IBM

Running Linux on IBM System z9 and zSeries under z/VM

Native SCSI support for z/VM

Networking enhancements for Linux guests

Layer 2 network support

Gregory Geiselhart
Klaus Egeler
Lutz Kühner
Markku Kymalainen
Esthon Medeiros
Karl-Erik Stenfors

Redbooks

ibm.com/redbooks

International Technical Support Organization

Running Linux on IBM System z9 and zSeries under z/VM

February 2006

SG24-6311-00

Note: Before using this information and the product it supports, read the information in "Notices" on page v.

First Edition (February 2006)

This edition applies to z/VM Version 5, Release 1 and multiple Linux distributions. SUSE Linux Enterprise 8 (SLES8) and Red Hat Enterprise Linux 3 (RHEL 3) are used for examples in this publication.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of 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.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:
This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

eserver®
pSeries®
xSeries®
z/OS®
z/VM®
zSeries®
z9™
Enterprise Storage Server®
ECKD™
FICON®
HiperSockets™
IBM®
Redbooks™
Redbooks (logo) ™
RACF®
S/390®
System z9™
Tivoli®
TotalStorage®
Virtualization Engine™

The following terms are trademarks of other companies:

IPX, and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States, other countries, or both.

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

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

Other company, product, and service names may be trademarks or service marks of others.

Preface

This IBM® Redbook discusses running Linux® under z/VM® on IBM System z9™ and zSeries® platforms. It describes enhancements introduced in z/VM Version 5.1.

We describe installing z/VM on FCP-attached SCSI disks. Configured as emulated Fixed Block Architecture (FBA) disks, z/VM 5.1 can use FCP-attached disks for its system paging, spooling, directory, and minidisks.

z/VM 5.1 adds new functions for Virtual Switches. For increased network security, guests must have authorization before connecting to a VSWITCH. z/VM 5.1 introduces VSWITCH Layer 2 support. Operating at Layer 2, a VSWITCH delivers and receives network traffic in Ethernet frames. This provides the ability to handle non-IP protocols such as SNA, NetBIOS, and IPX. In addition, Layer 2 support reduces network latency and CPU overhead.

The team that wrote this redbook

This redbook was produced by a team of specialists from around the world working at the International Technical Support Organization, Poughkeepsie Center.

Gregory Geiselhart is a Project Leader for Linux on zSeries at the International Technical Support Organization, Poughkeepsie Center.

Klaus Egeler is an IT Systems Management Specialist with IBM Global Services, Germany. He has more than 15 years of experience as a VSE and VM Systems Programmer. He has worked with Linux for zSeries and S/390® for more than three years. He has contributed to several Linux related IBM Redbooks™.

Lutz Kühner is an IT Specialist in Germany. He has 18 years of experience in the mainframe systems field. His areas of expertise include z/OS® and z/VM system programming. He has co-written a number of Redbooks in his fields of expertise.

Markku Kymalainen is an IT Specialist for IBM Global Services in Finland. He has 17 years of experience in the VM area and 30 years of experience in mainframes. He has been involved in major z/VM and Linux on zSeries installations in Finland from capacity sizing to installation and tuning. In addition to z/VM and Linux on zSeries, he has a good understanding of IBM @server zSeries hardware.

Esthon Medeiros is an IBM Certified Senior IT Specialist in Brazil. He has 25 years of experience in the IT field, mainly in mainframe, z/VM and Linux. As a Linux expert, his activities now support customers implementing Linux projects in mainframes.

Karl-Erik Stenfors is a Senior IT Specialist in the IBM Product and Solutions Support Centre (PSSC) in Montpellier, France. He has more than 30 years of experience in the mainframe systems field, both as a systems programmer and consultant with several IBM customers, since 1986 with IBM. His areas of expertise include zSeries hardware and operating systems (z/OS, z/VM, Linux) and IBM and Open Source middleware. He has co-authored a number of Redbooks in his fields of expertise. Karl-Erik has held two international assignments, and he is currently responsible for Early Support Programmes for zSeries in EMEA. He is member of an international IBM workgroup providing customer and field input to the Poughkeepsie lab to help create a vision for the future of zSeries.

Thanks to the following people for their contributions to this project:

Roy Costa, Bob Haimowitz, Octavian Lascu, Al Schwab
International Technical Support Organization, Poughkeepsie Center

IBM Endicott
Steve Wilkins, Joe Gregor

Dave Jones
Sine Nomine Associates

Jack Hoarau, Gerard Laumay, Sebastien Llaurency
IBM Product and Solutions Support Centre
Montpellier, France

Michael Mai, IBM Germany

Become a published author

Join us for a two- to six-week residency program! Help write an IBM Redbook dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You'll team with IBM technical professionals, Business Partners and/or customers.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you'll develop a network of contacts in IBM development labs, and increase your productivity and marketability.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our Redbooks to be as helpful as possible. Send us your comments about this or other Redbooks in one of the following ways:

- Use the online **Contact us** review redbook form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbook@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYJ Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

1

Installing z/VM 5.1

In this chapter, we describe how to install z/VM 5.1 from DVD to 3390 minidisks, and how to configure the VM TCP/IP stack. Topics include:

- Installation from DVD
- First level installation from DVD
- Second level installation from DVD
- Installation to FCP-attached SCSI disk
- TCP/IP configuration

1.1 Installation from DVD

Installation from DVD is now possible with z/VM 5.1. You can either use the Hardware Management Console (HMC) equipped with a DVD drive, or a workstation with a DVD drive accessible over FTP. DVD installation requires IBM Hardware Management Console Version, 1.8.0 or later. For details on installing z/VM 5.1 from DVD, consult *z/VM: Guide for Automated Installation and Service*, GC24-6099.

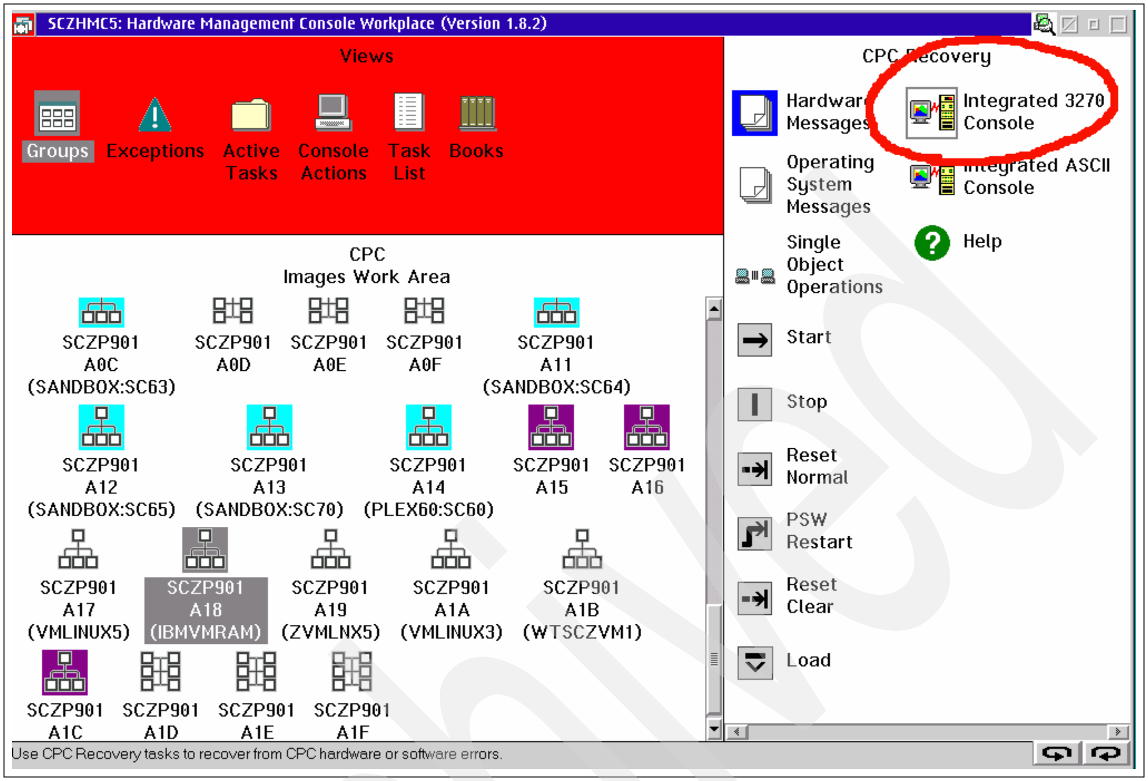
1.2 First level installation from DVD

In this section, we describe first level z/VM 5.1 installation from DVD. The steps involve:

1. Establish an Integrated 3270 Console session
2. Access the primary Support Element
3. Load the z/VM 5.1 RAMDISK
4. IPL the RAMDISK
5. IPL the installed z/VM 5.1 system
6. Apply the Recommended Service Upgrade

1.2.1 Establish an Integrated 3270 Console session

To establish an Integrated 3270 Console session to the LPAR, log on to the HMC as SYSPROG. Select **Task List** (from the **Views** area) →**CPC Recovery** (from the **Task List Work Area**) →**Groups** (from the **Views** area) →**CPC Images** (from the **Groups Work Area**). This takes you to the screen shown in Figure 1-1 on page 3.



Use CPC Recovery tasks to recover from CPC hardware or software errors.

Figure 1-1 Select Integrated 3270 Console option

Select the image where z/VM 5.1 is to be installed from the **CPC Images Work Area** (SCZP901 in this example), then double click **Integrated 3270 Console** from the **CPC Recovery** area.

Note: The Integrated 3270 Console may also be started by dragging the desired CPC image icon onto the **Integrated 3270 Console** icon. Hold the right mouse button to drag the selected CPC image icon.

1.2.2 Access the primary Support Element

To access the Support Element (SE), select **Task List** (from the **Views** area) → **CPC Recovery** (from the **Task List Work Area**) → **Groups** (from the **Views** area) → **Defined CPCs** (from the **Groups Work Area**). This navigates to the panel shown in Figure 1-2 on page 4.

