Installing a Usable z/VM System is Easy!

Bruce Hayden
IBM Washington Systems Center
bjhayden@us.ibm.com

IBM

Agenda

- Standard z/VM installation
  - History and current method
- A new z/VM system
  - What is there
  - What would be better?
- z/VM installation system
  - Modifying it
  - Creating disk images
  - Restoring the images
- Customization
  - System
  - Network
  - Input panel
- Our install process
  - Booting LPAR
  - Network
  - Input fields
  - System restore
  - Automated customization
- IPL new sytem
- Other items
  - ..like SSI

z/VM Installation

Overview and a little history

Originally tape based

- Restore a starter system using DDR
  - The system must match your DASD! (3330, 3350, 3380, 3390, FB-512)
- IPL that system
- Restore minidisks from tape (VMFPLC2 format)
- Problems with tape formats and compatibility

Now based on capabilities of the HMC

- Read files from DVD, USB, or server
- Load images into memory
- ”Restart” the LPAR
- Documented in the Hardware library
  ”Installing software by using Load from Removable Media or FTP server support”

Summary of the z/VM Installation

- Load VM nucleus and Ramdisk; Start LPAR
- Use HMC 3270 to fill in fields
- Start installation
  - Format, label, and allocate disks
  - Load minidisk image files
    - Pipeline stages created to do this (Not documented, for installation only)
    - Restores initial part of RES disk, initial spool
    - Restores minidisk by minidisk
- IPLs restored system (2nd level)
  - Init spool, init SFS, apply service (RSU)
- Install complete!

Installation Has Finished

Now, what do we have?

A “starter” z/VM system

- One spool volume
- One page volume
- No volumes define to add workload
- Minimal security (no RACF or TLS/SSL)
- No system automation, little monitoring
- Minimal performance monitoring
- Minimal network setup (no vswitch defined)
- No API capability
- No framework for Linux

<u>How do you proceed?</u>

- Maybe you’ve done this before?
  - Did you take notes?

Or

- Follow a redbook (cookbook)
- Instructions from a class
- You have Richard’s lab book
  (Hopefully you went yesterday!)

Fortunately, we upgrade systems more often than performing fresh installs.

<u>But what if this is your first z/VM system?</u>

Let’s Do Something Different!

Create a system ready to be used – a.k.a. a “Usable System”

First, we must create that system

- Install the basic system
- Configure networking, configure a vswitch
- Activate RACF, Dirmaint, Perfkit & configure
- Install z/VM Operations Manager & configure
- Install Linux
- Configure TLS/SSL, LDAP
- Configure and enable SMAPI
- Add some users
- Enable monitor data collection
- Lots of other miscellaneous customization

If you were going to deploy multiple systems, how would you build the “golden” system?

<u>OK, we built it!</u>

Now..
Dump it to tape
Send it in the mail!
Run standalone DDR
Ummm..

NOT!

How Can a Customer Restore Our System?

Especially to a "new" LinuxONE customer with no z/VM skills

Can I use the normal install process?
How does it work?

Boot up the install system, log on and look around
Gee, this looks like a small z/VM system!

- CMS
- Directory source file
- Parm disk
- SYSTEM CONFIG
- Minidisks
- Users

Our "DASD" is a RAM disk, FB-512 format

- This is one of the files loaded by the HMC

What is this RAM disk?

FBA disks are simple. 512 byte fixed blocks, numbered starting at zero. Mapped into memory.

I have a zPDT (emulated Z system.) Take the RAM disk file and mount as an FBA disk.

- Use the source directory and map minidisks
- Make some changes and reverse the process
- Test the modified RAM disk – works fine!

(Later I used Pipelines to read and write the RAM disk image to and from an emulated FBA disk. A zPDT is not required to work with it.)

The z/VM Install System

What can we learn about it?

Your z/VM system comes with source code!

- You don’t get all of it – but quite a lot.

Maybe I can scan it and look for ”ramdisk”

- Hint: Look in VMPSFS:MAINT730.CPDV.OBJECT
- In HCPIST ASSEMBLE at label HCPISTRD:

```
RAMSTART DC    XL4'1800000'  Start RAM Disk at 24MB line
RAMSIZE  DC    F'414208000'  Size in bytes of RAM Disk
```

- This is 809000 512-byte blocks, 395 MB
- The disk on the original install system:
  - 363400 blocks, 177 MB

There is room for a larger disk!

<u>What can we do with a 395 MB disk?</u>

- Add
  - New install code
  - Networking (really?)
  - More users to run background tasks
  - Operations Manager for z/VM
  - Automatic LUN discovery

- We dream of more, but it is a pretty small disk!
  - My dreams..
    - A Linux system hosting a GUI
    - SSL server enabling secure connections

Dumping the System

How can we do this?

Minidisk by minidisk? Or..
Entire volumes?

- I have PIPEDDR EXEC to dump entire volumes
- I don't have development's minidisk dumper

- The restore process uses Pipelines:
  dvdload | unpack | eckdrest. (load from server)
  dvddecod | unpack| eckdrest (from minidisk)

- Can I somehow do this for entire volumes?
  PIPE dvdload | pipeddr

  (PIPEDDR will unpack or uncompress)

"dvdload" uses Diag 2C0 to fetch files

- An API to an undocumented interface
- Documented in *CP Programming Services*
  - Appendix F. Reserved DIAGNOSE Codes
    Diag x'2C0' HMC Data Source Load
  - Each file is 4 MB (maximum) size
    - written to a memory buffer

- dvdload fetches a sequence of files
  - 6 character name, 2 character sequence
- Files have 4 byte record length then record
  - (just like variable length records.)
- Special end of file string.
- Fetch files until EOF. Simple!

Let’s Make This Happen!

- Modify PIPEDDR to call as Pipeline stage
  - The author agreed to it!
- Create files that DVDLOAD understands
  - Add 4 byte record length (“cms4” format)
  - Maximum 4 MB size
  - Write in sequence
  - Add End of File string
- SMOPP
  - Simple Matter of Pipeline Programming
  - Create stage named TODVD REXX
  - Takes PIPEDDR records, adds length, stops after 4 MB, writes file to ftp server
  - At End of File, write EOF string

PIPE rexx (pipeddr exec) dump …
  | todvd <basename> ftp://…

- Each file is 4 MB or less
- Create 8 character filenames; first 6 are base name, last 2 are sequence.
- Note: “base 36” sequence number

How do we restore a volume?

PIPE dvdload <basename>
  | rexx (pipeddr exec) restore …

- DVDLOAD uses Diag 2C0 to read a file
- Continues reading files in sequence
- Stops when it sees End of File string

Customizing the Cloned System

The system must be tailored to the environment

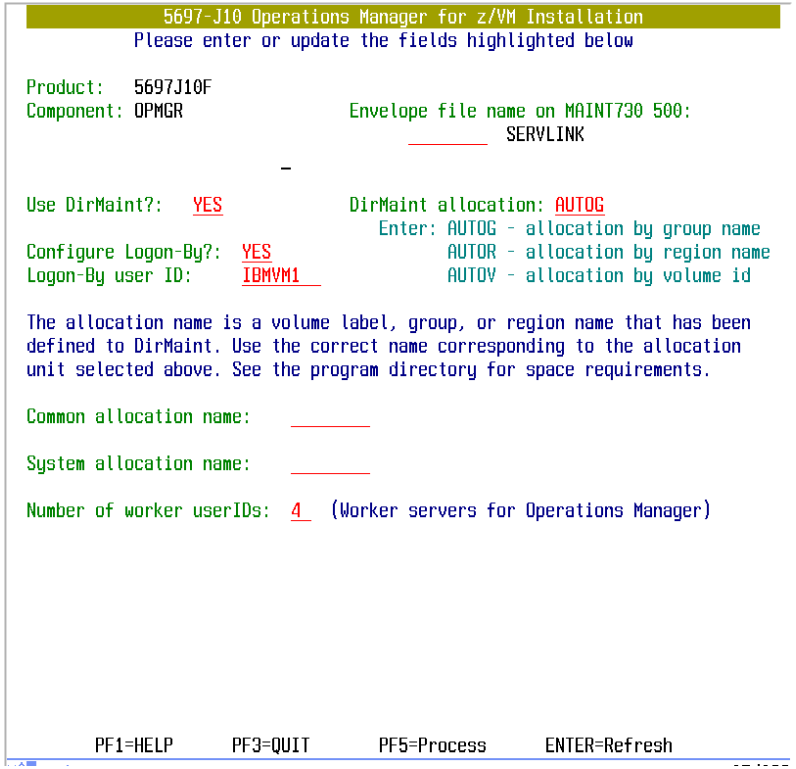
What may need to be changed after the restore?

- Real addresses (DASD and OSA)
- Disk labels
- System name
- IPL parameters
- IP info (IP address, VLAN, etc.)

Other changes we can make

- Add paging and spooling disks
- Add disks ready for Linux workloads
- Adjust to different size DASD or LUNs

- I developed an XEDIT based panel for INSTPROD EXEC; use that idea
- Ask for the values to be customized
- Verify values as valid, issue errors, etc.



Network customization

Set up a network connection on the Installer system

On a newly installed system, you run IPWIZARD

- IPWIZARD asks for:
  - Host name and domain
  - OSA address
  - VLAN id (if needed)
  - IP address
  - Netmask
  - MTU
  - Gateway IP address
  - DNS addresses
- It creates a basic network connection
  - No vswitch is configured
  - Telnet is enabled
  - The configuration is tested

<u>We can use this!</u>

Add users, minidisks, and files so that IPWIZARD can run on the installer system.

