IBM Systems & Technology Group

Linux on z/VM System Programmer Survival Guide

Jay Brenneman
rjbrenn@us.ibm.com

Trademarks

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

DB2*
DB2 Connect
DB2 Universal Database
e-business logo*
e-business on demand
HiperSockets
IBM*
IBM eServer
IBM logo*
IMS*
Resource Link
S/390*
Tivoli*
Tivoli Storage Manager
TotalStorage*
WebSphere*
z/OS*
z/VM*
zSeries*

*Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.

UNIX is a registered trademark of The Open Group in the United States and other countries.

SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Agenda

- **Missing Kernel or Initrd**
- **Missing Network connections**
- **Expanding a root LV**
- **LVM Recovery**
- **Customized Initrd**

Missing Kernel symptoms

00: I 201 CL
00: zIPL v1.3.2 interactive boot menu
00:
00: 0. default (linux)
00:
00: 1. linux
00:
00: Note: VM users please use '#cp vi vmsg <input>'
00:
00: Please choose (default will boot in 15 seconds):
00: Booting default (linux)...
00: HCPGIR453W CP entered; program interrupt loop

Missing Initrd symptoms

NET: Registered protocol family 1

NET: Registered protocol family 17

md: Autodetecting RAID arrays.

md: autorun ...

md: ... autorun DONE.

VFS: Cannot open root device "LABEL=/" or unknown-block(0,0)

Please append a correct "root=" boot option

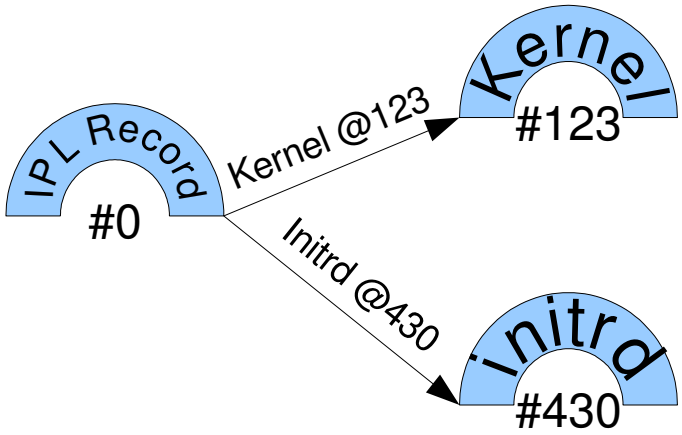
Kernel panic - not syncing: VFS: Unable to mount root fs on unknown-block(0,0)

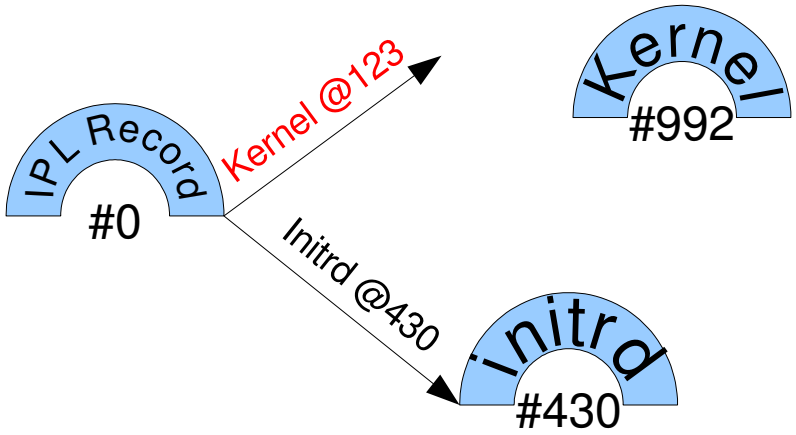
01: HCPGSP2629I The virtual machine is placed in CP mode due to a SIGP stop from CPU 00.

00: HCPGIR450W CP entered; disabled wait PSW 00020001 80000000 00000000 002F2F92

How did we end up here?

