/bin/bash parameter

```
CP I 150
zIPL v1.6.3 interactive boot menu
 0. default (ipl)
 1. ipl
 2. Failsafe
Note: VM users please use '#cp vi vmsg <number> <kernel-parameters>'
Please choose (default will boot in 10 seconds):

CP VI VMSG 0 INIT=/BIN/BASH
Booting default (ipl)...
Linux version 2.6.16.60-0.21-default (geeko@buildhost) (gcc version 4.1.2
    200701
15 (SUSE Linux)) #1 SMP Tue May
:
Kernel command line: root=/dev/disk/by-path/ccw-0.0.0150-part1   TERM=dumb
    vmhat=LOGOFF  vmpoff=LOGOFF init=/bin/bash BOOT_IMAGE=0
PID hash table entries: 4096 (order: 12, 131072 bytes)
:
```

Use the VI VMSG command to pass the parm

Alternate approach: login prompt and change the password

```
:
kjournald starting.  Commit interval 5 seconds
EXT3 FS on dasda1, internal journal
EXT3-fs: mounted filesystem with ordered data mode.
(none):/#
```

```
passwd
Changing password for root.
New Password: test9ng
```

Is the passwd command ever happy with your choice?

```
Bad password: it is based on a dictionary word
Reenter New Password: test9ng

Password changed.
(none):/#
```

Alternate approach: Signal shutdown and RelPL

```
CP SIGNAL SHUTDOWN WITHIN 30
md: stopping all md devices.
Restarting system.
.
HCPGIR450W CP entered; disabled wait PSW 00020001 80000000 00000000
   00000FFF
 :
CP IPL 200
zIPL v1.6.3 interactive boot menu
 0. default (ipl)
 1. ipl
 2. Failsafe
Note: VM users please use '#cp vi vmsg <number> <kernel-parameters>'
Please choose (default will boot in 10 seconds):
```

Alternate approach: IPL with default parameters

```
CP VI VMSG 0
Booting default (ipl)...
Linux version 2.6.16.60-0.21-default (geeko@buildhost) (gcc version 4.1.2
   200701
15 (SUSE Linux)) #1 SMP Tue May 6 12:41:02 UTC 2008

Welcome to SUSE Linux Enterprise Server 10 SP2 (s390x) - Kernel 2.6.16.60-
   0.21-d
efault (ttyS0).
svlvdc21 login: root
root
Password: test9ng
Last login: Wed Feb 17 15:31:44 EST 2010 on ttyS0
svlvdc21:~ #
```

So we don't wait for the time interval set in the boot menu

Password set from prior ipl using the init parm

Linux: runaway writes

- **Scenario: without separate partitions or quota controls a runaway write can fill up the root file system.**
- **Results:**
  - **limited use Linux machine**
  - **Um, sysadmin work a little difficult**
- **Solution:**
  - **Use a rescue system to link to the broken root file system, prepare a new disk for broken system, prepare a new /var partition.**

Linux: Runaway Writes causing file system full condition! VMRLNX03

Failing system: Fill up the root file system

```
vmrlnx03:~ # dd bs=1024 count=102240 if=/dev/zero of=/bigfiled
dd: writing `/bigfiled': No space left on device
1425+0 records in
1424+0 records out
1458176 bytes (1.5 MB) copied, 0.052277 seconds, 27.9 MB/s
vmrlnx03:~ # df
Filesystem           1K-blocks      Used Available Use% Mounted on
/dev/dasda1            2361920   2361920         0 100% /
udev                    252412        56    252356   1% /dev
```

Repair: Linux: Runaway Writes causing file system full condition! Move to new partition vi failure vmrlnx03

Failing system: Cannot create files

```
#vi new_file
E303: Unable to open swap file for new_file", recovery impossible
   E297: Write error in swap file
"new_file" [New File]
Press ENTER or type command to continue
```

Repair: Linux: Runaway Writes causing file system full condition! Move to new partition from another machine link online rescue system