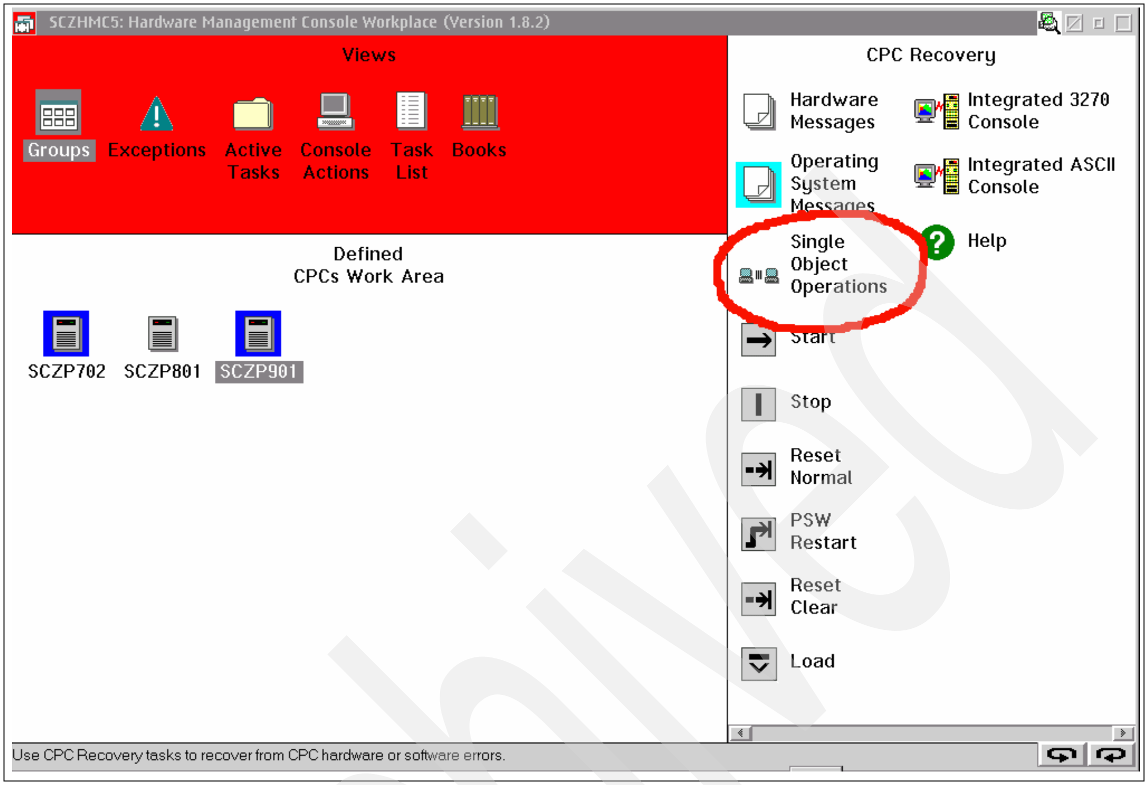
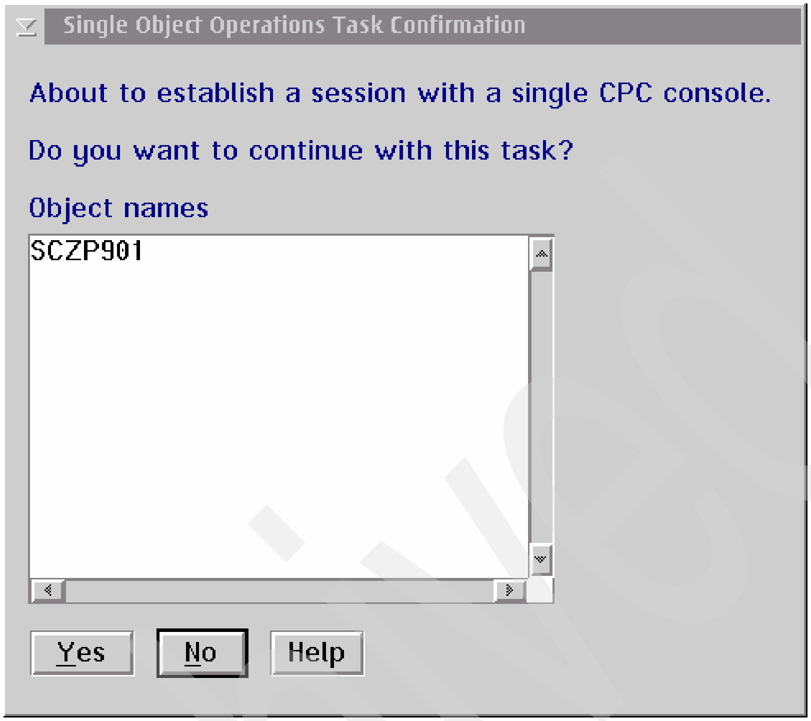


Figure 1-2 Select Single Object Operation option

Select the desired Central Processing Complex (CPC) from the **Defined CPCs Work Area** (SCZP901 in this example), then double click **Single Object Operations** from the **CPC Recovery** area. You are then prompted to confirm your selection in Figure 1-3 on page 5. Select **Yes** to continue.



Single Object Operations Task Confirmation

About to establish a session with a single CPC console.

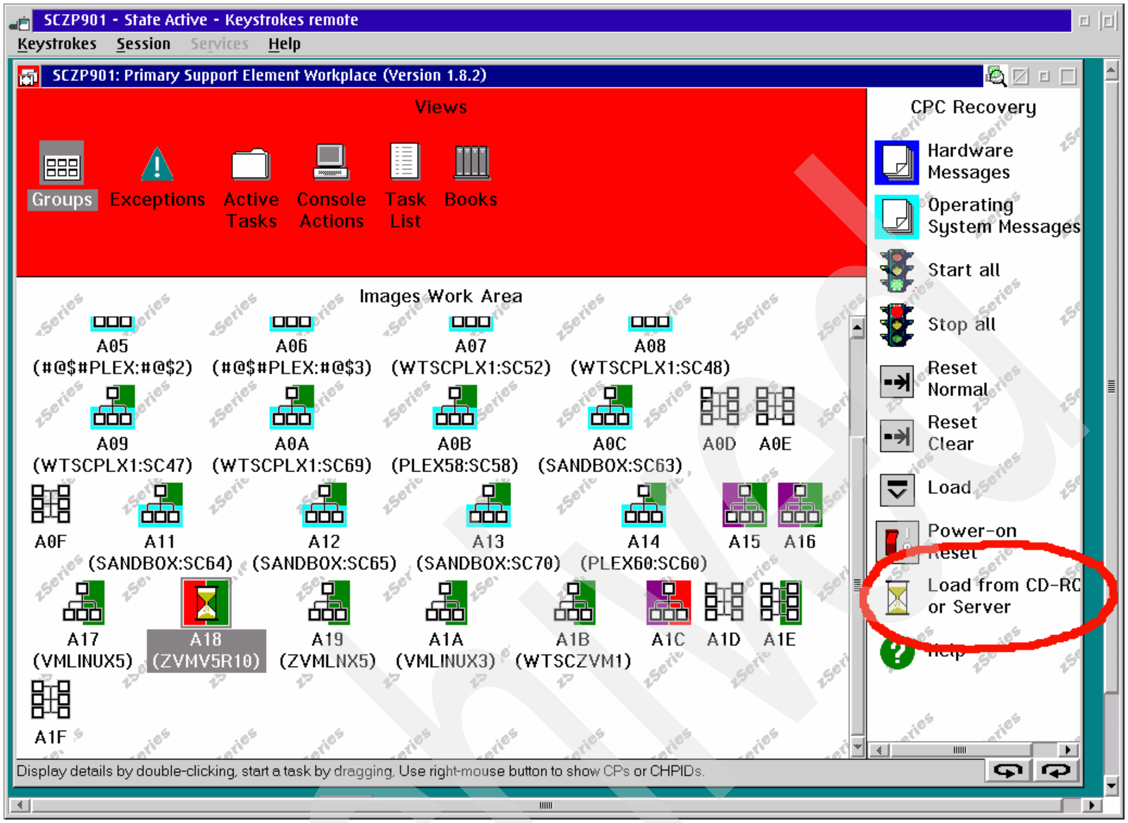
Do you want to continue with this task?

Object names

SCZP901

Figure 1-3 Confirm selection dialog

Select **Groups** (from the **Views** area) →**Images** (from the **Groups View Area**), and navigate to the **CPC Recovery** menu using the arrows in the lower right corner. This takes you to the panel shown in Figure 1-4 on page 6.