- We don’t need all of TCPIP
  - A bit of trial and error to see what is needed
- Once this runs successfully we have
  - A working IP connection
  - The values in IPWIZARD $FILE$ on 2CC

The installed system is configured with this info

- A working IP connection on the first IPL
- Uses a vswitch with failover and other networking best practices

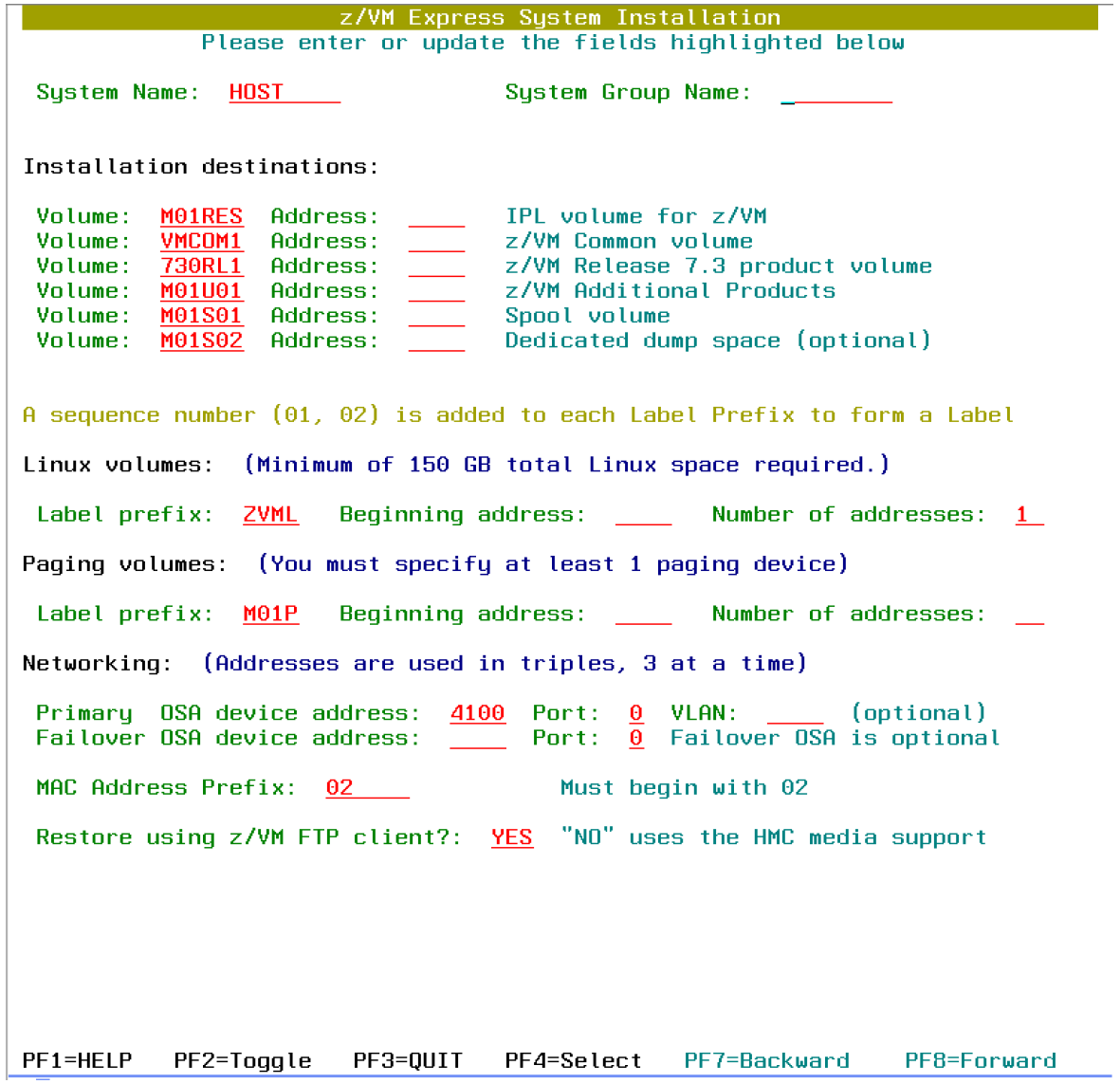
RESTORESYS panel

The information needed to restore and customize the system

The RESTORESYS command shows a 3270 panel

- The system customization information
  - System name
  - DASD real addresses and labels
  - OSA devices
  - Networking VLAN, MAC address prefix
- Ranges of devices can be specified
  - Paging disks
  - Linux workload disk
  - Redhat Openshift Cluster disks (not shown)
  - Input of specific addresses also possible

- System group name is not used by z/VM
  - The Linux guest system uses that value



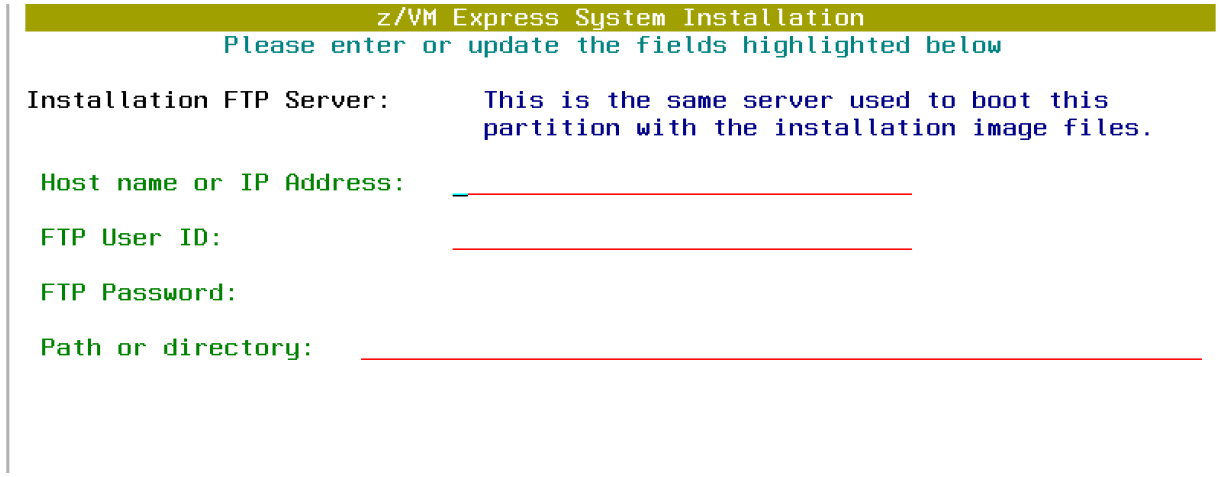
Alternate Restore Option

We can restore using the Pipelines ftp stage

Did you notice?

```
Restore using z/VM FTP client?:  YES  "NO" uses the HMC media support
```

- Because we created a network connection
  - we can use the ftp stage directly to a server
- This is MUCH faster than Diag 2C0
  - Restore in less than 10 minutes vs. an hour or so
- Some additional fields if this is enabled:



IPWIZARD, the sequel

- IPWIZARD works, but it is several screens
- I had a few more requirements:
  - Ask for a Linux IP address
  - Force layer 3
  - Ask if Telnet server should be disabled

- I created a single panel to ask all in one place
  - Called SETUPNET EXEC
- It still calls DTCIPWIZ EXEC for the setup
  - Same thing that IPWIZARD does
  - It still creates IPWIZARD $FILE$ C

What is Customized?

Disks:
- Format page and spool
- Label volumes and set owner
- Erase or release residual data at end (ECKD)
- Update IPL parms on RES disk

System config:
- Update DASD labels
- Add spool or page as specified
- Add Edevice defs (SCSI)
- Update vswitch (OSA addresses, VLAN)
- MAC prefix
- System name
- Threads (1 or 2), STP enabled (if available)
- CP variables with our z/VM ESI level

Directory:
- DASD labels
- VLAN (if used)
- Update the source as USER INPUT (erase Dirmaint's USER DIRECT file)
- Write the object directory

Dirmaint:
- Create EXTENT CONTROL

VMSES/E:
- Change the system name

TCPIP:
- IP address, MTU, netmask, default route
- TCPIP DATA: Host name and domain
- Name server addresses