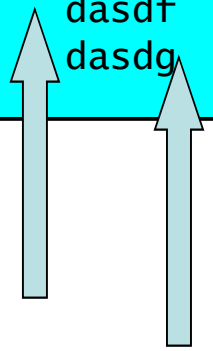
From a rescue system link to the failing Linux' root disk on mdisk f200 and to a newly created 201 on f201

```
VMR-host:~ # vmcp link vmrlnx03 200 f200 w write

VMR-host:~ # vmcp link vmrlnx03 201 f201 w write

VMR-host:~ # chccwdev --online f200
Setting device 0.0.f200 online
Done
VMR-host:~ # chccwdev --online f201
Setting device 0.0.f201 online
Done
VMR-host:~ #
```

IBM tools

Repair: Linux: Runaway Writes causing file system full condition! Move to new partition from another machine link online rescue system

```
VMR-host:~ # lsdasd
Bus-ID     Status      Name      Device  Type  BlkSz  Size      Blocks
==============================================================================
   ===
0.0.0200   active      dasda     94:0    ECKD  4096   7042MB    1802880
0.0.0191   active      dasdb     94:4    ECKD  4096   35MB      9000
0.0.0203   active      dasdc     94:8    FBA   512    125MB     256000
0.0.0201   active      dasdd     94:12   ECKD  4096   3515MB    900000
0.0.0202   active      dasde     94:16   ECKD  4096   3515MB    900000
0.0.f200   active      dasdf     94:20   ECKD  4096   2343MB    599940
0.0.f201   active      dasdg     94:24   ECKD  4096   351MB     90000
VMR-host:~ #
```



Rescue system

Linked from VMRLNX03

Repair: Linux: Runaway Writes causing file system full condition! Move to new partition from another machine dasdfmt f201

Rescue system: Mount the partitions and copy the contents of VMRLNX03 /var to its' eventual new /var on mdisk (F)201. (not shown: the dasdfmt, fdasd, and mke2fs).

```
VMR-host:~ # mount /dev/dasdf1 /mnt/vmrlnx03
mount /dev/dasdf1 /mnt/vmrlnx03
kjournald starting.  Commit interval 5 seconds
EXT3 FS on dasdf1, internal journal
EXT3-fs: mounted filesystem with ordered data mode.
VMR-host:~ # mount /dev/dasdg1 /mnt/newvar

mount /dev/dasdg1 /mnt/newvar
VMR-host:~ #
 cp -a /mnt/vmrlnx03/var/* /mnt/newvar
VMR-host:~ #
```

Repair: Linux: Runaway Writes causing file system full condition! Move to new partition from another machine updated fstab

Rescue system: display the new fstab (edit not shown) for VMRLNX03. Unmount and detach the rescued partitions (F200 and F201 mdisks).

```
VMR-host:/mnt/vmrlnx03/etc # cat /mnt/vmrlnx03/etc/fstab
/dev/disk/by-path/ccw-0.0.0200-part1 /    ext3       acl,user_xattr        1 1
/dev/disk/by-path/ccw-0.0.0201-part1 /var ext2       defaults              0 2
proc                 /proc                proc       defaults              0 0
sysfs                /sys                 sysfs      noauto                0 0
debugfs              /sys/kernel/debug    debugfs    noauto                0 0
devpts               /dev/pts             devpts     mode=0620,gid=5       0 0
VMR-host:/mnt/vmrlnx03/etc #
VMR-host:/ # umount /mnt/vmrlnx03/
VMR-host:/ # umount /mnt/newvar/
VMR-host:/ # chccwdev --offline f200 f201
Setting device 0.0.f200 offline
Done
Setting device 0.0.f201 offline
Done
VMR-host:/ #
```

Repair: Linux: Runaway Writes causing file system full condition! Move to new partition from another machine mounts

Repaired system: IPL

```
CP IPL 200 CL
zIPL v1.6.3-0.24.5 interactive boot menu
 0. default (ipl)
 1. ipl
 2. Failsafe

Note: VM users please use '#cp vi vmsg <input>'
Please choose (default will boot in 10 seconds):
```

Repair: Linux: Runaway Writes causing file system full condition! Move to new partition from another machine mounts