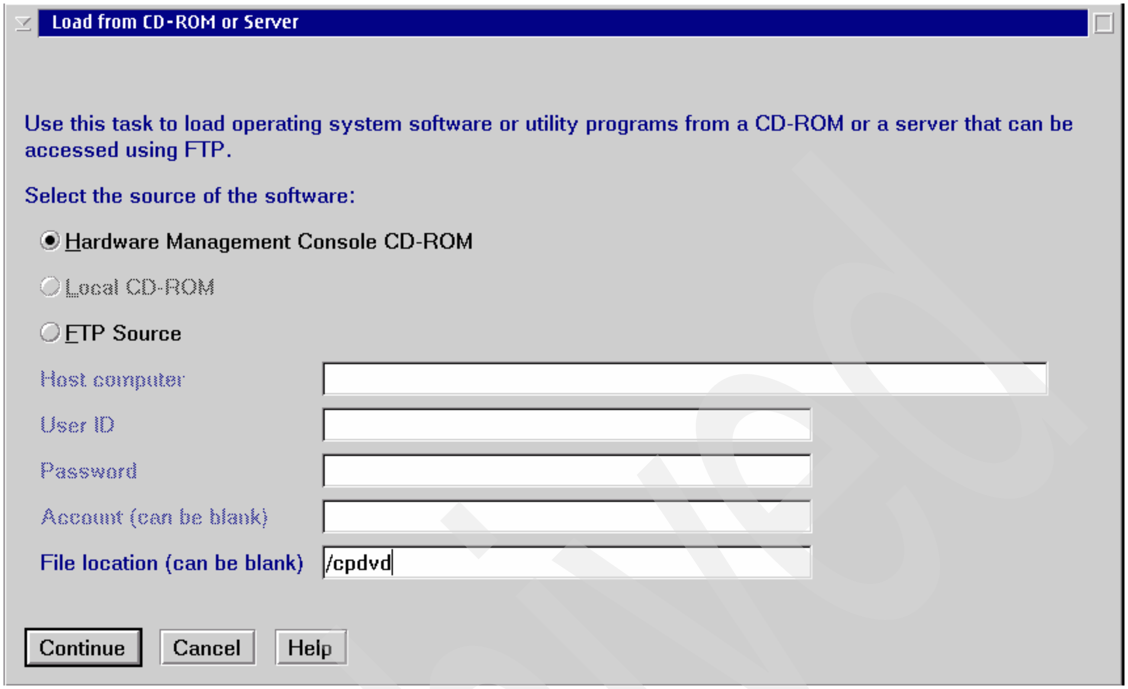
Display details by double-clicking, start a task by dragging. Use right-mouse button to show CPs or CHPIDs.

Figure 1-4 Select the DVD drive for installation

Select the LPAR on which to install (A18 in this example), and double click the **Load from CD-ROM or Server** icon.

1.2.3 Load the z/VM 5.1 RAMDISK

At this point, load the z/VM 5.1 DVD into the DVD drive and return to the panel shown in Figure 1-5 on page 7.


Load from CD-ROM or Server

Use this task to load operating system software or utility programs from a CD-ROM or a server that can be accessed using FTP.

Select the source of the software:

● Hardware Management Console CD-ROM

○ Local CD-ROM

○ FTP Source

Host computer:

User ID:

Password:

Account (can be blank):

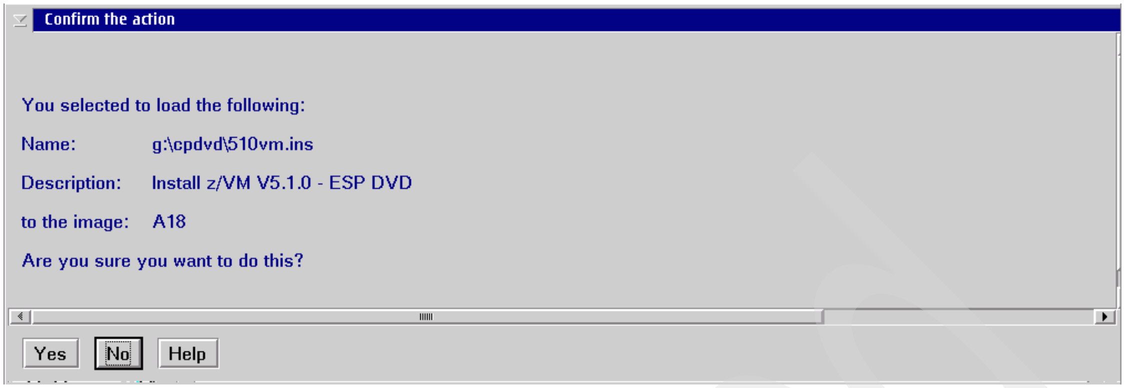
File location (can be blank): /cpdvd

Figure 1-5 Load from CD-ROM or Server panel

We select the **Hardware Management Console CD-ROM** option and provide the directory name where the installation image is found on the DVD (/cpdvd).

> **Note:** It is possible to load the installation image using the DVD drive of a remote workstation. The DVD drive must be accessible from FTP, and a TCP/IP path must exist between the FTP server and the Support Element. In this redbook, we only document installation from the HMC DVD drive.

Click **Continue** and choose the 510vm.ins installation image when prompted. In Figure 1-6 on page 8, you are asked to confirm your selection before proceeding.


Confirm the action

You selected to load the following:

Name: g:\cpdvd\510vm.ins

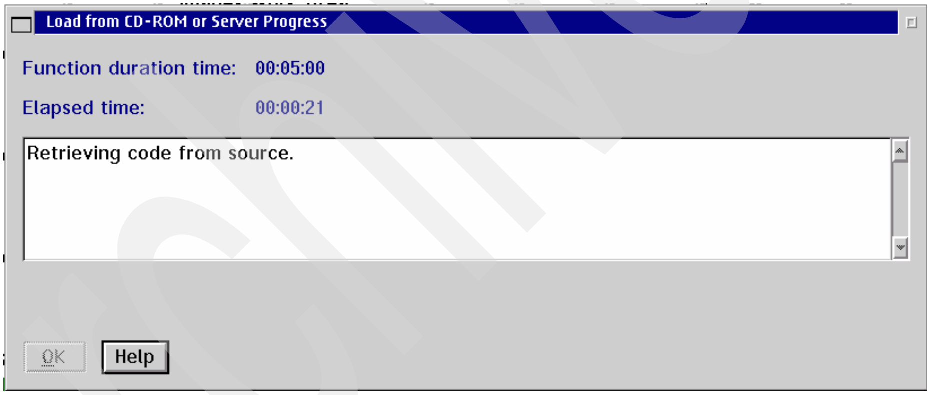
Description: Install z/VM V5.1.0 - ESP DVD

to the image: A18

Are you sure you want to do this?

Figure 1-6 Confirm DVD installation

Hit **Yes** to continue. As installation proceeds, the panel shown in Figure 1-7 is displayed.


Load from CD-ROM or Server Progress

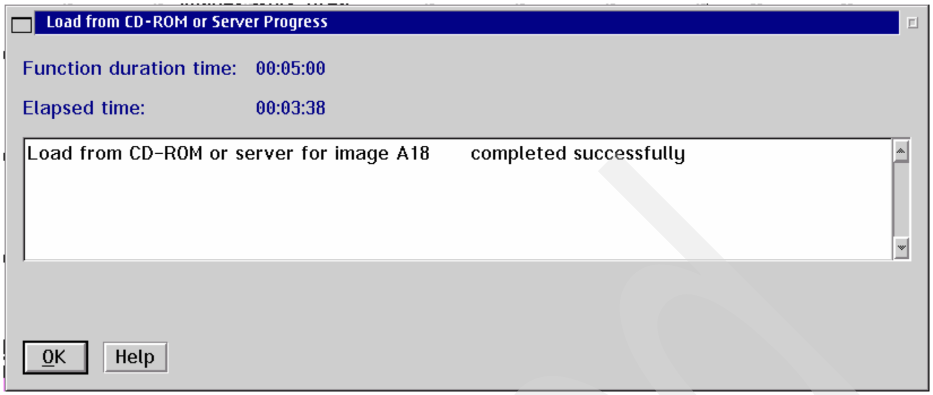
Function duration time: 00:05:00

Elapsed time: 00:00:21

Retrieving code from source.

Figure 1-7 Installation messages

When the image installation is complete, the panel shown in Figure 1-8 on page 9 is displayed.


Load from CD-ROM or Server Progress

Function duration time: 00:05:00

Elapsed time: 00:03:38

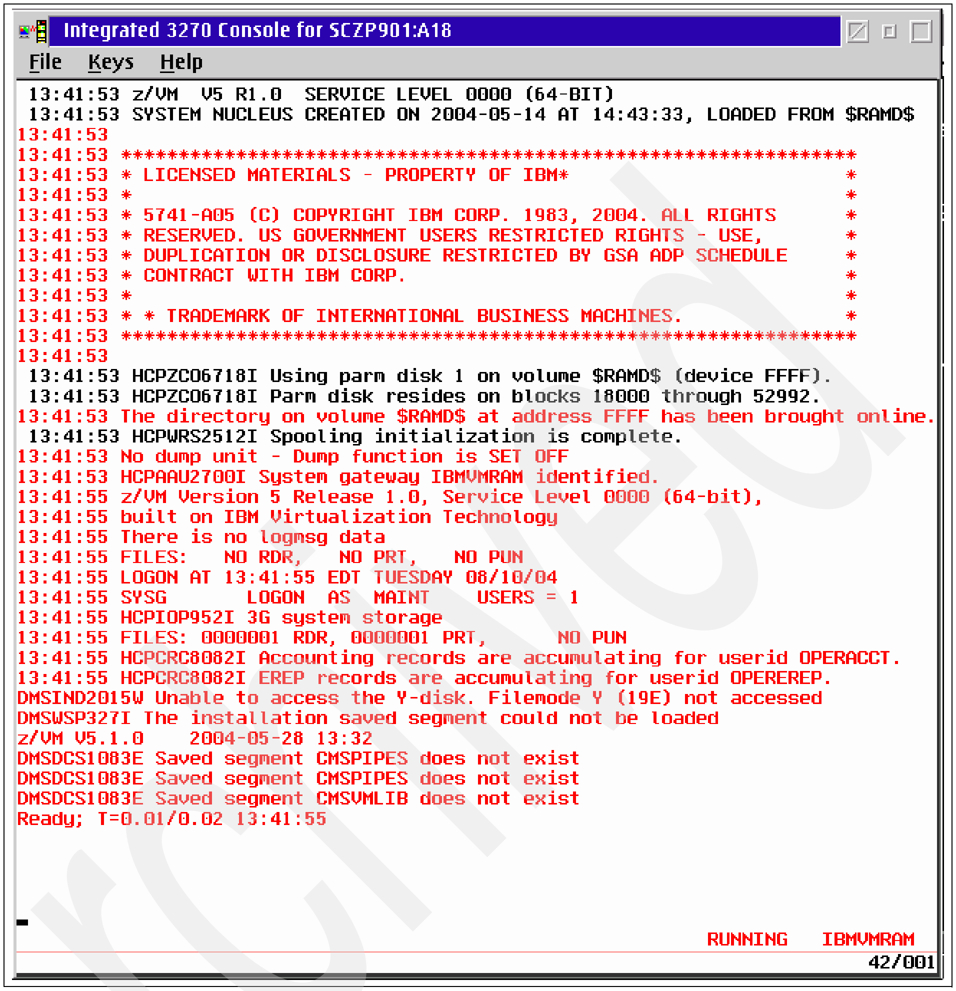
Load from CD-ROM or server for image A18 completed successfully

Figure 1-8 DVD load complete

Click **OK** to close and IPL from the RAMDISK.