The Install Process, Start to Finish

Let's show each step in the process

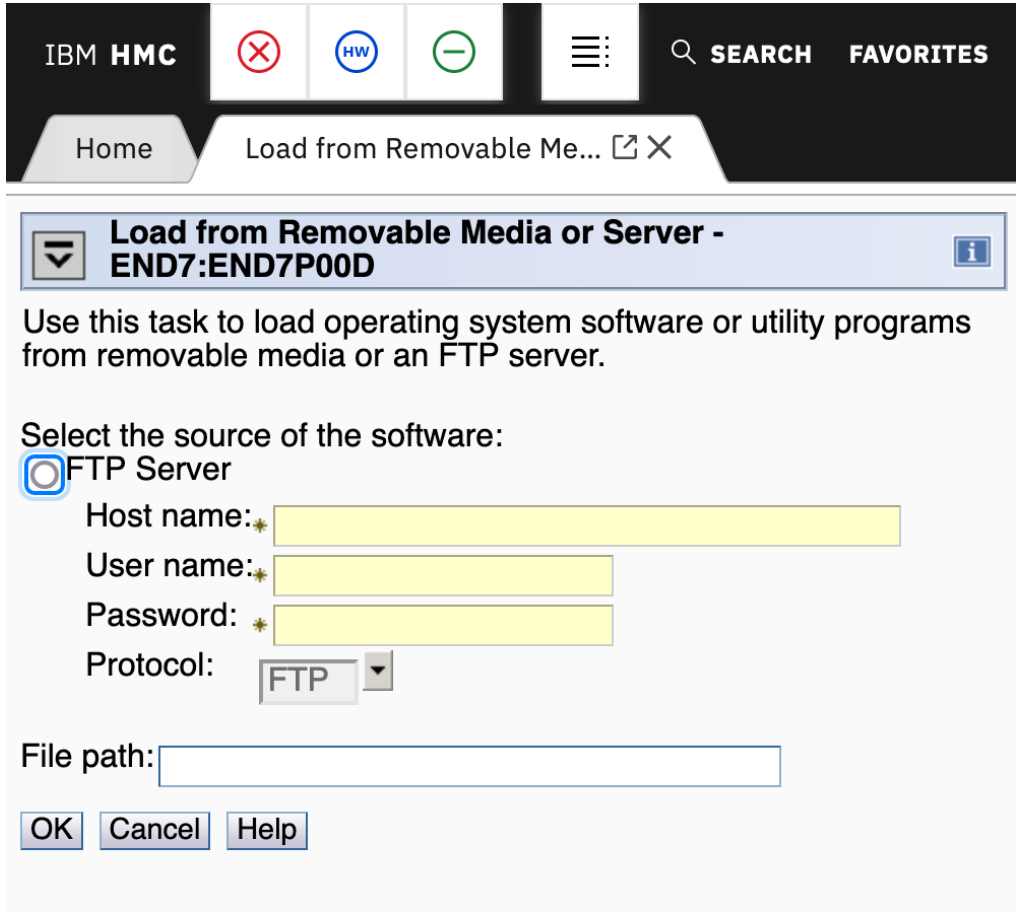
Assumptions:

- You have an LPAR defined and ready, or
- A Partition created (DPM mode)
- You have the installation files either
  - On a server accessible by the HMC
  - Loaded as an ISO file (DPM mode only)
  - On a USB inserted into the HMC

To start:

– "Load from Removable Media or Server"

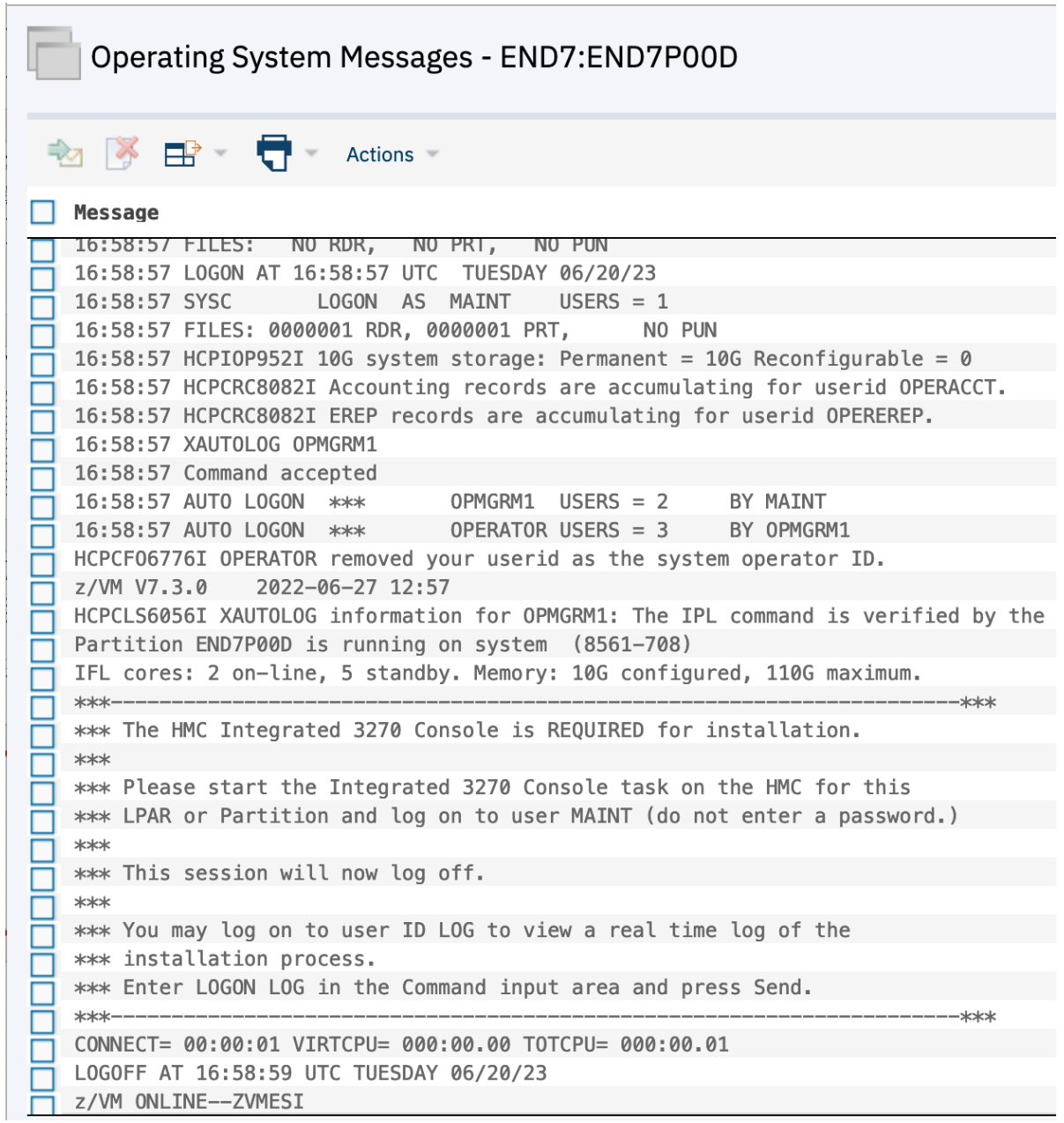
- Protocols FTP, SFTP, and FTPS supported



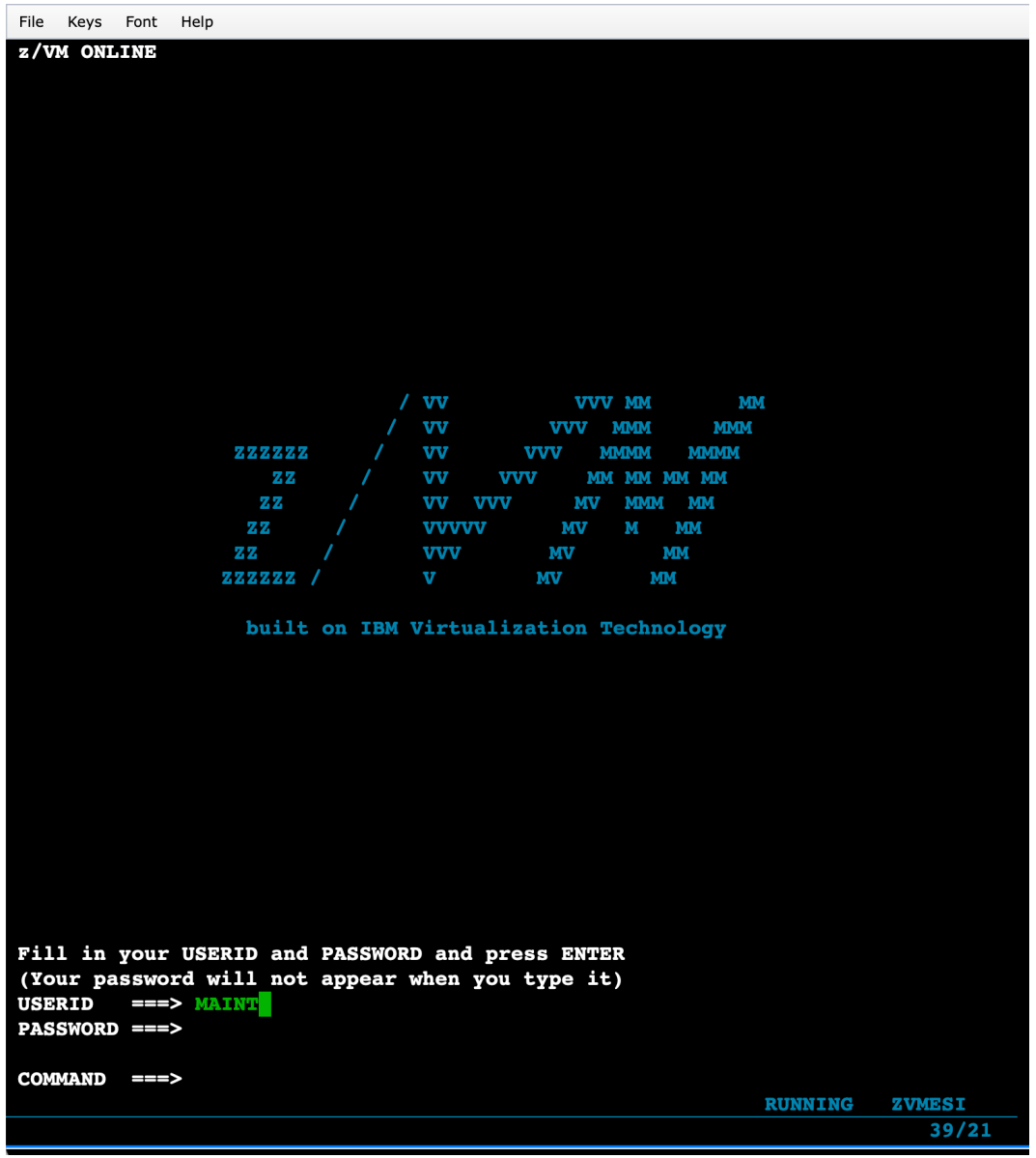
- The Integrated 3270 Console is required
  - It displays the fields and allows input
  - This is just like a normal z/VM install