Repaired system: pass in dasd parm to pick up new 201 mdisk …

```
CP VI VMSG 0 DASD=200-201
Booting default (ipl)...
Linux version 2.6.16.60-0.54.5-default (geeko@buildhost) (gcc version 4.1.2
   2007
0115 (SUSE Linux)) #1 SMP Fri Sep 4 01:28:03 UTC 2009
We are running under VM (64 bit mode)
Detected 1 CPU's
Boot cpu address  0
Built 1 zonelists
Kernel command line: root=/dev/disk/by-id/ccw-IBM.68000000033025.0300.0c-
   part1
 TERM=dumb dasd=200-201 BOOT_IMAGE=0
```

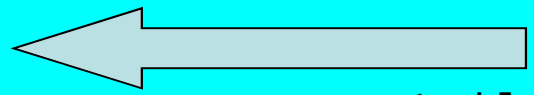
Repair: Linux: Runaway Writes causing file system full condition! Move to new partition from another machine mounts

Repaired system: dasd 200 and 201 found in boot …

```
:
Loading dasd_mod
Loading dasd_eckd_mod
dasd(eckd): 0.0.0200: 3390/0C(CU:3990/01) Cyl:3333 Head:15 Sec:224
dasd(eckd): 0.0.0200: (4kB blks): 2399760kB at 48kB/trk compatible disk
   layout
 dasda:VOL1/  0X0200: dasda1
dasd(eckd): 0.0.0201: 3390/0C(CU:3990/01) Cyl:500 Head:15 Sec:224
dasd(eckd): 0.0.0201: (4kB blks): 360000kB at 48kB/trk compatible disk
   layout
 dasdb:VOL1/  0XF201: dasdb1
```

Repaired system: show the mounted file systems and utilization

```
mount
mount
/dev/dasda1 on / type ext3 (rw,acl,user_xattr)
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
debugfs on /sys/kernel/debug type debugfs (rw)
udev on /dev type tmpfs (rw)
devpts on /dev/pts type devpts (rw,mode=0620,gid=5)
/dev/dasdb1 on /var type ext2 (rw)
securityfs on /sys/kernel/security type securityfs (rw)
df
```