1.2.4 IPL the RAMDISK

The system comes up with MAINT logged on. Messages are displayed on the Integrated 3270 Console as shown in Figure 1-9 on page 10.



```
 13:41:53 z/VM  V5 R1.0  SERVICE LEVEL 0000 (64-BIT)
 13:41:53 SYSTEM NUCLEUS CREATED ON 2004-05-14 AT 14:43:33, LOADED FROM $RAMD$
13:41:53
13:41:53 ****************************************************************
13:41:53 * LICENSED MATERIALS - PROPERTY OF IBM*                        *
13:41:53 *                                                              *
13:41:53 * 5741-A05 (C) COPYRIGHT IBM CORP. 1983, 2004. ALL RIGHTS      *
13:41:53 * RESERVED. US GOVERNMENT USERS RESTRICTED RIGHTS - USE,       *
13:41:53 * DUPLICATION OR DISCLOSURE RESTRICTED BY GSA ADP SCHEDULE     *
13:41:53 * CONTRACT WITH IBM CORP.                                      *
13:41:53 *                                                              *
13:41:53 * * TRADEMARK OF INTERNATIONAL BUSINESS MACHINES.              *
13:41:53 ****************************************************************
13:41:53
 13:41:53 HCPZCO6718I Using parm disk 1 on volume $RAMD$ (device FFFF).
 13:41:53 HCPZCO6718I Parm disk resides on blocks 18000 through 52992.
13:41:53 The directory on volume $RAMD$ at address FFFF has been brought online.
 13:41:53 HCPWRS2512I Spooling initialization is complete.
13:41:53 No dump unit - Dump function is SET OFF
13:41:53 HCPAAU2700I System gateway IBMVMRAM identified.
13:41:55 z/VM Version 5 Release 1.0, Service Level 0000 (64-bit),
13:41:55 built on IBM Virtualization Technology
13:41:55 There is no logmsg data
13:41:55 FILES:   NO RDR,    NO PRT,    NO PUN
13:41:55 LOGON AT 13:41:55 EDT TUESDAY 08/10/04
13:41:55 SYSG       LOGON  AS  MAINT    USERS = 1
13:41:55 HCPIOP952I 3G system storage
13:41:55 FILES: 0000001 RDR, 0000001 PRT,      NO PUN
13:41:55 HCPCRC8082I Accounting records are accumulating for userid OPERACCT.
13:41:55 HCPCRC8082I EREP records are accumulating for userid OPEREREP.
DMSIND2015W Unable to access the Y-disk. Filemode Y (19E) not accessed
DMSWSP327I The installation saved segment could not be loaded
z/VM V5.1.0    2004-05-28 13:32
DMSDCS1083E Saved segment CMSPIPES does not exist
DMSDCS1083E Saved segment CMSPIPES does not exist
DMSDCS1083E Saved segment CMSVMLIB does not exist
Ready; T=0.01/0.02 13:41:55
```

Figure 1-9 IPL from RAMDISK

At this point, log off the Primary Service Element (click the upper left corner of the window and select **Logoff)**. To install on 3390 DASD, issue the **INSTPLAN 3390** command from the Integrated 3270 Console.

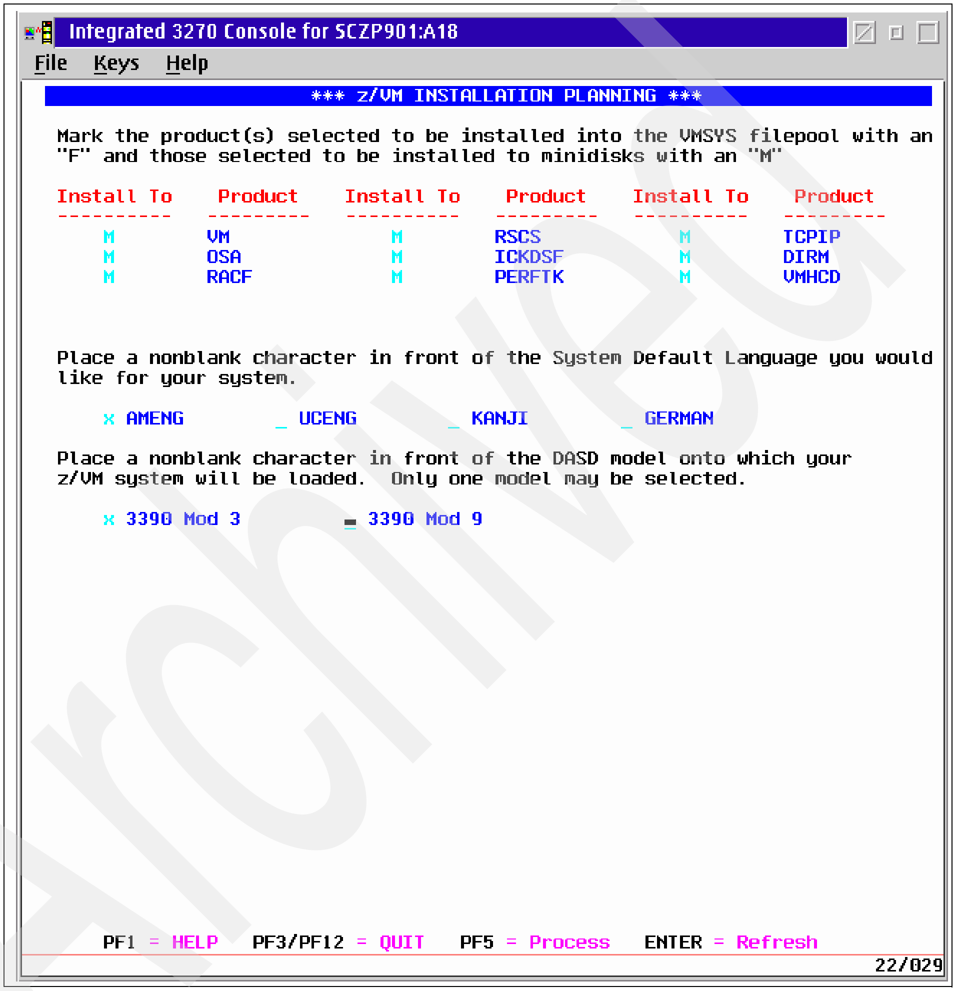
Note: For installation on FCP-attached SCSI, use the INSTPLAN FBA command. We discuss SCSI installation in 1.4, “Installation to FCP-attached SCSI disk” on page 20.

In Figure 1-10 on page 11, choose where the products are to be installed:

- **M** for VM minidisk

- **F** for SFS VMSYS file pool

Place an non-blank character next to the System Default Language for your system, and a non-blank character in front of the DASD model that matches your system installation.



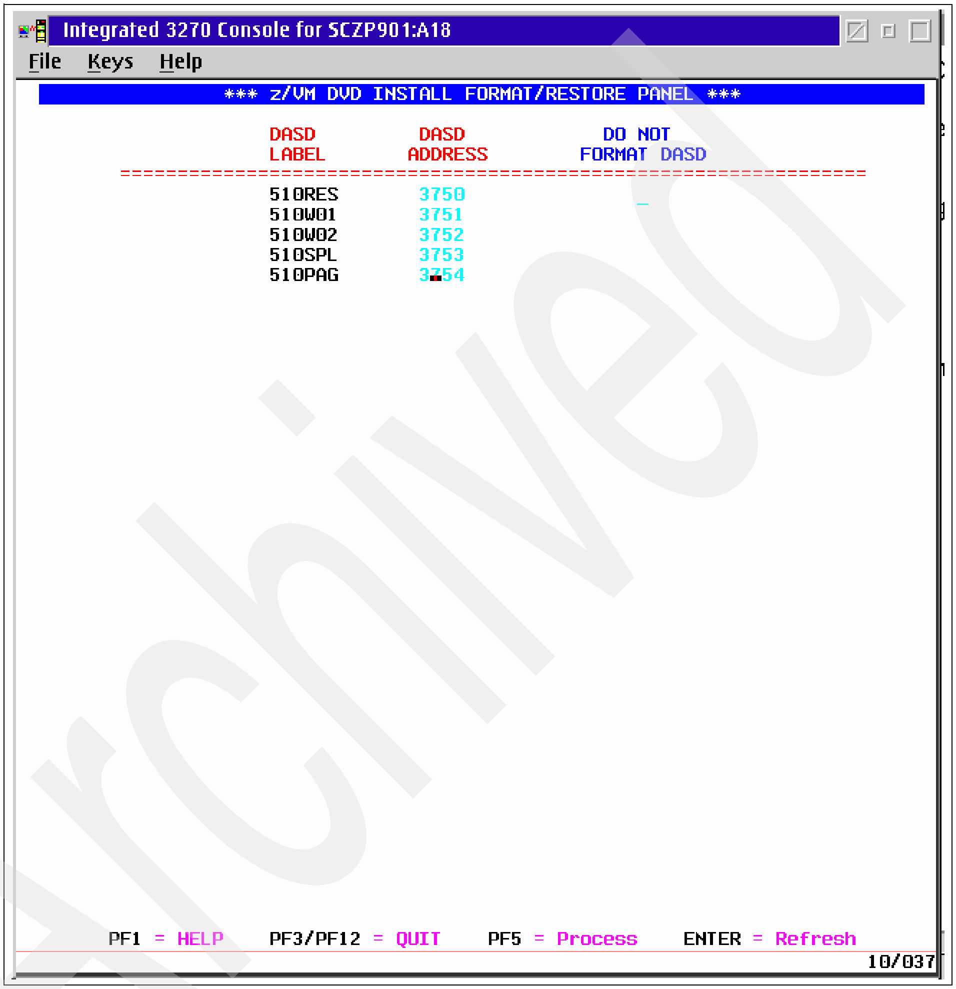
```
                       *** z/VM INSTALLATION PLANNING ***

 Mark the product(s) selected to be installed into the VMSYS filepool with an
 "F" and those selected to be installed to minidisks with an "M"

 Install To    Product    Install To    Product    Install To    Product
 ----------   ---------   ----------   ---------   ----------   ---------
     M          VM            M          RSCS          M          TCPIP
     M          OSA           M          ICKDSF        M          DIRM
     M          RACF          M          PERFTK        M          VMHCD


 Place a nonblank character in front of the System Default Language you would
 like for your system.

     x AMENG          _ UCENG          _ KANJI          _ GERMAN

 Place a nonblank character in front of the DASD model onto which your
 z/VM system will be loaded.  Only one model may be selected.

     x 3390 Mod 3           _ 3390 Mod 9


     PF1  = HELP    PF3/PF12 = QUIT    PF5 = Process    ENTER = Refresh
                                                                       22/029
```

Figure 1-10 The INSTPLAN dialog

Press **PF5** to complete the planing step.