- You are not required to start it first
  - If you forget, these messages are shown
    - Operating System Messages task
  - Start up the 3270 and log on
  - The user id is MAINT

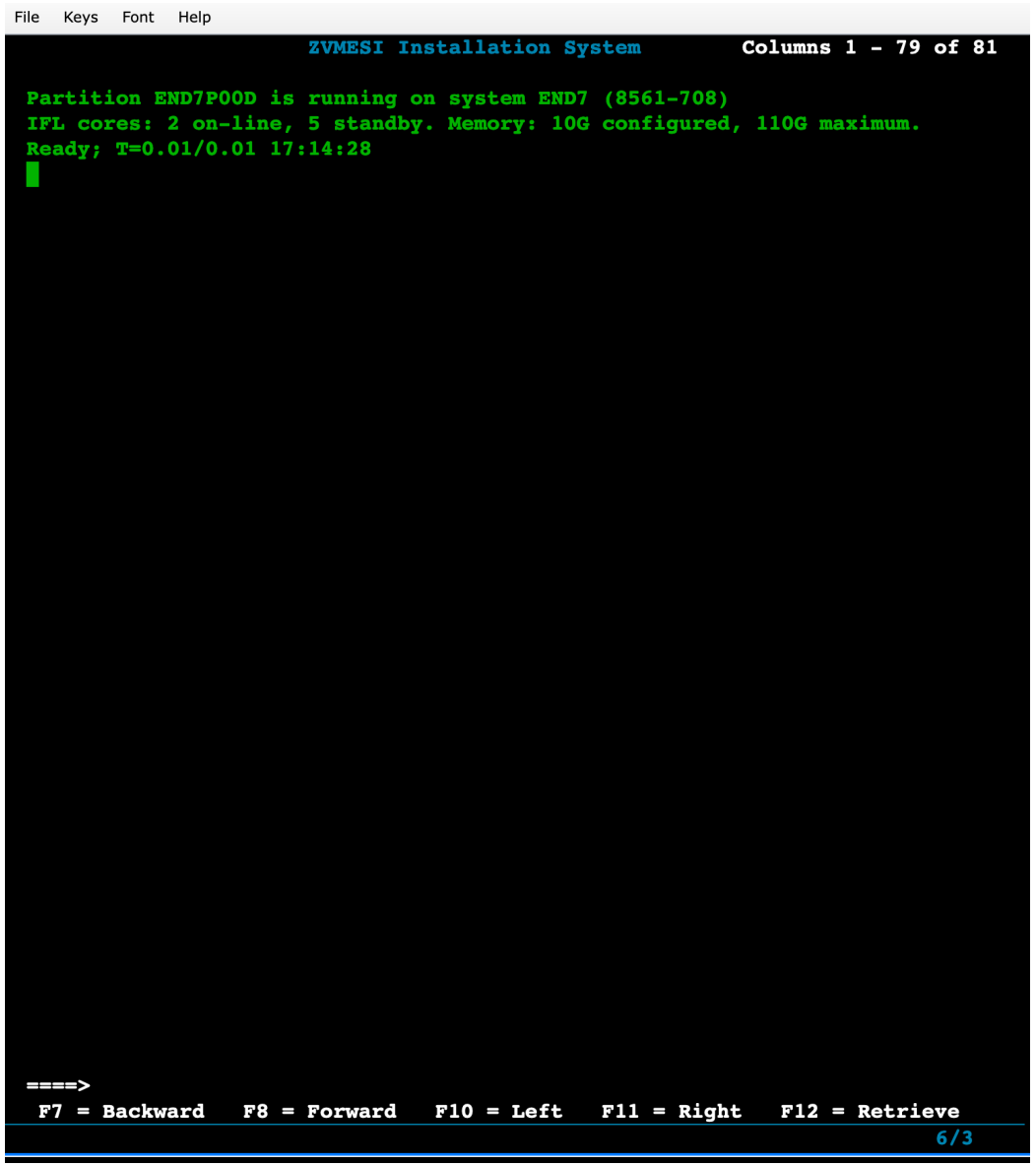
- Notice that you can log on to user id LOG here
  - This is truly a line mode interface
    - Remember the typewriter consoles?
  - The install process will send messages to this console to track progress
  - This is helpful if there is a login timeout on your HMC