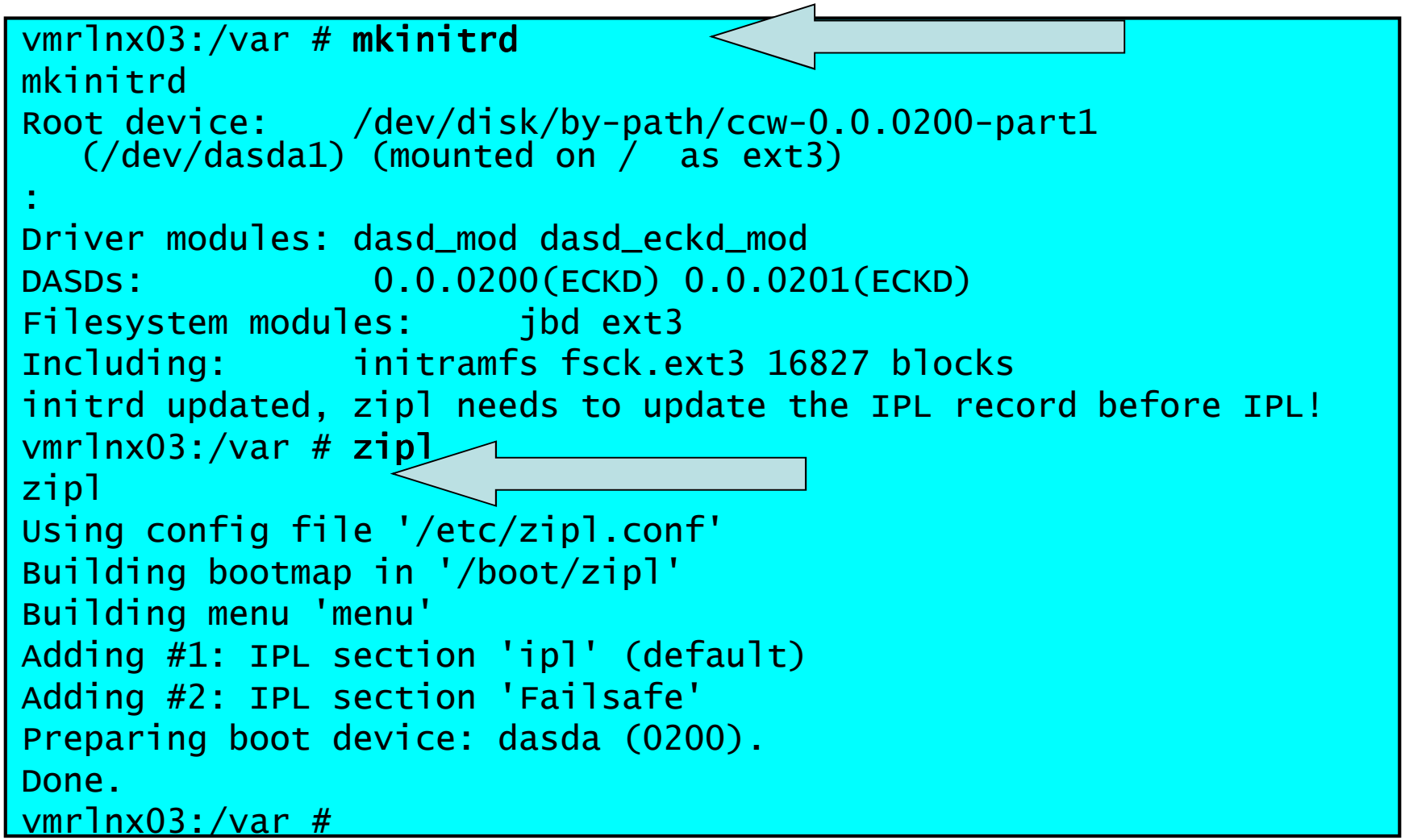


```
df
Filesystem           1K-blocks      Used Available Use% Mounted on
/dev/dasda1            2361920   2133628    108312  96% /
udev                    252412        64    252348   1% /dev
/dev/dasdb1             348608    183980    146636  56% /var
vmrlnx03:/ #
```

Repaired system: vi works

```
~
~
~
~
~
~
~
~
~
"new_file" [New File]
  0,0-1         All
```

Repaired system: run mkinitrd and zip to remember dasd configuration



Linux administration: package removal

- **Scenario:**
  - **Many packages installed from default installation of Linux on System distributions**
- **Results:**
  - **Package bloat**
  - **Unnecessary cycle usage**
- **Solution:**
  - **Remove unneeded packages as part of creating “golden image”**

Reducing Linux footprint

- SuSE distribution from Novell has 3500+ packages
- After removing (not disabling !) unnecessary packages, should finish with about 150 packages
  - Why removing :
    - Can’t be restarted
    - Can’t be hacked
    - Can’t be replaced
- Secure your Linux footprint with RACF (golden image and shared instance in R/O mode)

Pro's to a smaller Linux footprint

- **Fewer packages means :**
  - **Less maintenance to apply**
    - **No unnecessary packages (= no maintenance)**
      - **Security threats (ex.FTP, NNTP), GUI Desktops, Firefox, …**
      - **32-bit packages when using the 64-bit version**
      - **Tools for unused FileSystems**
      - **Compilers, SDK Java, Perl, Yast**
    - **Maintenance to apply**
      - **Very little (ex. zDrivers, surviving packages, …)**
      - **10's of patches to install vs 1000's available**
  - **Less management (2 maintenance projects per year)**
  - **Less chance for intrusions with hardening activities**
    - **Closing ports (21, 80, …)**
  - **Less resources needed (disk & memory)**
  - **Reduce backup (space & time)**
  - **Easier DR exercise**