1.2.5 Load the system image

To load the system image from DVD, execute the INSTDVD command. The panel shown in Figure 1-11 is displayed.



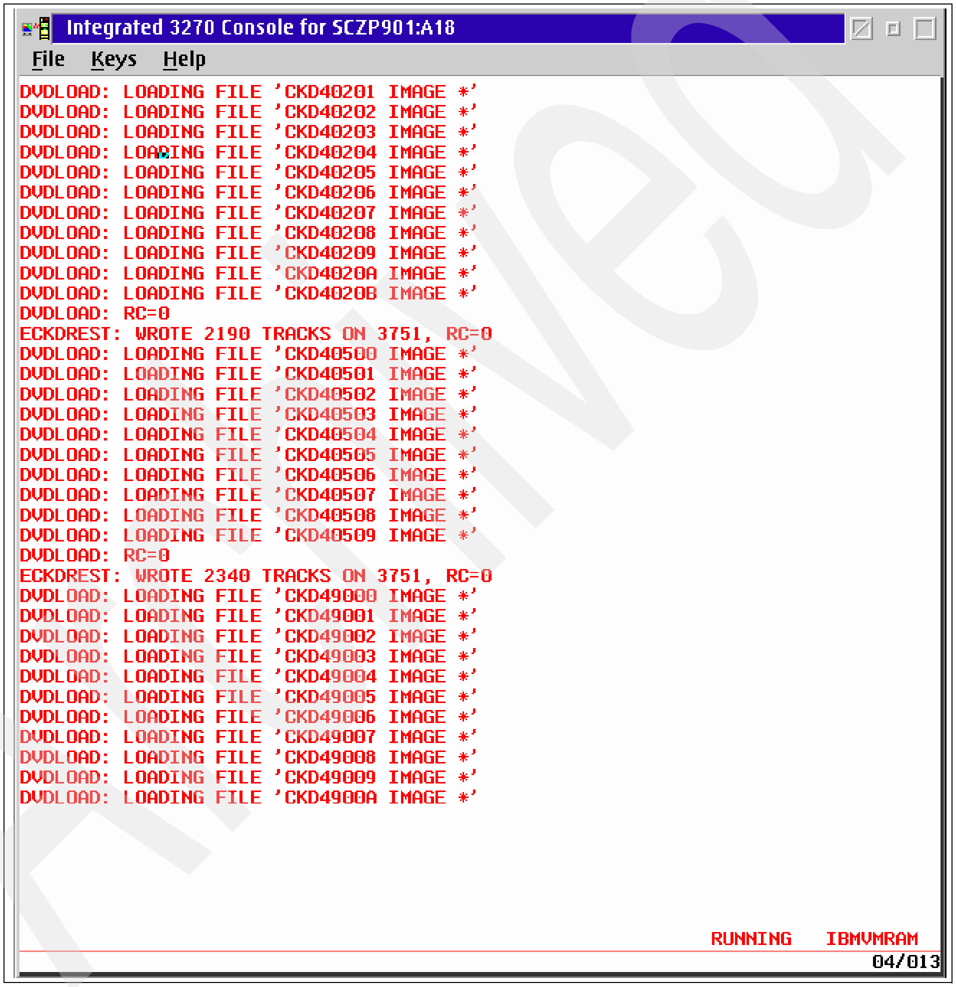
```
                    *** z/VM DVD INSTALL FORMAT/RESTORE PANEL ***

                    DASD          DASD              DO  NOT
                    LABEL        ADDRESS          FORMAT  DASD
          =================================================================
                    510RES        3750                 _
                    510W01        3751
                    510W02        3752
                    510SPL        3753
                    510PAG        3754




        PF1  = HELP      PF3/PF12  = QUIT     PF5  = Process     ENTER  = Refresh
                                                                          10/037
```

Figure 1-11 The INSTDVD dialog

Provide the DASD addresses specific to your installation. Place a non-blank character in the **DO NOT FORMAT DASD** column if the DASD is already formatted. Press **PF5** to process.

Tip: DASD formatting is performed sequentially. Depending on your devices, this can require some time (the five RVA disks used in this redbook required about 180 Minutes to format). To shorten the elapsed time, format the disks before installation using up to five userids operating in parallel (one for each DASD). In our case, this reduced elapsed time to about 25-30 Minutes.

Figure 1-12 shows the console messages that appear as installation proceeds.



```
DVDLOAD: LOADING FILE 'CKD40201 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40202 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40203 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40204 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40205 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40206 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40207 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40208 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40209 IMAGE *'
DVDLOAD: LOADING FILE 'CKD4020A IMAGE *'
DVDLOAD: LOADING FILE 'CKD4020B IMAGE *'
DVDLOAD: RC=0
ECKDREST: WROTE 2190 TRACKS ON 3751, RC=0
DVDLOAD: LOADING FILE 'CKD40500 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40501 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40502 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40503 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40504 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40505 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40506 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40507 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40508 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40509 IMAGE *'
DVDLOAD: RC=0
ECKDREST: WROTE 2340 TRACKS ON 3751, RC=0
DVDLOAD: LOADING FILE 'CKD49000 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49001 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49002 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49003 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49004 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49005 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49006 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49007 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49008 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49009 IMAGE *'
DVDLOAD: LOADING FILE 'CKD4900A IMAGE *'



                                                      RUNNING    IBMVMRAM
                                                                   04/013
```

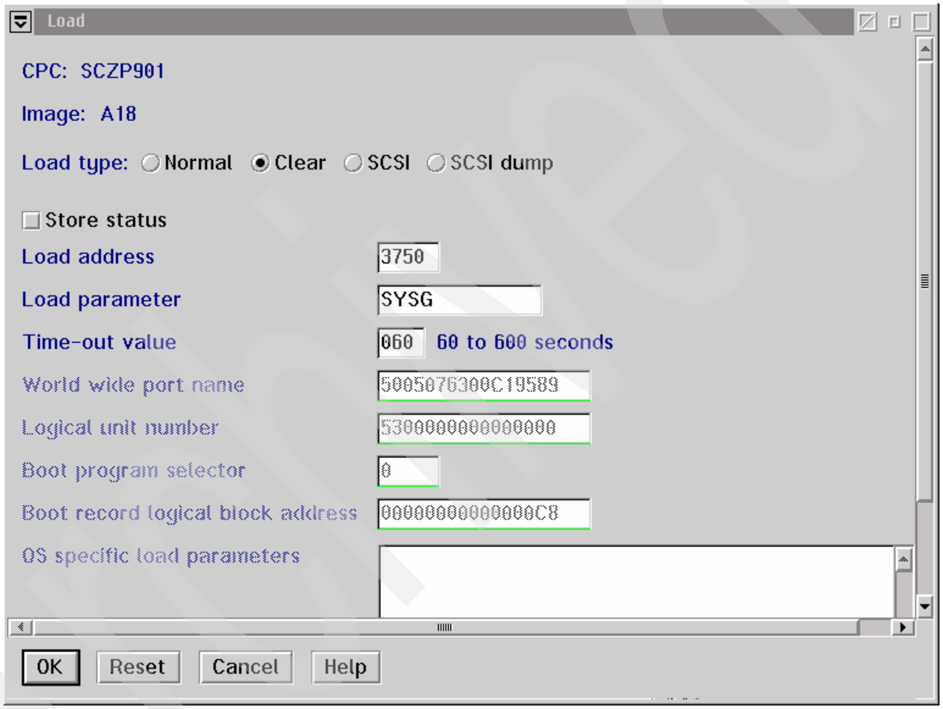
Figure 1-12   Installation messages

During system load, you are prompted to place the System RSU DVD in the DVD drive. Once the RSU DVD is loaded, press Enter. When the installation is complete, the following message is displayed:

```
HCPIDV8392I INSTDVD EXEC ENDED SUCCESSFULLY
```

1.2.6 IPL the installed z/VM 5.1 system

To IPL the installed z/VM system, navigate to the LOAD panel in the HMC shown in Figure 1-13.