The Integrated 3270 Console screen

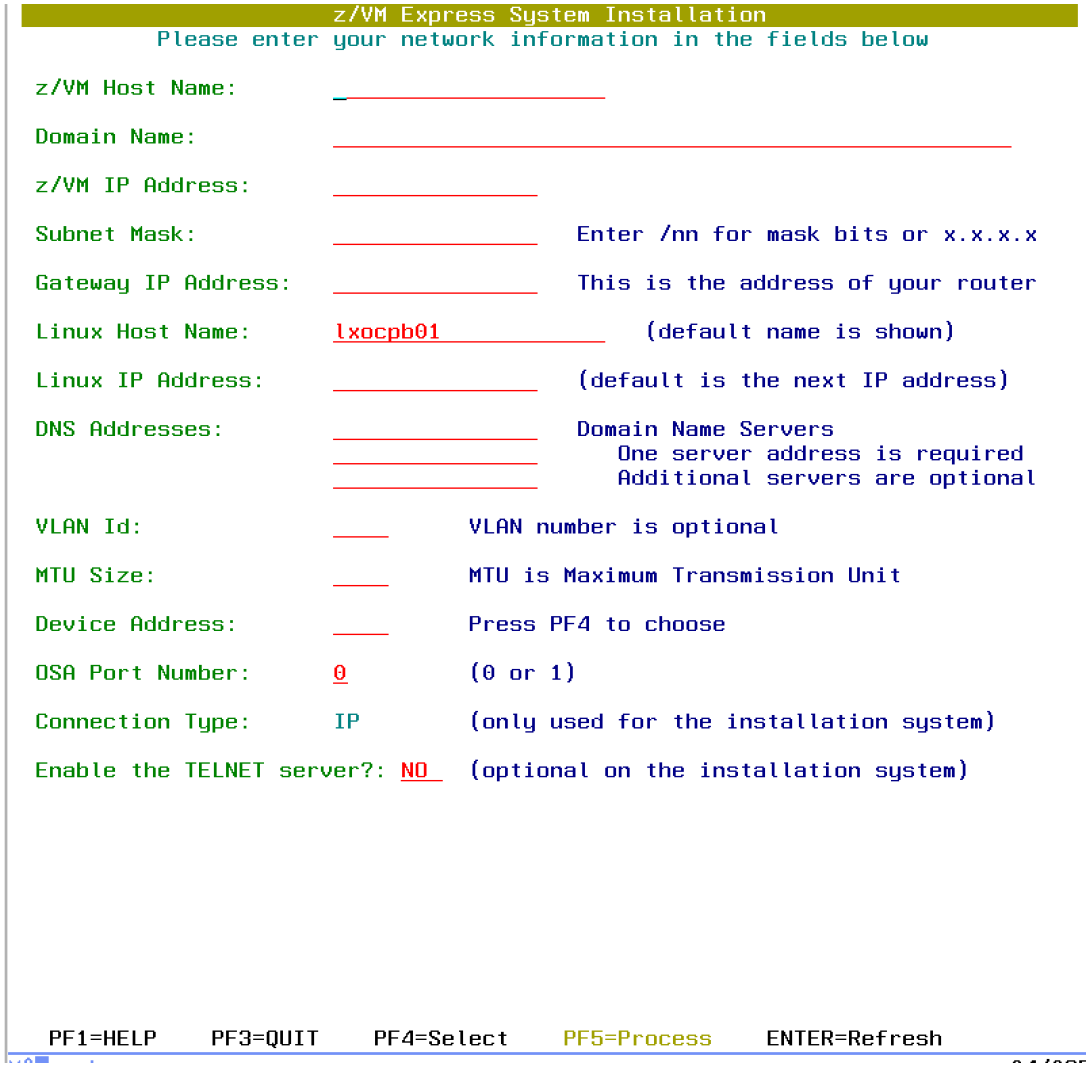


Notice that Fullscreen CMS is used



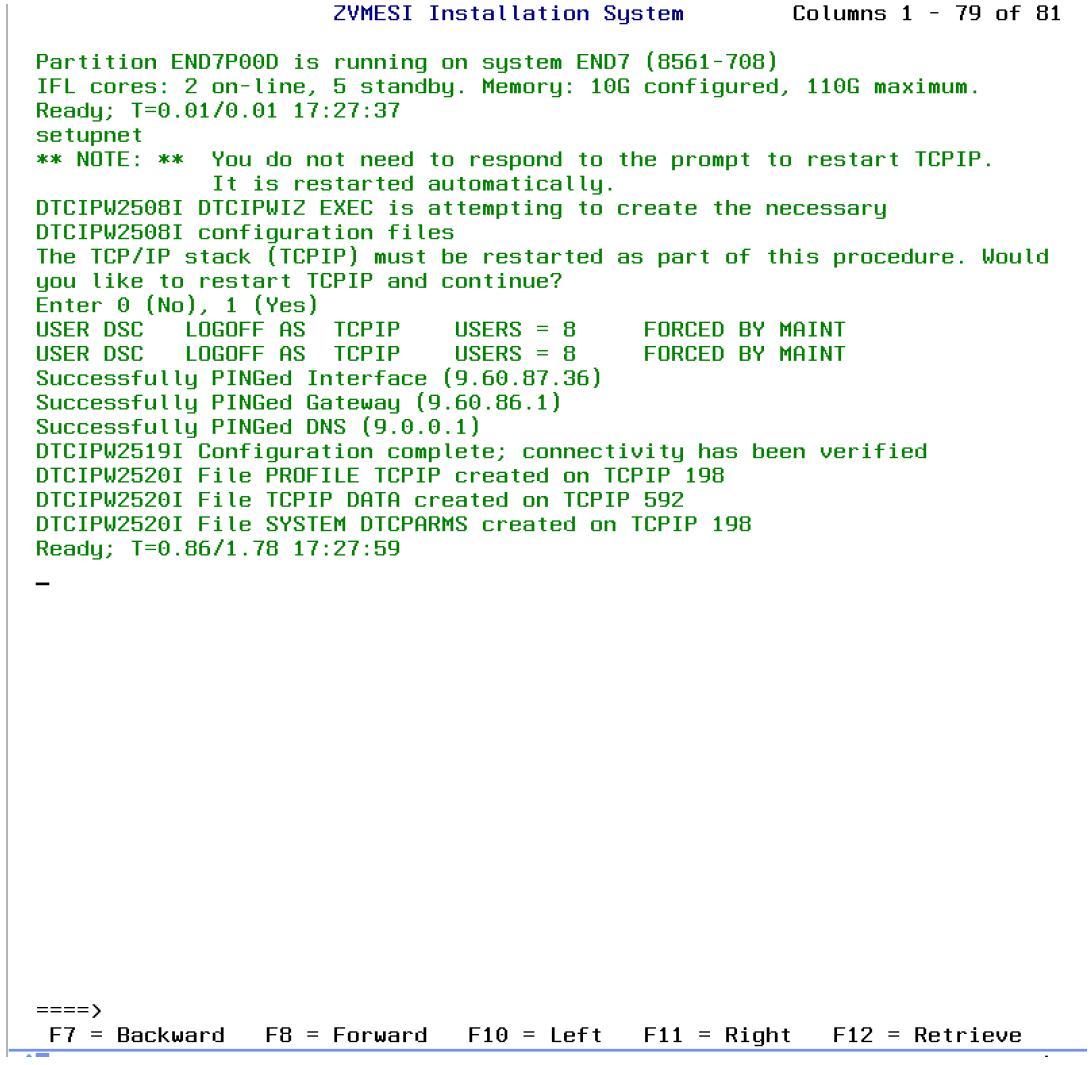
Setting up the network – Run SETUPNET

- All the values are on 1 screen
- IPWIZARD can also be used



After the values are entered, Press PF5

- DTCIPWIZ sets up TCPIP
- The connections are validated



Start the restore process – Enter RESTORESYS

Make a choice on the type of system

```
                    ZVMESI Installation System        Columns 1 - 79 of 81

Partition END7P00D is running on system END7 (8561-708)
IFL cores: 2 on-line, 5 standby. Memory: 10G configured, 110G maximum.
Ready; T=0.01/0.01 17:27:37
setupnet
** NOTE: **  You do not need to respond to the prompt to restart TCPIP.
             It is restarted automatically.
DTCIPW2508I DTCIPWIZ EXEC is attempting to create the necessary
DTCIPW2508I configuration files
The TCP/IP stack (TCPIP) must be restarted as part of this procedure. Would
you like to restart TCPIP and continue?
Enter 0 (No), 1 (Yes)
USER DSC    LOGOFF AS   TCPIP     USERS = 8      FORCED BY MAINT
USER
Succe
Succe
Succe
DTCIP
DTCIP
DTCIP
DTCIP
Ready
resto
DMSAC
DMSAC
```

```
                    z/VM Express System Installation

  Enter a selection for the intended workload of this system:

  1. Red Hat OpenShift Container Platform
  2. IBM Cloud Infrastructure Center
  3. Other Linux based workload

  Enter your selection (1 to 3):  _
```

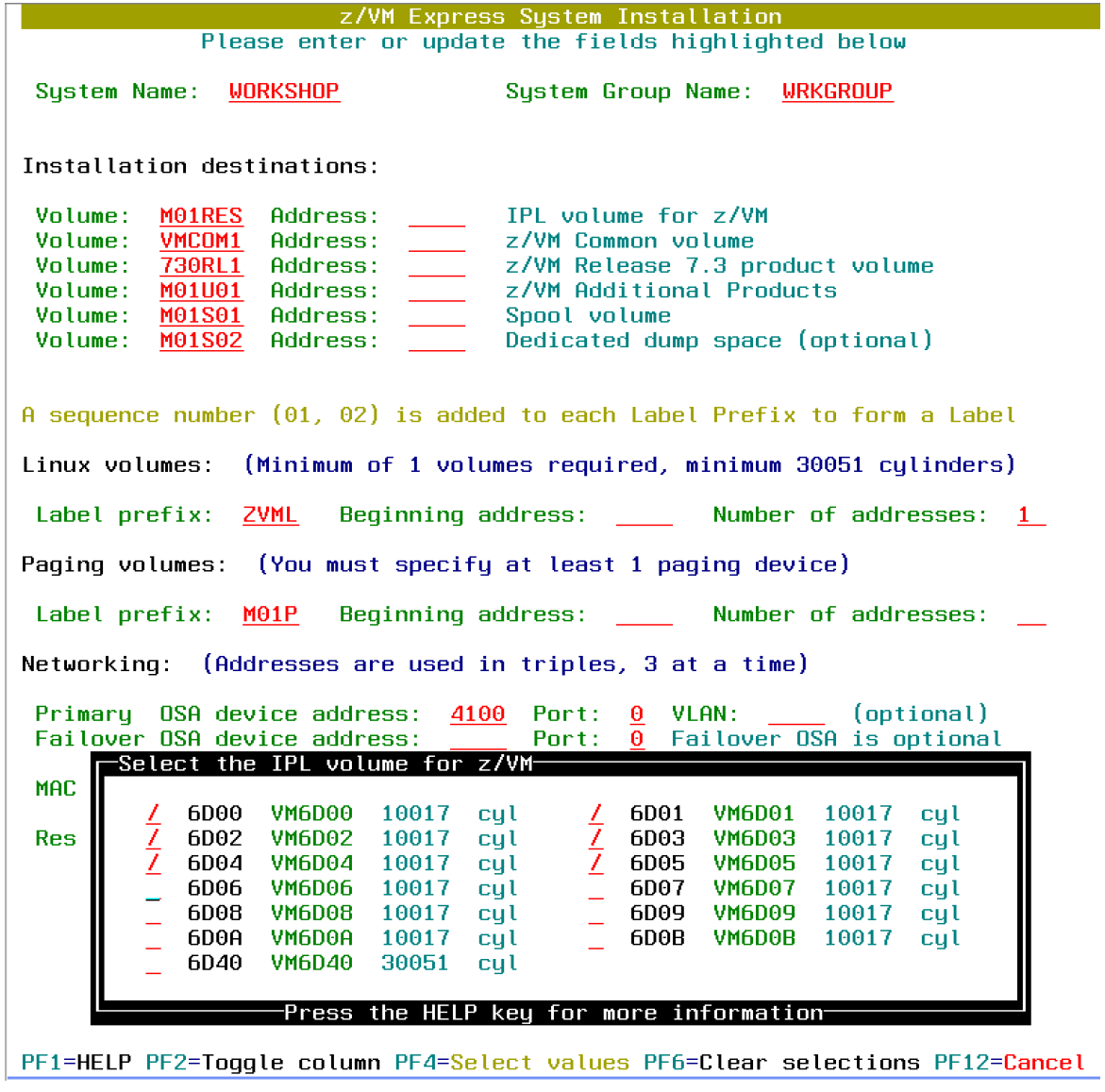
```
====>
 F7 = Backward    F8 = Forward    F10 = Left    F11 = Right    F12 = Retrieve
```

The panel is shown

```
                    z/VM Express System Installation
              Please enter or update the fields highlighted below

 System Name:  _________              System Group Name:  _________


Installation destinations:

 Volume:  M01RES  Address:  ____   IPL volume for z/VM
 Volume:  VMCOM1  Address:  ____   z/VM Common volume
 Volume:  730RL1  Address:  ____   z/VM Release 7.3 product volume
 Volume:  M01U01  Address:  ____   z/VM Additional Products
 Volume:  M01S01  Address:  ____   Spool volume
 Volume:  M01S02  Address:  ____   Dedicated dump space (optional)


A sequence number (01, 02) is added to each Label Prefix to form a Label

Linux volumes:  (Minimum of 1 volumes required, minimum 30051 cylinders)

 Label prefix:  ZVML   Beginning address:  ____   Number of addresses:  1_

Paging volumes:  (You must specify at least 1 paging device)

 Label prefix:  M01P   Beginning address:  ____   Number of addresses:  __

Networking:  (Addresses are used in triples, 3 at a time)

 Primary  OSA device address:  4100  Port:  0  VLAN:  ____  (optional)
 Failover OSA device address:  ____  Port:  0  Failover OSA is optional

 MAC Address Prefix:  02____          Must begin with 02

 Restore using z/VM FTP client?:  NO_  "NO" uses the HMC media support


PF1=HELP    PF2=Toggle    PF3=QUIT    PF4=Select    PF5=Process    ENTER=Refresh
```