- Installed Kernel service without running zipl
  - Might take days or weeks to notice if you don't IPL right away
- Ran mkinitrd without then running zipl
- Physically moved any of the files in /boot





Fixing a missing Kernel or Initrd

- Fixing a broken linux system requires another linux system
  - Any other Linux system with access to the broken system's DASD
  - If no other systems have access to the broken system's DASD, use the initial install media to get something running.
- mount the broken system's partitions
- chroot into the system
- run zipl

Getting the RHEL 4 & 5 installer to see DASD

- IPL the installer
- follow the prompts on the 3270 console to set up the network
- SSH to the install system when prompted
  - The script which runs after SSHing in loads the dasd modules
- Close the SSH session once the "Choose a language" prompt appears
- Open another SSH session to get a shell prompt

Getting the SLES 9 installer to see DASD

- IPL the installer
- follow the prompts on the 3270 console to set up the network
- Choose 0 when asked “Please specify the installation Source:”
  - Cancels out of the installer
- SSH to the install system to get a shell prompt
- Run: modprobe dasd_eckd_mod

Getting the SLES 10 installer to see DASD

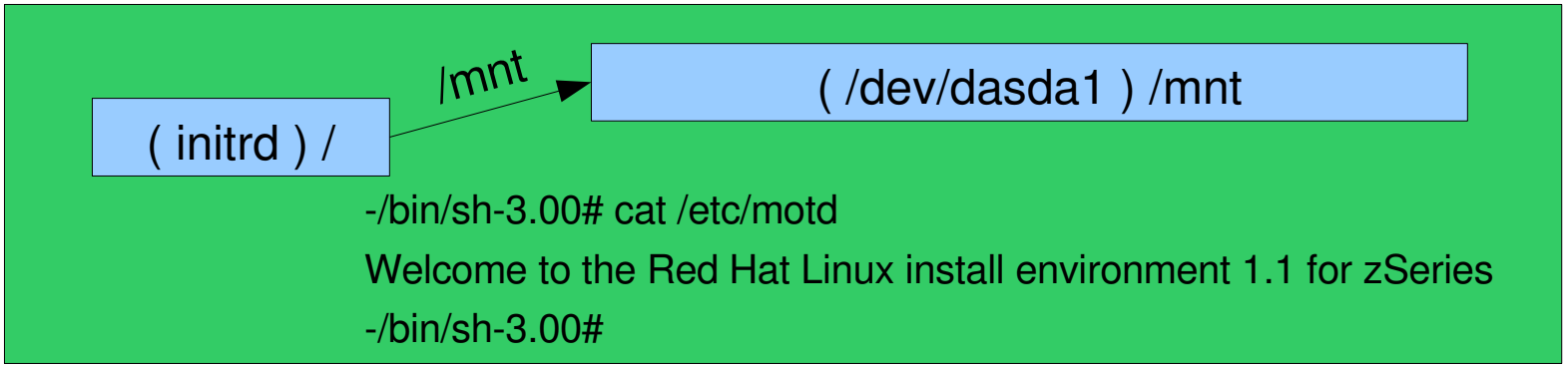
- IPL the installer
- Follow the prompts to install a new system, choose network install and SSH display type
  - The system will load another ram disk from the install server
- SSH in to the system when prompted
- Run: modprobe dasd_eckd_mod