Sample of packages to be removed

Sample packages to be removed						If YAST is to be removed	
audit-libs-32bit	expat-32bit	libjpeg-32bit	nscd	ppp	yast2-dhcp-server		
autofs	file-32bit	liblcms-32bit	ntfsprogs	pptp	yast2-dns-server	yast2-trans-en_US-2.14.1-36.3	yast2-samba-server-2.13.24-0.3
autoyast2	freetype2-32bit	libmng-32bit	openct	prctl	yast2-heartbeat	yast2-mail-aliases-2.13.8-0.11	yast2-storage-lib-2.13.103-0.3
bind-libs-32bit	gdb-32bit	libnl	openct-32bit	providers	yast2-http-server	yast2-slp-2.13.4-0.17	yast2-hardware-detection-2.13.8-1.2.3
binutils-32bit	gdbm-32bit	libnscd-32bit	OpenIPMI	readline-32bit	yast2-instserver	yast2-perl-bindings-2.13.11-0.22	yast2-country-2.13.56-0.3
bootsplash	gettext-32bit	libpcap-32bit	openldap2-client-32bit	resmgr-32bit	yast2-kdump	yast2-bootloader-2.13.155-0.3	yast2-iscsi-server-2.13.26-0.3
bootsplash-theme-SuSE-SLES	glib	libpng-32bit	opensc	rug	yast2-ldap-server	yast2-mouse-2.13.7-16.2	yast2-online-update-2.13.67-0.3
bzip2-32bit	glib2-32bit	libtiff-32bit	opensc-32bit	sash	yast2-mail	yast2-printer-2.13.32-1.13	yast2-s390-2.13.24-0.3
convmv	glibc-locale-32bit	libtool-32bit	openslp-32bit	scpm	yast2-nis-client	autoyast2-installation-2.13.101-0.3	yast2-xml-2.13.4-0.8
cpp	gnome-icon-theme	libusb-32bit	openslp-server	sharutils	yast2-nis-server	yast2-security-2.13.7-0.11	yast2-kerberos-client-2.13.11-0.10
cpufrequtils	gpm	libxcrypt-32bit	openssl-32bit	siga	yast2-ntp-client	yast2-installation-2.13.211-0.3	yast2-update-2.13.54-0.3
cpufrequtils-32bit	gpm-32bit	libxml2-32bit	opie	smpppd	yast2-online-update-frontend	yast2-storage-2.13.103-0.3	yast2-users-2.13.60-0.3
cracklib-32bit	hal-32bit	libxml2-python	pam-32bit	sqlite	yast2-powertweak	yast2-iscsi-client-2.14.42-0.3	yast2-network-2.13.119-0.3
cups-client	hfsutils	libxslt-32bit	pam-modules-32bit	sqlite-32bit	yast2-profile-manager	yast2-trans-stats-2.11.0-21.18	
cups-libs-32bit	initviocons	libzypp-zmd-backend	parted-32bit	strace-32bit	yast2-registration	yast2-theme-NLD-0.4.5-3.26	
curl-32bit	ipmitool	limal-ca-mgm	pcre-32bit	suseRegister	yast2-repair	yast2-core-2.13.43-0.3	
cyrus-sasl-32bit	ivman	limal-ca-mgm-perl	pcsc-lite	sysfsutils-32bit	yast2-restore	yast2-transfer-2.13.4-0.10	
db-32bit	jfsutils	log4net	perl-32bit	tcpd-32bit	yast2-runlevel	yast2-ncurses-2.13.68-0.3	
dbus-1-32bit	joe	m4	perl-Bit-Vector	usbutils	yast2-schema	yast2-pkg-bindings-2.13.123-0.3	
dbus-1-glib-32bit	krb5-32bit	make	perl-Carp-Clan	utempter-32bit	yast2-slp-server	yast2-2.13.111-0.3	
dbus-1-mono	libacl-32bit	mdadm	perl-Config-IniFiles	vlock	yast2-support	yast2-firewall-2.13.15-0.10	
dbus-1-python	libaio	mkisofs	perl-Date-Calc	w3m	yast2-sysconfig	yast2-ldap-2.13.6-0.3	
delayacct-utils	libaio-32bit	mono-core	perl-TimeDate	wvdial	yast2-tftp-server	yast2-nfs-client-2.13.4-0.11	
deltarpm	libattr-32bit	mono-core-32bit	perl-URI	wvstreams	yast2-tune	yast2-pam-2.13.5-0.11	
device-mapper-32bit	libcap-32bit	mono-data	perl-XML-Parser	xntp	ypbind	yast2-ldap-client-2.13.32-0.3	
dhcpcd	libcom_err-32bit	mono-web	perl-XML-Writer	yast2-autofs	yp-tools	yast2-samba-client-2.13.40-0.3	
dmraid	libgcrypt-32bit	mono-winforms	pmtools	yast2-backup	zisofs-tools	yast2-packager-2.13.179-0.3	
dosfstools	libgpg-error-32bit	mutt	popt-32bit	yast2-boot-server	zlib-32bit	yast2-inetd-2.13.8-0.6	
e2fsprogs-32bit	libgssapi-32bit	ncurses-32bit	powersave-libs	yast2-ca-management	zmd	yast2-nfs-server-2.13.10-0.3	
evms	libidn-32bit	netdate	powersave-libs-32bit	yast2-cd-creator	zmd-inventory		
					zypper		