CPC: SCZP901

Image: A18

Load type: ○ Normal ● Clear ○ SCSI ○ SCSI dump

□ Store status

Load address 3750

Load parameter SYSG

Time-out value 060 60 to 600 seconds

World wide port name 5005076300C19589

Logical unit number 5300000000000000

Boot program selector 0

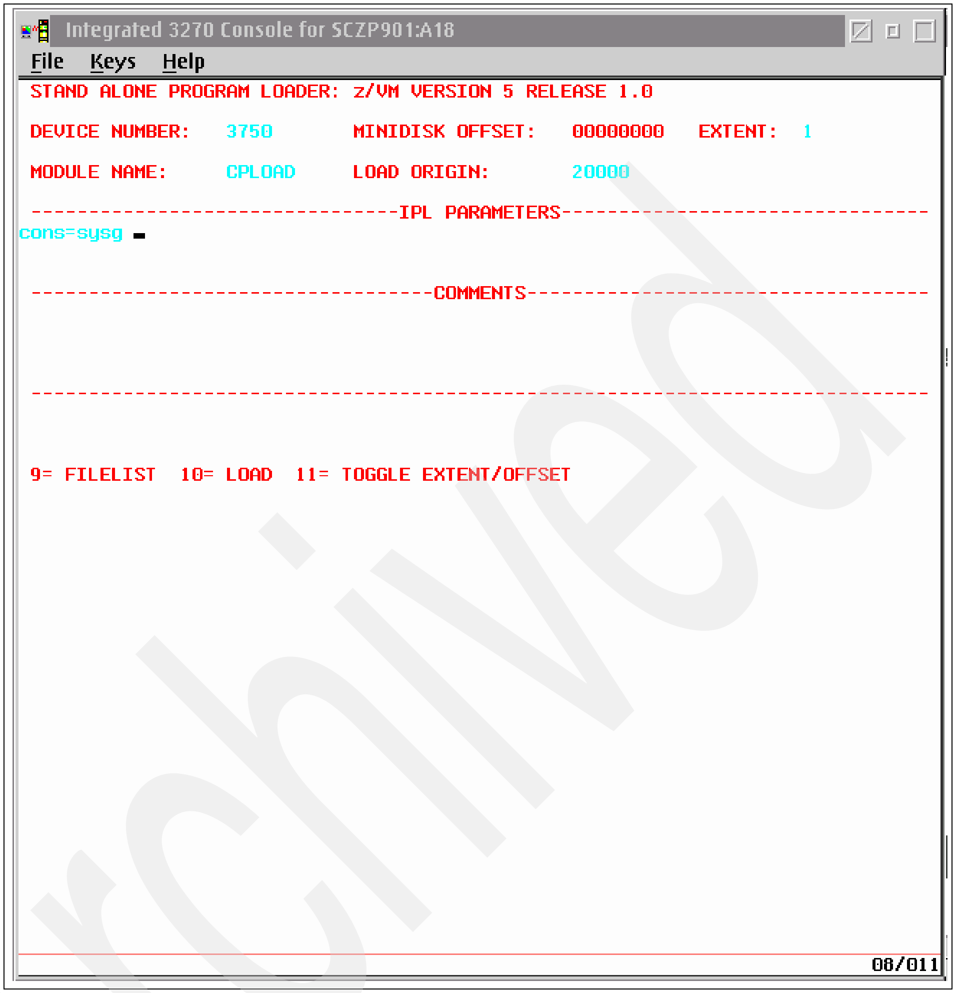
Boot record logical block address 00000000000000C8

OS specific load parameters

Figure 1-13 Load panel on the HMC

Choose **Load type** as **clear**, provide the address of the 510RES volume as the **Load address** (3750 in this example), and specify **SYSG** as **Load parameter**.

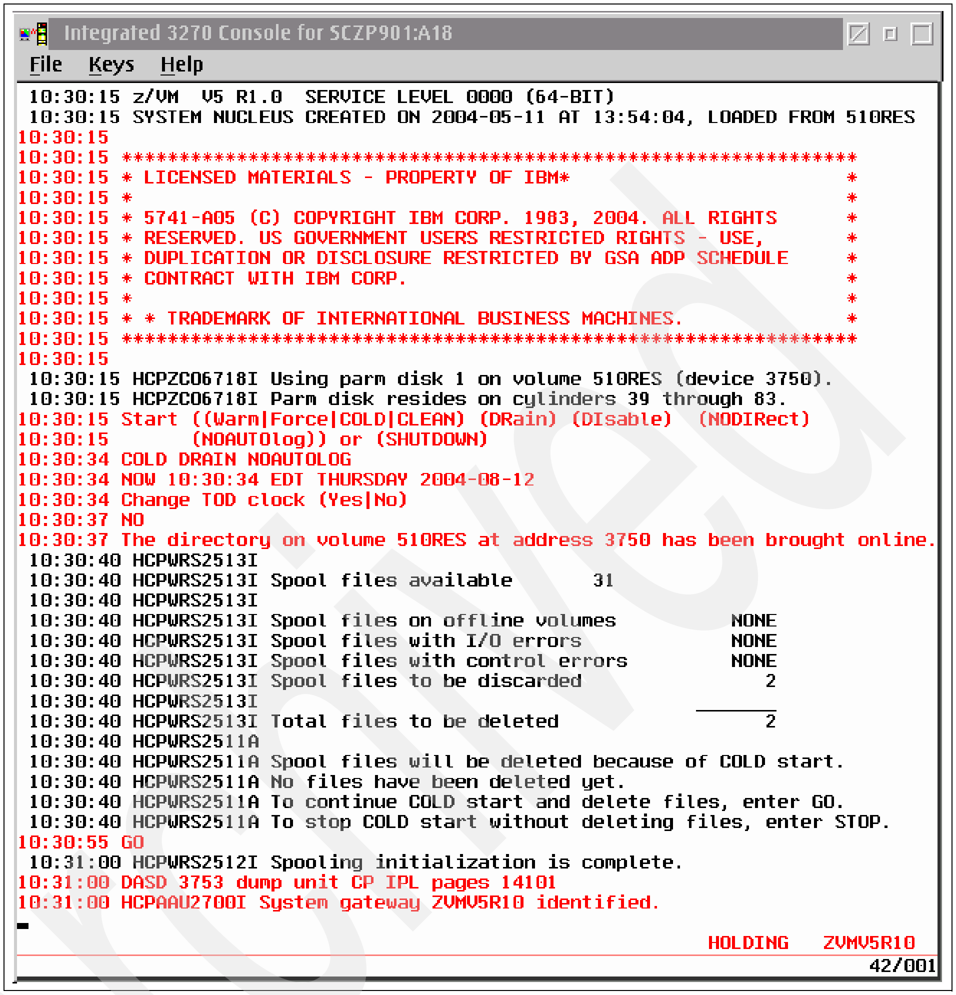
Click **OK** to proceed to the Stand Alone Program Loader panel in Figure 1-14 on page 15.



```
STAND ALONE PROGRAM LOADER: z/VM VERSION 5 RELEASE 1.0

DEVICE NUMBER:    3750       MINIDISK OFFSET:    00000000    EXTENT:  1

MODULE NAME:      CPLOAD     LOAD ORIGIN:        20000

--------------------------------IPL PARAMETERS--------------------------------
cons=sysg

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



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


9= FILELIST   10= LOAD   11= TOGGLE EXTENT/OFFSET
















                                                                        08/011
```

Figure 1-14 The standalone program loader

Provide `CONS=SYSG` as the **IPL Parameter** and press **PF10** to load z/VM. When prompted in Figure 1-15 on page 16, specify COLD DRAIN NOAUTOLOG.



```
 10:30:15 z/VM  V5 R1.0  SERVICE LEVEL 0000 (64-BIT)
 10:30:15 SYSTEM NUCLEUS CREATED ON 2004-05-11 AT 13:54:04, LOADED FROM 510RES
10:30:15
10:30:15 ****************************************************************
10:30:15 * LICENSED MATERIALS - PROPERTY OF IBM*                        *
10:30:15 *                                                              *
10:30:15 * 5741-A05 (C) COPYRIGHT IBM CORP. 1983, 2004. ALL RIGHTS      *
10:30:15 * RESERVED. US GOVERNMENT USERS RESTRICTED RIGHTS - USE,       *
10:30:15 * DUPLICATION OR DISCLOSURE RESTRICTED BY GSA ADP SCHEDULE     *
10:30:15 * CONTRACT WITH IBM CORP.                                      *
10:30:15 *                                                              *
10:30:15 * * TRADEMARK OF INTERNATIONAL BUSINESS MACHINES.              *
10:30:15 ****************************************************************
10:30:15
 10:30:15 HCPZCO6718I Using parm disk 1 on volume 510RES (device 3750).
 10:30:15 HCPZCO6718I Parm disk resides on cylinders 39 through 83.
10:30:15 Start ((Warm|Force|COLD|CLEAN) (DRain) (DIsable)  (NODIRect)
10:30:15       (NOAUTOlog)) or (SHUTDOWN)
10:30:34 COLD DRAIN NOAUTOLOG
10:30:34 NOW 10:30:34 EDT THURSDAY 2004-08-12
10:30:34 Change TOD clock (Yes|No)
10:30:37 NO
10:30:37 The directory on volume 510RES at address 3750 has been brought online.
 10:30:40 HCPWRS2513I
 10:30:40 HCPWRS2513I Spool files available       31
 10:30:40 HCPWRS2513I
 10:30:40 HCPWRS2513I Spool files on offline volumes           NONE
 10:30:40 HCPWRS2513I Spool files with I/O errors              NONE
 10:30:40 HCPWRS2513I Spool files with control errors          NONE
 10:30:40 HCPWRS2513I Spool files to be discarded                 2
 10:30:40 HCPWRS2513I                                      -------
 10:30:40 HCPWRS2513I Total files to be deleted                   2
 10:30:40 HCPWRS2511A
 10:30:40 HCPWRS2511A Spool files will be deleted because of COLD start.
 10:30:40 HCPWRS2511A No files have been deleted yet.
 10:30:40 HCPWRS2511A To continue COLD start and delete files, enter GO.
 10:30:40 HCPWRS2511A To stop COLD start without deleting files, enter STOP.
10:30:55 GO
 10:31:00 HCPWRS2512I Spooling initialization is complete.
10:31:00 DASD 3753 dump unit CP IPL pages 14101
10:31:00 HCPAAU2700I System gateway ZVMV5R10 identified.
_
                                                      HOLDING   ZVMV5R10
                                                                 42/001
```

Figure 1-15   IPL console messages

Disconnect from the operator userid and logon to MAINT. To complete the installation, issue the INSTVM DVD command. When complete, the following message is displayed:

```
HCPIVM8392I INSTVM EXEC ENDED SUCCESSFULLY.
```

1.2.7 Apply the Recommended Service Upgrade

Next, we load the service files from the Recommended Service Upgrade (RSU) servlink. Logon to MAINT and access the disk containing the RSU servlink as C:

```
ACCESS 500 c
```

Get the filenames of the RSU envelopes on the 500 disk:

```
LISTFILE * SERVLINK C
```

Run the SERVICE command:

```
SERVICE ALL fn1 fn2
```

Use the filenames of the envelope files (as reported from the previous LISTFILE command) for the ***fn1*** and ***fn2*** parameters. When complete, the following message is displayed:

```
VMFSRV2760I SERVICE processing completed successfully
```

To place the products into production, execute the PUT2PROD command as MAINT. When complete, the following message is displayed:

```
VMFP2P2760I PUT2PROD processing completed successfully
```

Shutdown and re-IPL z/VM. When the system returns, installation is complete. Now configure TCP/IP for an Initial Network Connection as described in 1.5, “TCP/IP configuration” on page 28.