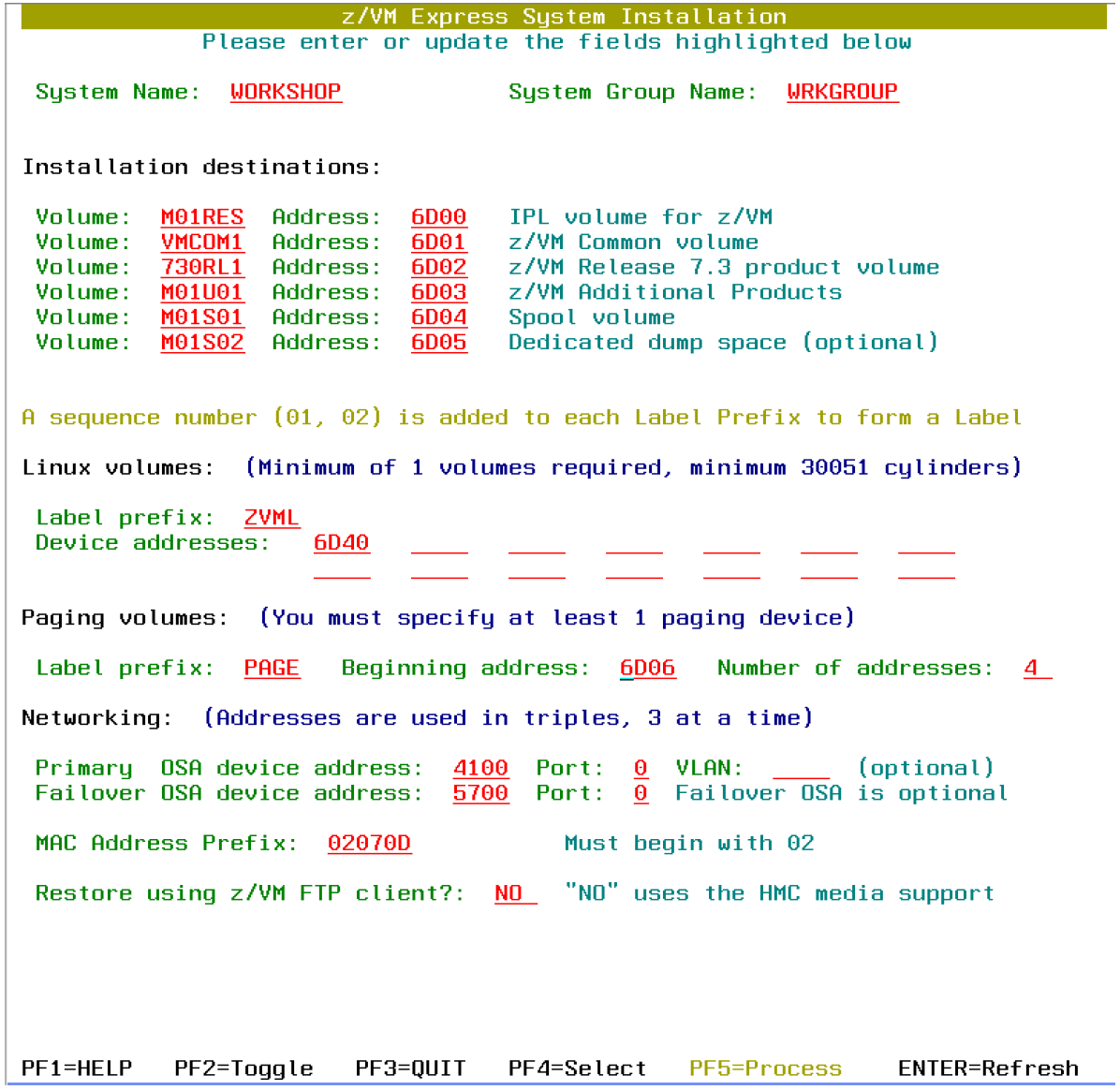
Devices can be selected from a menu

Multiple devices can be selected



DASD for paging and Linux can have multiple disks

Either a sequence of addresses or enter specific ones



If you make an error, a message is shown

```
                    z/VM Express System Installation
            Please enter or update the fields highlighted below

 System Name:  WORKSHOP            System Group Name:  WRKGROUP


Installation destinations:

 Volume:  M01RES  Address:  6D00   IPL volume for z/VM
 Volume:  VMCOM1  Address:  6D01   z/VM Common volume
 Volume:  730RL1  Address:  6D02   z/VM Release 7.3 product volume
 Volume:  M01U01  Address:  6D03   z/VM Additional Products
 Volume:  M01S01  Address:  6D04   Spool volume
 Volume:  M01S02  Address:  6D05   Dedicated dump space (optional)


A sequence number (01, 02) is added to each Label Prefix to form a Label

Linux volumes:  (Minimum of 1 volumes required, minimum 30051 cylinders)

 Label prefix:  ZVML
 Device addresses:   6D44   ____   ____   ____   ____   ____   ____
                     ____   ____   ____   ____   ____   ____   ____

Paging volumes:  (You must specify at least 1 paging device)

 Label prefix:  PAGE   Beginning address:  6D06   Number of addresses:  4

Networking:  (Addresses are used in triples, 3 at a time)

 Primary  OSA device address:  4100  Port:  0  VLAN:  9876  (optional)
 Failover OSA device address:  5700  Port:  0  Failover OSA is optional

 MAC Address Prefix:  02____          Must begin with 02

 Restore using z/VM FTP client?:  NO   "NO" uses the HMC media support

 VLAN number must not be greater than 4095
 MAC address prefix must be a 6 digit value starting with 02
 Linux device 1 device 6D44 does not exist

PF1=HELP   PF2=Toggle   PF3=QUIT   PF4=Select   PF5=Process   ENTER=Refresh
```

When it is correct, press PF5 to process

DASD list is shown for verification

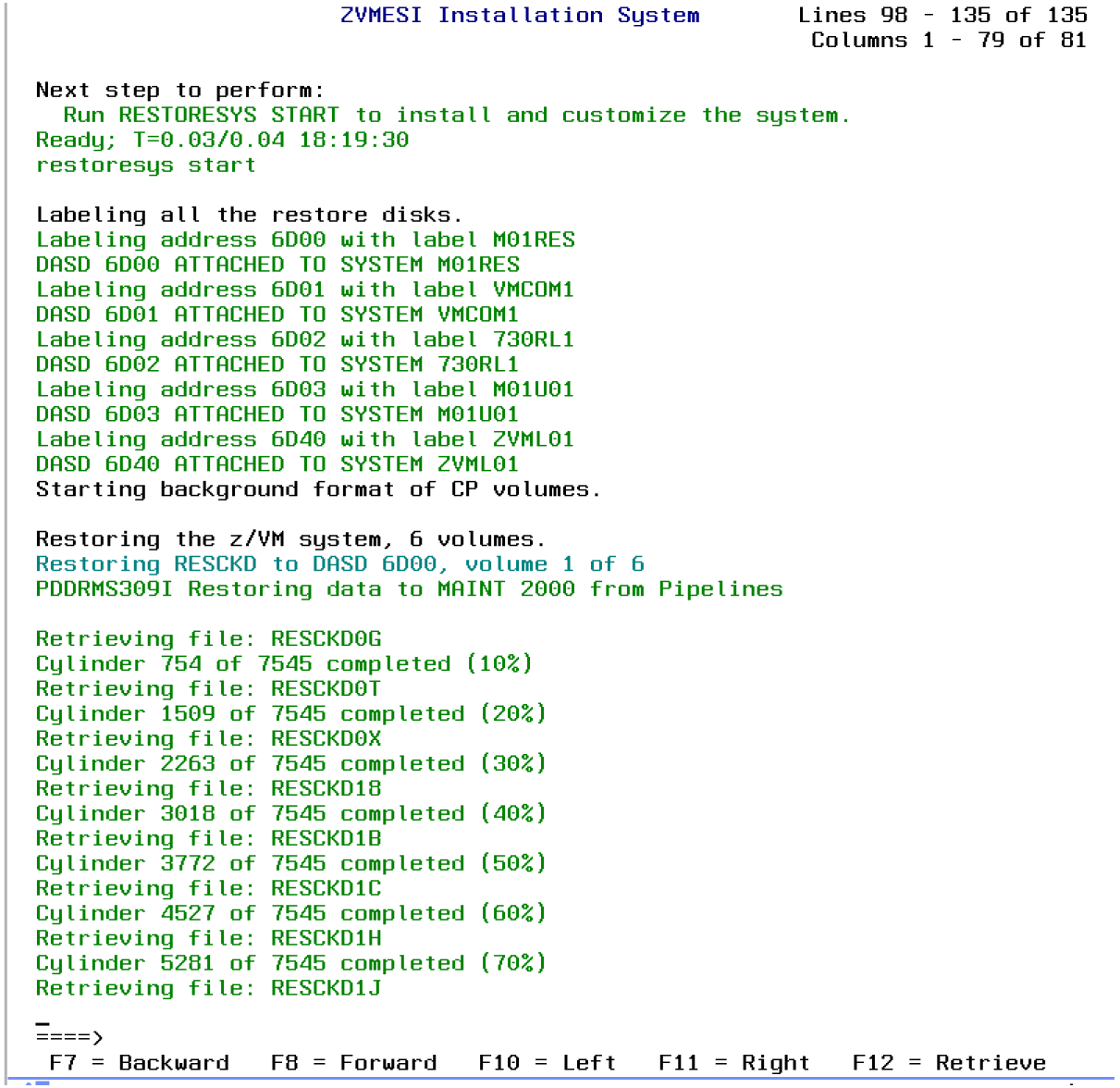
```
                        ZVMESI Installation System      Lines 75 - 101 of 101
                                                        Columns 1 - 79 of 81
Installation selections and values are saved.

Partition END7P00D is running on system END7 (8561-708)
IFL cores: 2 on-line, 5 standby. Memory: 10G configured, 110G maximum.

If these addresses are incorrect, run RESTORESYS CONFIG and make corrections.

These volumes will be initialized and the basic system is restored to:
ECKD DASD  6D00 VM6D00, new label will be M01RES
ECKD DASD  6D01 VM6D01, new label will be VMCOM1
ECKD DASD  6D02 VM6D02, new label will be 730RL1
ECKD DASD  6D03 VM6D03, new label will be M01U01
ECKD DASD  6D04 VM6D04, new label will be M01S01

These volumes will be initialized and formatted for Paging or Dump space:
ECKD DASD  6D05 VM6D05, new label will be M01S02
ECKD DASD  6D06 VM6D06, new label will be PAGE01
ECKD DASD  6D07 VM6D07, new label will be PAGE02
ECKD DASD  6D08 VM6D08, new label will be PAGE03
ECKD DASD  6D09 VM6D09, new label will be PAGE04

These volumes will be initialized and allocated for the Linux:
ECKD DASD  6D40 VM6D40, new label will be ZVML01

Next step to perform:
  Run RESTORESYS START to install and customize the system.
Ready; T=0.03/0.04 18:19:30




====>
 F7 = Backward   F8 = Forward   F10 = Left   F11 = Right   F12 = Retrieve
```