Scenario: Not Using a Performance Monitor!

- **Scenario:**
  - Without a performance monitor unable to evaluate expensive and critical resources
  - Operating system commands are just not enough
- **Results:**
  - You do not have the data based knowledge required for intelligent decision making
  - Delay in repairing problems
  - Cannot take prevent defense to fend off problems
  - Reliance on other shops' best practices will only get you so far
- **Solution:**
  - Install, care, and maintain your performance monitor!

Without a Good Performance Monitor....

- **You won't get system-wide and Linux virtual machine information:**
  - System Utilization
  - DASD usage
  - LPAR Utilization
  - PAGING and SPOOLING Activity
  - Real Storage Usage
  - TCP/IP Throughput
  - Workload activity
- … and you need both!

Without a Good Performance Monitor

- Carefully define more virtual CPUs for a Linux guest:
- The use of more than one processor requires software locks so that data or control blocks are not updated by more than one processor at a time.
- Set the number of virtual processors based on need and not simply match the number of real that are available.

Without a Good Performance Monitor....example

- You may define the Linux virtual machine larger than you need.
- Excessive virtual machine sizes negatively impact performance.
- Linux uses any extra storage for caching of data. For shared resources, this is an impact.
- Reduce the size of the Linux guest until it starts to swap.
- A good guess at virtual machine size is the z/VM scheduler's assessment of the Linux guest's working set size.

Avoiding the SIW: General Practices

- Documentation
  - Use product documentation
  - Google is not a reference manual!
    - But it is darn useful!
- Checklists
  - Even for us guys …
- Apply service regularly
  - Tell management
- Slow down, breathe, think
- Technical knowledge
- Consultation

Avoiding the SIW: Holistics

Avoid problems via holistic prevention:

- **Always:**
  - **Treat all tasks as a complete entity:**
    - **Why you are doing it**
    - **What tools you need to use to accomplish the task**
    - **What happens if you fail**
    - **All tasks are serious tasks**
    - **Document it for the next time!**
    - **Make things better**
- **Never:**
  - **Work blind**
  - **Work without approvals**

Treating our work as a profession and a craft

Avoiding the SIW: Heuristics

- Avoid problems via heuristic prevention:
  - Experience counts
  - Your existing skill can be adapted to other environments:
    - Copying a file before changing it is equally important in CMS (experienced) as in Linux (new)
    - Coding practices adaptable: REXX (experienced) to bash script (new)
  - Stop! Think! Breathe!
- Appreciate that nuance cannot be adapted but must be acquired

Avoiding the SIW: z/VM Practices

- **Stingy use of privileged commands**
  - **Changing powerful command to override class**
- **Always run CPSYNTAX after SYSTEM CONFIG changes**

Avoiding the SIW: Linux Practices

- Lighten the software load by package removal
- Keep a rescue system handy at all times

Thanks

Thanks to my friends and colleagues for helping me put this together:

- **Dominic Coulombe**
- **Jocelyn Hamel**
- **Dave Jones**
- **Rich Smrcina**
- **George Sobierajski**