1.3 Second level installation from DVD

Second level z/VM 5.1 installation from DVD is possible only from a first level z/VM 5.1 system. To install z/VM 5.1 second level from an older z/VM version, you must use the DDR installation method.

Note: Second level DVD installation requires the DVDPRIME EXEC (a new command for z/VM 5.1) to define the FTP connection to the DVD. This command is not available on older z/VM versions.

Second level installation from DVD differs only slightly different from first level installation. In this section, we point out the differences.

1.3.1 Set up the userid for installation

In the first level system, create a new userid on which to install z/VM 5.1. The userid should have privilege classes B and G, and be defined with a minimum of 64 MB virtual storage.

Attention: You should not grant privilege class A authority to the user in order to prevent the second level user from accidentally shutting down the first level system.

Verify the userid has the following resources defined to it:

- 191 A-disk accessed in read/write mode.
- Read access to the MAINT 2CC minidisk.
- 22CC minidisk with the size of 5 cylinders when using 3390 (7200 blocks when using FBA).
- 2CF1 minidisk with the size of 45 cylinders when using 3390 (68400 blocks when using FBA).

Access the MAINT 2CC minidisk as C:

```
ACCESS 2CC C
```

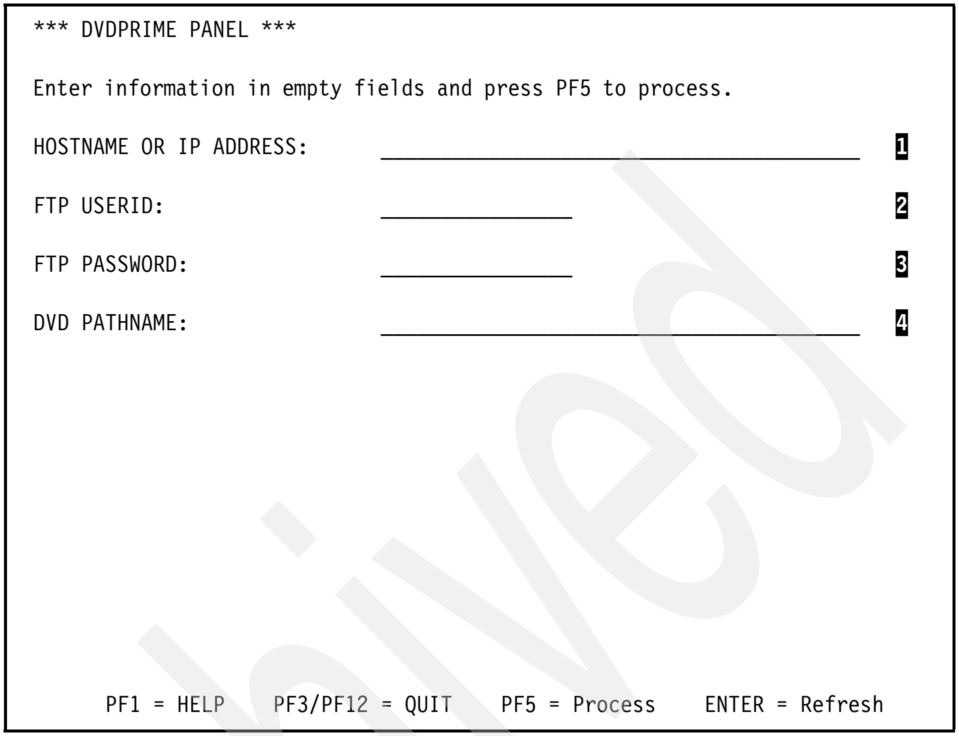
1.3.2 Run the DVDPRIME EXEC

To begin installation, load the z/VM 5.1 DVD in the DVD drive. Execute the DVDPRIME EXEC to define an FTP connection to the DVD drive:

```
DVDPRIME 3390
```

Note: For installation on FCP-attached SCSI, use the DVDPRIME FBA command. We discuss SCSI installation in 1.4, “Installation to FCP-attached SCSI disk” on page 20.

This invokes the DVDPRIME panel shown in Figure 1-16 on page 19.



```
*** DVDPRIME PANEL ***

Enter information in empty fields and press PF5 to process.

HOSTNAME OR IP ADDRESS:       ________________________________________   1

FTP USERID:                   ________________                           2

FTP PASSWORD:                 ________________                           3

DVD PATHNAME:                 ________________________________________   4










      PF1 = HELP    PF3/PF12 = QUIT    PF5 = Process    ENTER = Refresh
```

Figure 1-16 The DVDPRIME panel

In the panel, provide information specific to your environment:

1. The hostname or IP address of the server where the DVD is mounted.
2. The userid used to access the FTP server.
3. The password to logon to the FTP server.
4. The pathname of the DVD drive with /CPDVD directory appended to the end.

Press **PF5** to process. When execution is complete, following message is displayed:

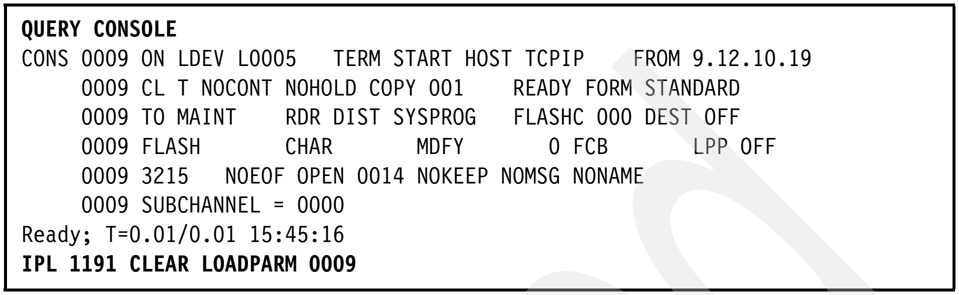
```
HCPDVP8392I DVDPRIME EXEC ENDED SUCCESSFULLY
```

At this point, follow the steps outlined in 1.2.4, “IPL the RAMDISK” on page 9 through 1.2.7, “Apply the Recommended Service Upgrade” on page 16.

From the first level user, issue the commands:

```
SYSTEM CLEAR
TERMINAL CONMODE 3270
```

To IPL the second level system, query the console to determine the virtual console address. IPL the system using that console address as shown in Figure 1-17.



```
QUERY CONSOLE
CONS 0009 ON LDEV L0005   TERM START HOST TCPIP    FROM 9.12.10.19
     0009 CL T NOCONT NOHOLD COPY 001    READY FORM STANDARD
     0009 TO MAINT    RDR DIST SYSPROG   FLASHC 000 DEST OFF
     0009 FLASH       CHAR       MDFY       0 FCB       LPP OFF
     0009 3215   NOEOF OPEN 0014 NOKEEP NOMSG NONAME
     0009 SUBCHANNEL = 0000
Ready; T=0.01/0.01 15:45:16
IPL 1191 CLEAR LOADPARM 0009
```

Figure 1-17 Second level z/VM 5.1 IPL

In this example, the virtual console is assigned to device 009.

1.4 Installation to FCP-attached SCSI disk

z/VM 5.1 supports Fibre Channel Protocol (FCP) attached SCSI disks for use as both system and guest storage devices. This enables z/VM to be installed on and operate from either SCSI or traditional extended count key data (ECKD™) disks.

Prior to z/VM 5.1, SCSI disk support was limited to Linux guests configured for FCP SCSI devices. Native SCSI disk support z/VM 5.1 is provided for logical units defined in an IBM TotalStorage® Enterprise Storage Server® (ESS) connected to a Fibre Channel fabric.

> **Note:** For details on how to configure and use FCP SCSI devices in a z/VM Linux guest, see *Linux for zSeries: Fibre Channel Protocol Implementation Guide*, SG24-6344.

When used by CMS and CP, SCSI disks are emulated as 9336 model 20 fixed-block-architecture (FBA) disks. With z/VM 5.1, SCSI disks can be used for system paging, spooling, directory services, and minidisks. z/VM guests that support FBA disks (such as CMS and Linux) can use SCSI disks without requiring specific SCSI support (using FBA emulation). Figure 1-18 on page 21 illustrates SCSI support in z/VM 5.1.