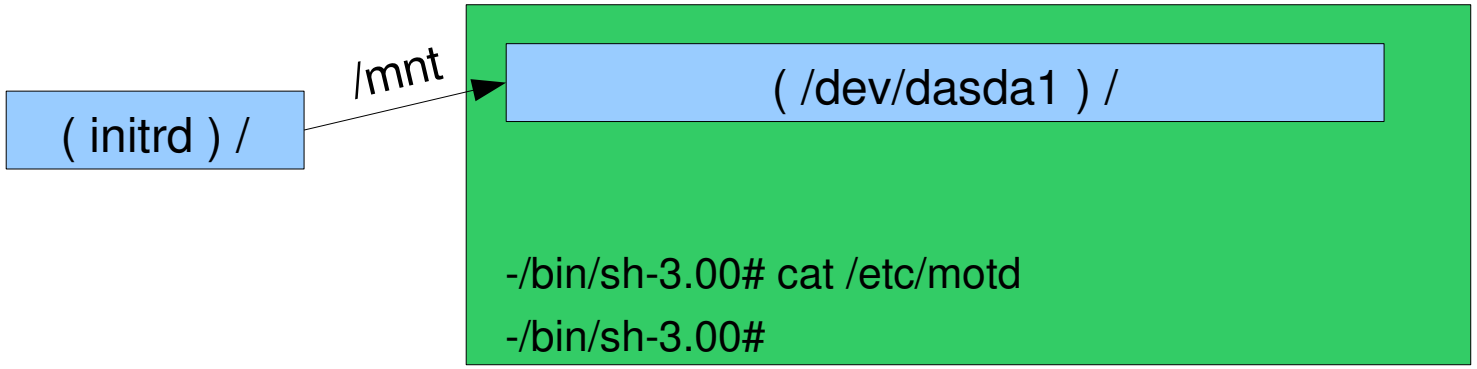
From the shell prompt - 1

- Vary the volumes online
  - echo 1 > /sys/bus/ccw/drivers/dasd-eckd/0.0.0201/online
  - cat /proc/dasd/devices
    - 0.0.0201(ECKD) at (94: 0) is dasda : active at blocksize: 4096, 599400 blocks, 2341 MB
- RHEL 4 & 5 – create the device nodes
  - mknod /dev/dasda b 94 0
  - mknod /dev/dasda1 b 94 1
- Mount the / and /boot file systems
  - Might have to create the mountpoints on RHEL 5
  - mount /dev/dasda1 /mnt

From the shell prompt - 2

- Chroot into the mounted / file system
  - chroot /mnt





From the shell prompt - 3

- Run zipl in the chroot environment to fix the IPL record
  - /sbin/zipl

```
Using config file '/etc/zipl.conf'
Building bootmap '/boot//bootmap'
Building menu 'rh-automatic-menu'
Adding #1: IPL section 'linux' (default)
Preparing boot device: 0201.
Done.
```

- exit the chroot environment
  - exit
- Unmount the partitions
  - umount /mnt
- Vary the devices offline
  - echo 0 > /sys/bus/ccw/drivers/dasd-eckd/0.0.0201/online

Missing Network Connection

- Can't ssh, ftp, ping...
- 3270 console access only
  - ASCII console may be available depending on hardware & z/VM release
- No GUI tools
- 3270 is line mode only – no curses tools, no vi, no emacs*

* no one should be using emacs anyway – it rots your brain and is the leading cause of arthritis among systems programmers

Fixing a Missing Network Connection

- Just get something working so you can ssh in
- Determine if it is a hardware or a software problem
- Define the device triplet to the qeth driver
  - echo “0.0.1f00,0.0.1f01,0.0.1f02” > /sys/bus/ccwgroup/drivers/qeth/group
- Set the portname if needed
  - echo '9DOTLAN' > /sys/bus/ccwgroup/drivers/qeth/0.0.1f00/portname
- Set Layer2 if needed
  - echo 1 > /sys/bus/ccwgroup/drivers/qeth/0.0.1f00/layer2

Fixing a Missing Network Connection

- Vary it online
  - echo 1 > /sys/bus/ccwgroup/drivers/qeth/0.0.1f00/online
- Ifconfig the device up manually
  - ifconfig eth0 9.12.20.154 netmask 255.255.255.0 up
- Add any needed routes
  - route add default gw 9.12.20.1
- SSH in and fix it permanently with the distro's tools

Expanding a root logical volume

- If you use auto partition when you install RHEL 4 your / will be on a LV
- It is not possible to unmount /
- It is not yet possible to resize an ext3 volume while its mounted
  - Resize2fs can expand a volume online
    - Kernel >= 2.6.10
    - E2fsprogs >=1.39-1
    - RHEL 5 can do it

Expanding a root logical volume

- Shutdown the owning system
  - You've already added the 202 disk as a PV to the VG and grown the LV containing the root fs
- Link the volumes to another Linux system
  - vmcp link testa001 201 901 mr
  - vmcp link testa001 202 902 mr
- Vary the volumes online
  - echo 1 > /sys/bus/ccw/drivers/dasd-eckd/0.0.0901/online
  - echo 1 > /sys/bus/ccw/drivers/dasd-eckd/0.0.0902/online
- Pvscan to discover the volume group name
  - pvscan
    - PV /dev/dasdc2 VG VolGroup00 lvm2 [2.19 GB / 0 free]
    - PV /dev/dasdd1 VG VolGroup00 lvm2 [2.28 GB / 0 free]
    - Total: 2 [4.47 GB] / in use: 2 [4.47 GB] / in no VG: 0 [0]

Expanding a root logical volume

- Import the volume group
  - vgexport VolGroup00 && vgimport VolGroup00
- Activate the volume group
  - vgchange VolGroup00 -a y
    - 2 logical volume(s) in volume group "VolGroup00" now active
- Fsck the logical volume
  - e2fsck -f /dev/VolGroup00/LogVol00
  - I'm assuming here (yes, I know the joke) that the logical volume has already been expanded, just not the filesystem