Begin the Restore and Customization Process

- All the disks are labeled with the new labels
- Space for Paging and Spool are formatted
  - This is done by other virtual machines
  - Done in parallel with the restore process

- Each volume is restored
  - The progress of each volume is shown
  - Verification that the process is working

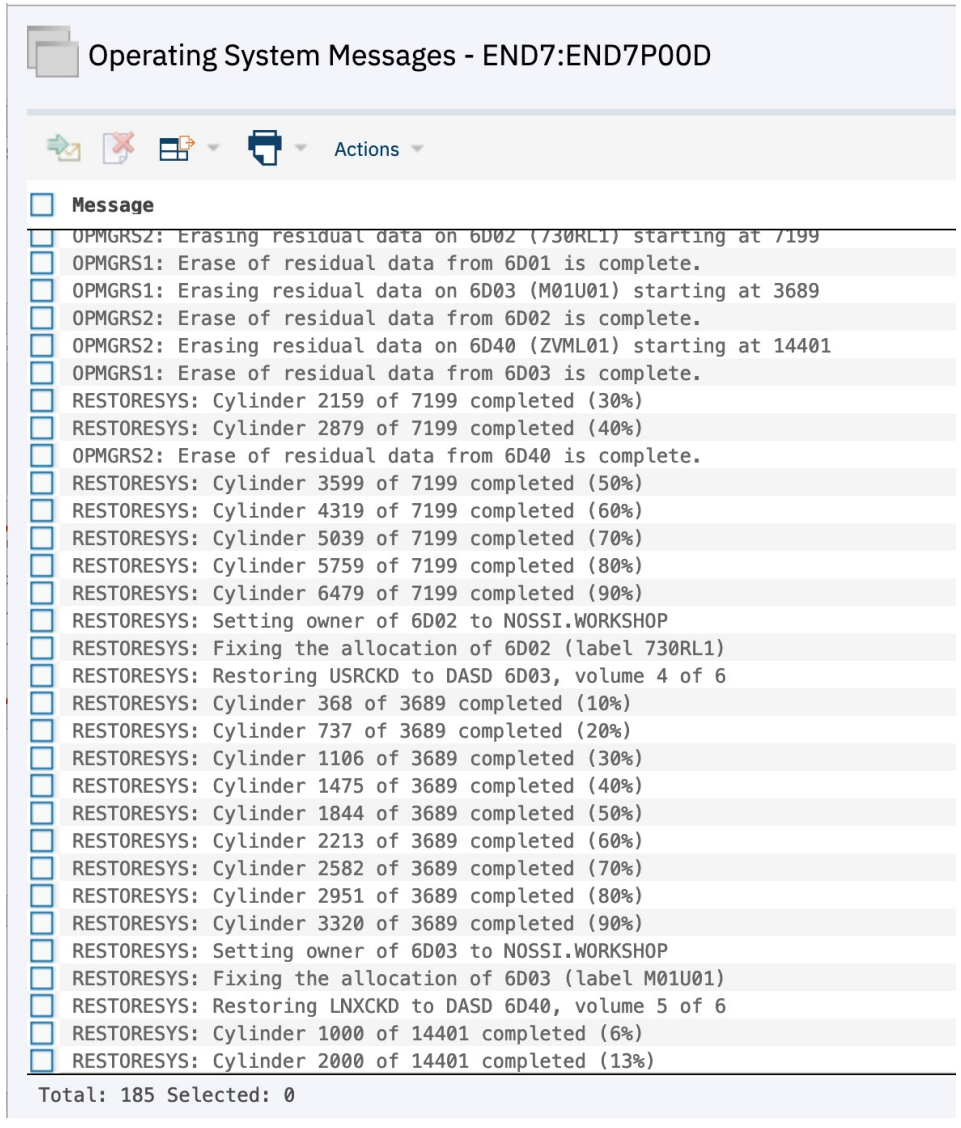
- After each volume is completed
  - The disk allocation is updated
  - The label is rewritten (if needed)
  - The owning system is set



Operating System Messages shows a log

We logged on to a user id named LOG to see this

These are the same messages seen on the 3270 display



Customizing the Restored System

This is done automatically once the restore is complete

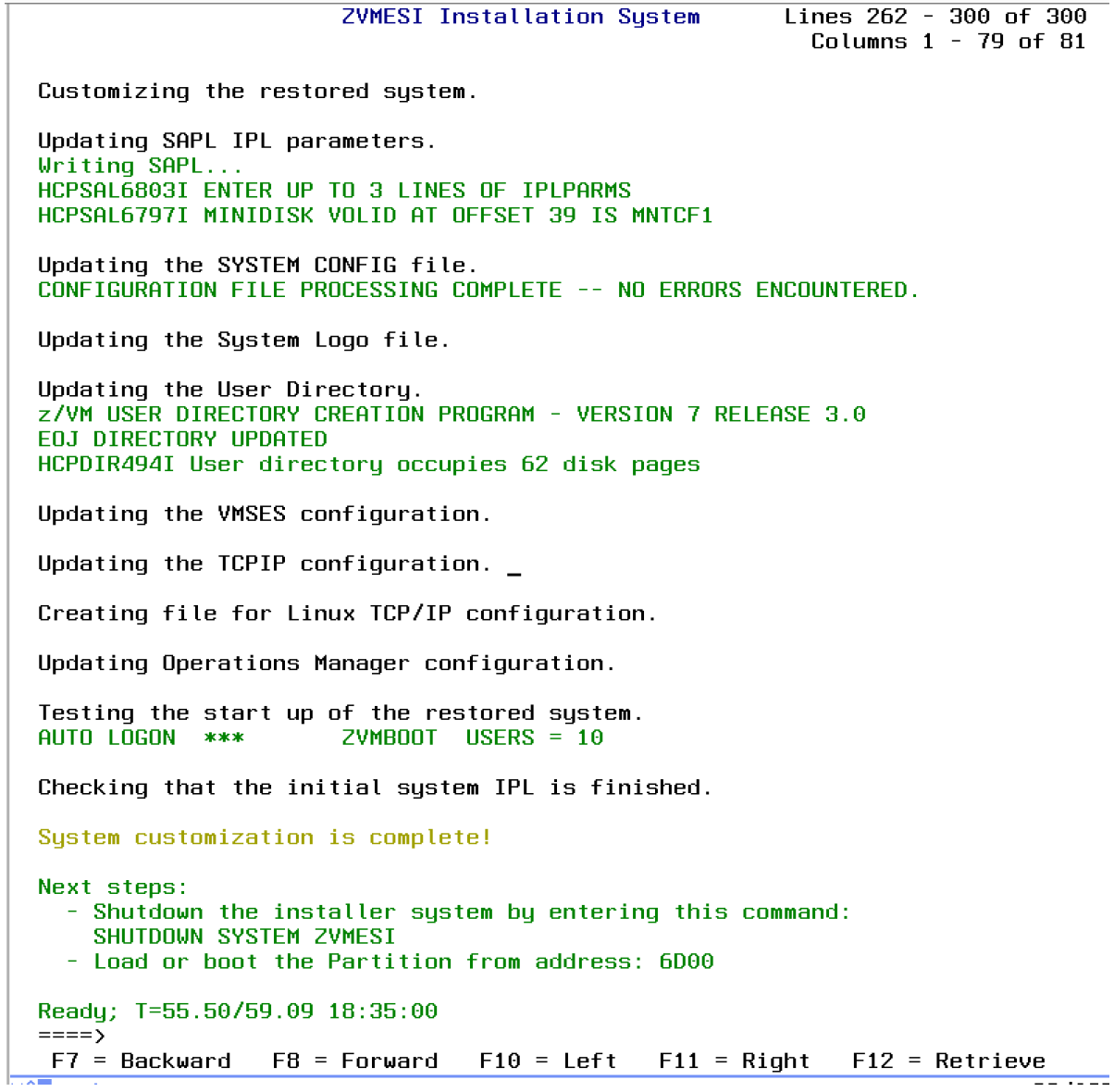
The list of customization tasks was shown earlier

The installed system is booted 2nd level

- The spool disks have changed
  - The size and maybe the label
- It is necessary to perform a FORCE start
  - It starts then shuts down

The final messages help with the next steps

- Shutdown the installer system
- How to IPL the new system



Now IPL Your New System!