Expanding a root logical volume

- Expand the logical volume
  - resize2fs /dev/VolGroup00/LogVol00
    - resize2fs 1.35 (28-Feb-2004)
    - Resizing the filesystem on /dev/VolGroup00/LogVol00 to 1064960 (4k) blocks.
    - The filesystem on /dev/VolGroup00/LogVol00 is now 1064960 blocks long.
- Vary the volumes offline and detach them
  - echo 0 > /sys/bus/ccw/drivers/dasd-eckd/0.0.0902/online
  - echo 0 > /sys/bus/ccw/drivers/dasd-eckd/0.0.0901/online
  - vmcp det 901
  - vmcp det 902

```
Scanning logical volumes
  Reading all physical volumes.  This may take a while...
  Couldn't find device with uuid 'JzhQlZ-k0ko-1Mgt-qGzI-tvNf-GNS6-feLWdR'.
  Couldn't find all physical volumes for volume group VolGroup00.
  Couldn't find device with uuid 'JzhQlZ-k0ko-1Mgt-qGzI-tvNf-GNS6-feLWdR'.
  Couldn't find all physical volumes for volume group VolGroup00.
  Volume group "VolGroup00" not found
ERROR: /bin/lvm exited abnormally! (pid 190)
Activating logical volumes
  Couldn't find device with uuid 'JzhQlZ-k0ko-1Mgt-qGzI-tvNf-GNS6-feLWdR'.
  Couldn't find all physical volumes for volume group VolGroup00.
  Couldn't find device with uuid 'JzhQlZ-k0ko-1Mgt-qGzI-tvNf-GNS6-feLWdR'.
  Couldn't find all physical volumes for volume group VolGroup00.
  Volume group "VolGroup00" not found
ERROR: /bin/lvm exited abnormally! (pid 191)
Creating root device
Mounting root filesystem
mount: error 6 mounting ext3
mount: error 2 mounting none
Switching to new root
switchroot: mount failed: 22
umount /initrd/dev failed: 2
Kernel panic - not syncing: Attempted to kill init!
01: HCPGSP2629I The virtual machine is placed in CP mode due to a SIGP stop from
 CPU 00.
00: HCPGIR450W CP entered; disabled wait PSW 00020001 80000000 00000000 00040DA0
```


Fixing an LVM group which is missing volumes

- All volumes must be attached and online when the “Scanning logical volumes” step runs in the initrd
- The initrd must be updated to bring those devices online by default when it loads the dasd driver
- But – we have yet another un-bootable system...
- Link the volumes to another system and vary them online to Linux
- Activate the volume group with vgchange

Fixing an LVM group which is missing volumes

- Mount the root logical volume and any sub trees
  - mount /dev/VolGroup00/LogVol00 /mnt
  - mount /dev/dasdc1 /mnt/boot
- Chroot into the broken system
  - Chroot /mnt
- Edit /etc/modprobe.conf to add the missing volumes to the dasd list
  - options dasd_mod dasd=201-202

Fixing an LVM group which is missing volumes

- Make the new initrd with mkinitrd
  - May have to force mkinitrd to load the dasd drivers with a parameter
  - Use the -v flag to generate verbose output
    - Look for the messages for dasd_mod
    - Look for the messages for the file system type
    - Look for the messages for device mapper modules

```
mkinitrd -v --with dasd_eckd_mod /boot/initrd-2.6.9-42.EL.img.new 2.6.9-42.EL
Creating initramfs
Looking for deps of module ide-disk
Looking for deps of module ext3  jbd
Looking for deps of module jbd
Looking for deps of module dm-mod
Looking for deps of module dm-mirror     dm-mod
Looking for deps of module dm-mod
Looking for deps of module dm-zero       dm-mod
Looking for deps of module dm-mod
Looking for deps of module dm-snapshot   dm-mod
Looking for deps of module dm-mod
Looking for deps of module dasd_eckd_mod         dasd_mod
Looking for deps of module dasd_mod
Using modules:  ./kernel/fs/jbd/jbd.ko ./kernel/fs/ext3/ext3.ko ./kernel/drivers/md/dm-mod.ko ./kernel/drivers/md/dm-mirror.ko
./kernel/drivers/md/dm-zero.ko ./kernel/drivers/md/dm-snapshot.ko ./kernel/drivers/s390/block/dasd_mod.ko ./kernel/drivers/s390/
block/dasd_eckd_mod.ko

...some output trimmed here...

Loading module jbd
Loading module ext3
Loading module dm-mod
Loading module dm-mirror
Loading module dm-zero
Loading module dm-snapshot
Loading module dasd_mod with options dasd=201-202
Loading module dasd_eckd_mod
```