Most of us know what that looks like

Some details about the restored system

- RACF/VM is active
  - IDs have Password Phrases
  - All are expired
  - List of IDs and initial phrases in document
- Operations Manager for z/VM is installed
  - Monitoring consoles of running VMs
  - Cleaning up spool and log files
- Passwords are removed from several IDs
  - Some are revoked – no need to log on
  - LOGON BY enabled for some others
- CP Monitor data collection is active
  - Collects 5 to 6 days of data
  - Automatically erases old files
- TLS/SSL is configured and active
  - We created our own small CA to issue cert
    - CA is Certificate Authority
    - Certificates can be downloaded from the included Linux system
  - Secure Telnet is enabled
    - Unsecure is allowed for initial logon
  - LDAP is configured

We have a document that describes the differences.

Some Additional Features

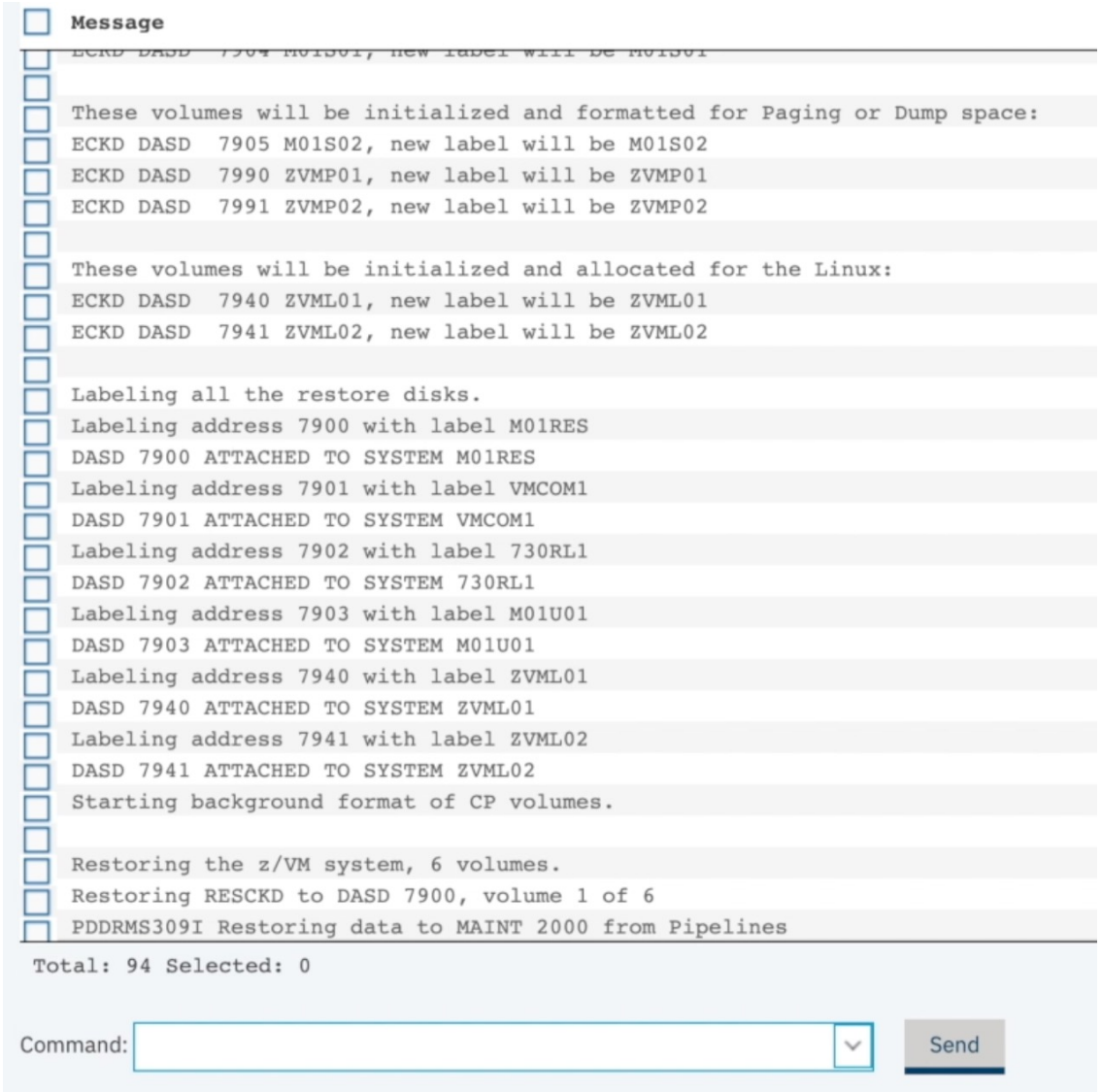
We've added some options for special situations

Configuration using a Spreadsheet

- For planning in advance or multiple installs
  - Fill out a spreadsheet we provide
  - Values are the same as you see on the 3270
  - Save it in CSV format
  - Put it in the same directory as install files
- Fills in the fields on the 3270 panels
  - Otherwise, the process is the same

Then we took that one step further

- Fully automatic installation
- No need to log on to a 3270
  - (assuming all input data is correct)
- See progress on Operating System Messages



What About z/VM SSI?

Our target customers usually install a single system

SSI is great, but setting up the CTCs is harder

- The CTC connections are required!
- We hope DPM will offer configuration help

Converting an “SSI-ready” z/VM system to a single “SSI-enabled” one is documented

- Could it be done automatically?
- Maybe after restoring our system?
- Sure!
  - It needs more disks for RACF
  - SSI cluster name is needed
  - SYSTEM CONFIG and directory changes
  - Create the PDR
  - SFS config, VMSES/E
- Use the cloning process for more members
  - Documented in the z/VM Library

```
                    z/VM Express System Installation
              Please enter or update the fields highlighted below

 System Name:  WORKSHOP            System Group Name:  WRKGROUP
 Enable SSI?:  YES                 SSI Cluster Name:   SHOPSSI

Installation destinations:

 Volume:  M01RES  Address:  ____   IPL volume for z/VM
 Volume:  VMCOM1  Address:  ____   z/VM Common volume
 Volume:  730RL1  Address:  ____   z/VM Release 7.3 product volume
 Volume:  M01U01  Address:  ____   z/VM Additional Products
 Volume:  M01S01  Address:  ____   Spool volume
 Volume:  RACFDB  Address:  ____   RACF Primary database volume
 Volume:  RACFBK  Address:  ____   RACF Backup database volume
 Volume:  M01S02  Address:  ____   Dedicated dump space (optional)

A sequence number (01, 02) is added to each Label Prefix to form a Label

Linux volumes:  (Minimum of 1 volumes required, minimum 30051 cylinders)

 Label prefix:  ZVML   Beginning address:  ____   Number of addresses:  1

Paging volumes:  (You must specify at least 1 paging device)

 Label prefix:  M01P   Beginning address:  ____   Number of addresses:  __

Networking:  (Addresses are used in triples, 3 at a time)

 Primary  OSA device address:  4100  Port:  0  VLAN:  ____  (optional)
 Failover OSA device address:  ____  Port:  0  Failover OSA is optional

 MAC Address Prefix:  02____          Must begin with 02

 Restore using z/VM FTP client?:  NO_  "NO" uses the HMC media support

PF1=HELP   PF2=Toggle   PF3=QUIT   PF4=Select   PF5=Process   ENTER=Refresh
```

Thanks for Listening!

Thanks to our Skunkworks team
- Ernie Horn
- Vic Cross
- Paul Novak
- Bruce Hayden

With help from
- Fred Bader
- Jacob Emery
- Jay Brenneman
- Justice Heughan
- Matt Mondics
- Stephanie Rivero

Be sure to attend our other sessions:
- *Easily Extending a Usable z/VM System!*
  - Friday 2:15 PM Cartoon room
- *Documenting, testing, and packaging of an automated bundle like z/VM ESI*
  - Saturday 9:45 AM Cartoon room

Contact information:
Bruce Hayden
bjhayden@us.ibm.com
IBM Z Washington Systems Center