Fixing an LVM group which is missing volumes

- Edit /etc/zipl.conf – add a new section using the new initrd

```
[defaultboot]
default=linux
target=/boot/
[linux1]
        image=/boot/vmlinuz-2.6.9-42.EL
        ramdisk=/boot/initrd-2.6.9-42.EL.img.new
        parameters="root=/dev/VolGroup00/LogVol00"
[linux]
        image=/boot/vmlinuz-2.6.9-42.EL
        ramdisk=/boot/initrd-2.6.9-42.EL.img
        parameters="root=/dev/VolGroup00/LogVol00"
```

- Run zipl to update the boot record to include the new initrd

```
Using config file '/etc/zipl.conf'
Building bootmap '/boot//bootmap'
Building menu 'rh-automatic-menu'
Adding #1: IPL section 'linux1'
Adding #2: IPL section 'linux' (default)
Preparing boot device: 0901.
Done.
```

Fixing an LVM group which is missing volumes

- Exit the chroot
  - exit
- Unmount the volumes
  - unmount /mnt/boot
  - unmount /mnt
- Vary the devices offline to Linux and detach them
- IPL your fixed system

Creating a custom initrd

- An initrd is one of 2 things:
  - A file which contains a file system that is gzipped
  - A cpio archive of files that is gzipped
- In either case, it is possible to add software to the initrd
  - Bacula Client
  - TSM client
  - Lightweight editor
  - Dasd utilities

Creating a custom initrd

- Copy the install initrd to a linux system
- Rename the initrd to something that ends in .gz
- Uncompress the file with gunzip
- Test the resulting file to see what it is
  - # file initrd.sles10
    - initrd.sles10: ASCII cpio archive (SVR4 with no CRC)
  - # file initrd.img
    - initrd.img: Linux rev 1.0 ext2 filesystem data

Creating a custom initrd

- XOR
  - extract files from the cpio archive
    - # cpio -id < ../initrd.sles10
      - 49077 blocks
  - mount the ext2 file system
    - mount initrd.img initrd -o loop
- Copy the software into the initrd with tar or cpio
- Check the software's library needs with ldd

```
# ldd ./dsmc
        libcrypt.so.1 => /lib/libcrypt.so.1 (0x77fa4000)
        libpthread.so.0 => /lib/libpthread.so.0 (0x77f8e000)
        libdl.so.2 => /lib/libdl.so.2 (0x77f8a000)
        libstdc++-libc6.2-2.so.3 => /usr/lib/libstdc++-libc6.2-2.so.3 (0x77f3b000)
        libm.so.6 => /lib/libm.so.6 (0x77ea9000)
        libc.so.6 => /lib/libc.so.6 (0x77d7a000)
        /lib/ld.so.1 (0x77fe6000)
```

Creating a custom initrd

- Make sure all the indicated libraries exist within the initrd
  - Copy them in too if not
  - You may run in to space problems with a loopback mounted file
- XOR
  - rebuild the cpio archive
    - # find . | cpio -o > ../initrd.sles10.withtsm
      - 88183 blocks
  - unmount the ext2 file system
    - unmount initrd.img

Creating a custom initrd

- Gzip the resulting file
- Rename it to show it is not the normal initrd
  - initrd.img.withtsm
- Keep it somewhere safe till it's needed

Using a custom initrd

- Follow the instructions in the “Getting the XXXX V installer to see the DASD” slides earlier in this presentation
  - Substitute the altered initrd in place of the default one
- Ssh in to the initrd system
- Attach and vary the DASD online to linux
- Pvscan and vgchange to bring LVM online if needed
- Mount the file systems
- Use the programs you added to the initrd to save the day

Useful Sources of information

- **IBMVM list:**
  - http://listserv.uark.edu/scripts/wa.exe?SUBED1=ibmvm&A=1
- **LINUX390 list:**
  - http://www2.marist.edu/htbin/wlvindex?linux-390
- **Redbooks:**
  - **RHEL 4 cookbook**
    http://www.redbooks.ibm.com/abstracts/sg247272.html?Open
  - **SLES 9 cookbook**
    http://www.redbooks.ibm.com/abstracts/sg246695.html?Open

This space intentionally left blank