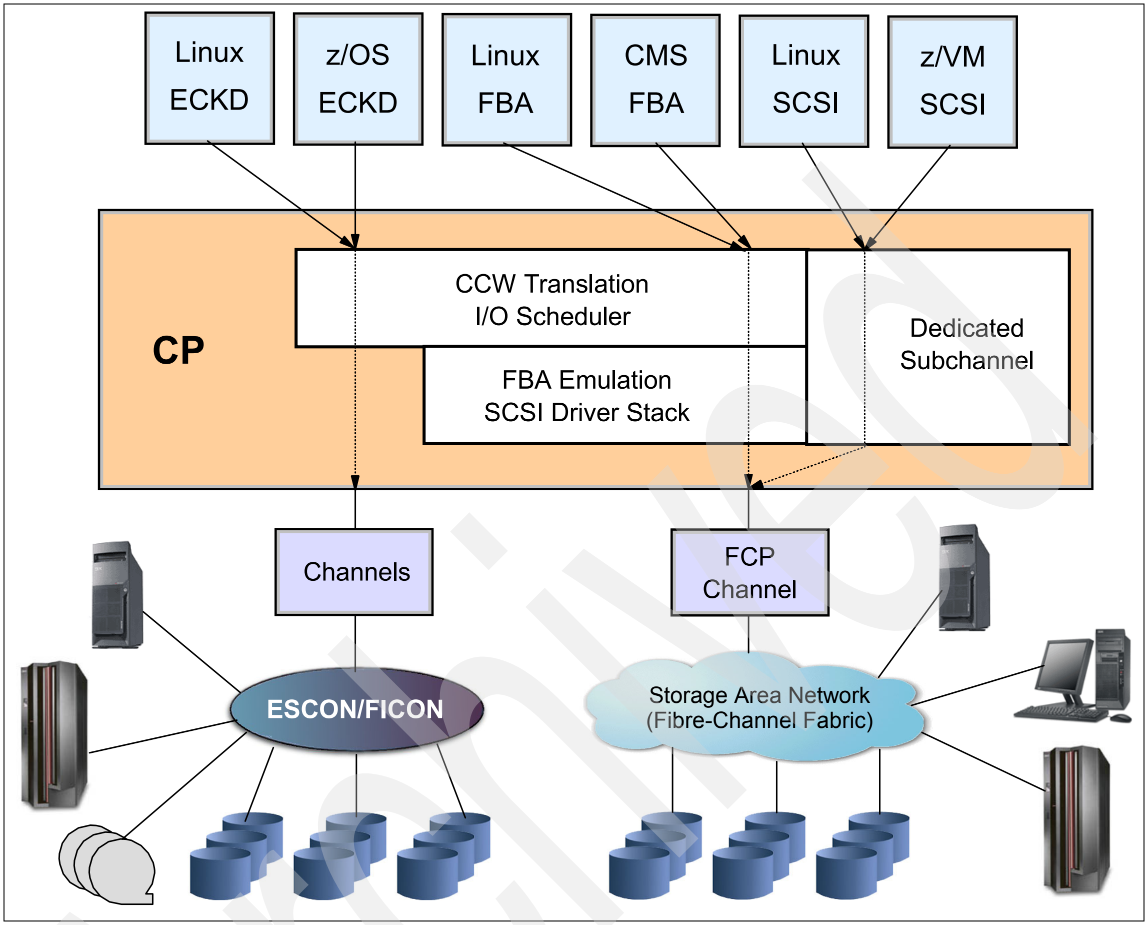


Figure 1-18 SCSI support in z/VM 5.1

z/VM supports emulated FBA disks up to 1 terabyte minus 1 page in size. However, directory, paging, and spool storage must be allocated within the first 64GB of a CP-formatted volume. Other CP allocations (such as TDSK, PERM, and PARM) may be allocated past the first 64GB.

Note: The maximum size of FBA SCSI disks allocated for use by CMS or GCS guests is 381GB. However, FBA SCSI disks used by CMS should not be larger than 22GB in size. CMS file system control and status data structures must reside below 16MB in virtual storage. With larger minidisks, the system may not be able to obtain sufficient virtual storage below 16MB to access the disks. For more details, consult *z/VM: CP Planning and Administration*, SC24-6043.

1.4.1 Planning for installation

Figure 1-19 on page 23 illustrates our system configuration for installing z/VM 5.1 to SCSI disk. In the ESS, three Logical Unit Numbers (LUNs) are defined for the z/VM system volumes:

- LUN 5300 is allocated for the 510RES volume.
- LUN 5301 is allocated for the 510W01 volume.
- LUN 5302 is allocated for the 510SPL volume.
- LUN 5303 is allocated for the 510PAG volume.

The LUNs are connected to the Fibre Channel fabric using Worldwide Port Name 5005076300C19589. The z/VM 5.1 system defines FCP devices B000-B003 connected to FCP CHPID 000B. The CHPID connects to the fabric through a FICON® or FICON Express card.

Note: For details on FCP concepts and defining FCP-attached SCSI devices, consult *Linux for zSeries: Fibre Channel Protocol Implementation Guide*, SG24-6344.

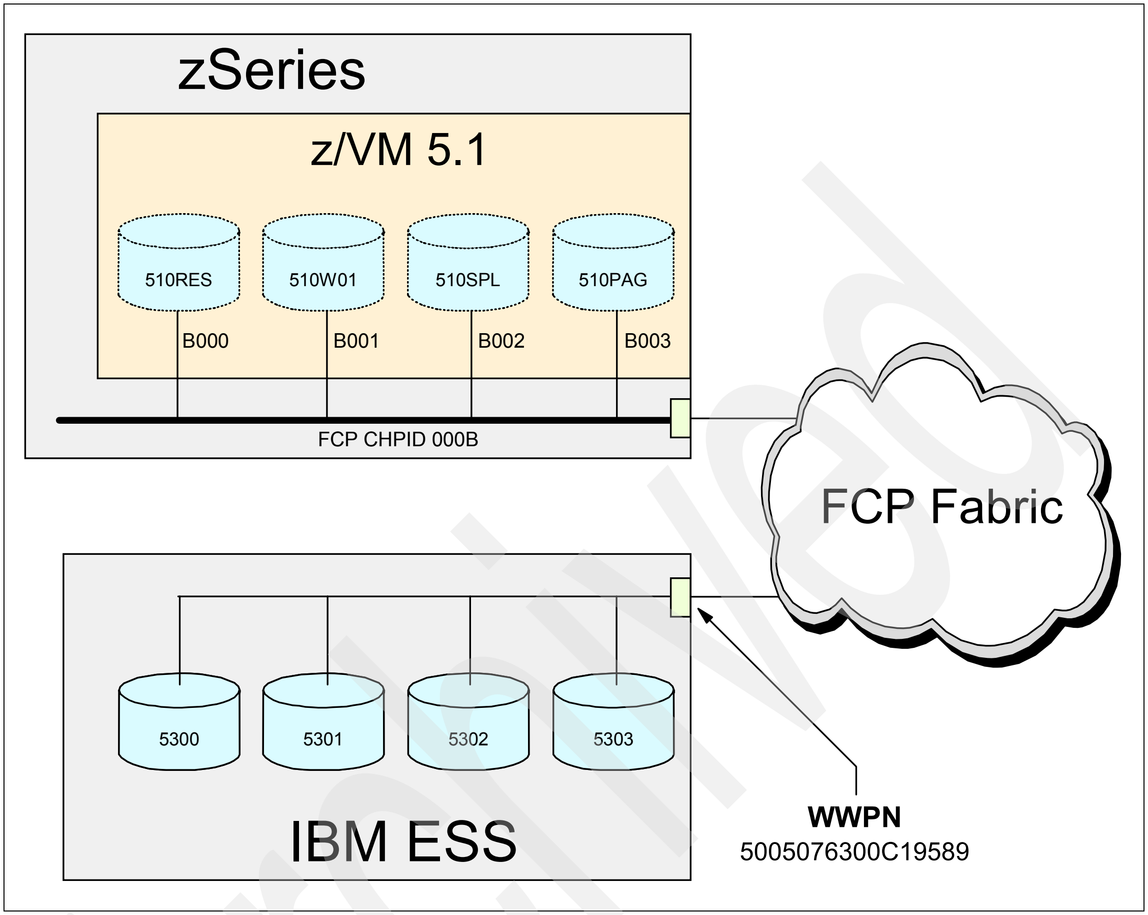


Figure 1-19 System configuration for installation to SCSI disk

1.4.2 Emulated FBA disks

In order for z/VM 5.1 to use a SCSI disk, an emulated FBA device must be created for the disk. z/VM 5.1 provides two methods to define the emulated device:

- The EDEVICE statement in the SYSTEM CONFIG file (described in *z/VM: CP Planning and Administration*, SC24-6043)
- The CP SET EDEVICE command (described in *z/VM: CP Command and Utility Reference*, SC24-6008)

> **Note:** Emulated FBA SCSI disks can also be defined using z/VM's Hardware Configuration Manager (HCM) and Hardware Configuration Definition (HCD) support. For details, see *z/OS and z/VM: Hardware Configuration Manager User's Guide*, SC33-7989 and *z/VM: I/O Configuration*, SC24-6100.

To create an emulated FBA device for an FCP-attached SCSI disk, we need to supply:

- The Worldwide Port Name (WWPN) used to access the SCSI disk
- The Logical Unit Number (LUN) address of the SCSI disk
- The FCP device number used access the SCSI disk
- An emulated device number to access the emulated FBA disk

Table 1-1 contains the information need to define four emulated FBA devices for installing z/VM 5.1.

Table 1-1 SCSI disks used for z/VM 5.1 installation

Volume	WWPN	LUN	Device number	
			FCP	**Emulated**
510RES	5005076300C19589	5300000000000000	B000	5300
510W01	5005076300C19589	5301000000000000	B001	5301
510SPL	5005076300C19589	5302000000000000	B002	5302
510PAG	5005076300C19589	5303000000000000	B003	5303

The SET EDEVICE command syntax to define an emulated FBA device is:

```
SET EDEVICE rdev TYPE FBA ATTR 2105 fcpdev WWPN wwpn LUN lun
```

Parameters to the command include:

rdev The emulated device number to create

fcpdev The FCP device used to access the SCSI disk

wwpn The WWPN on the ESS used to access the SCSI disk

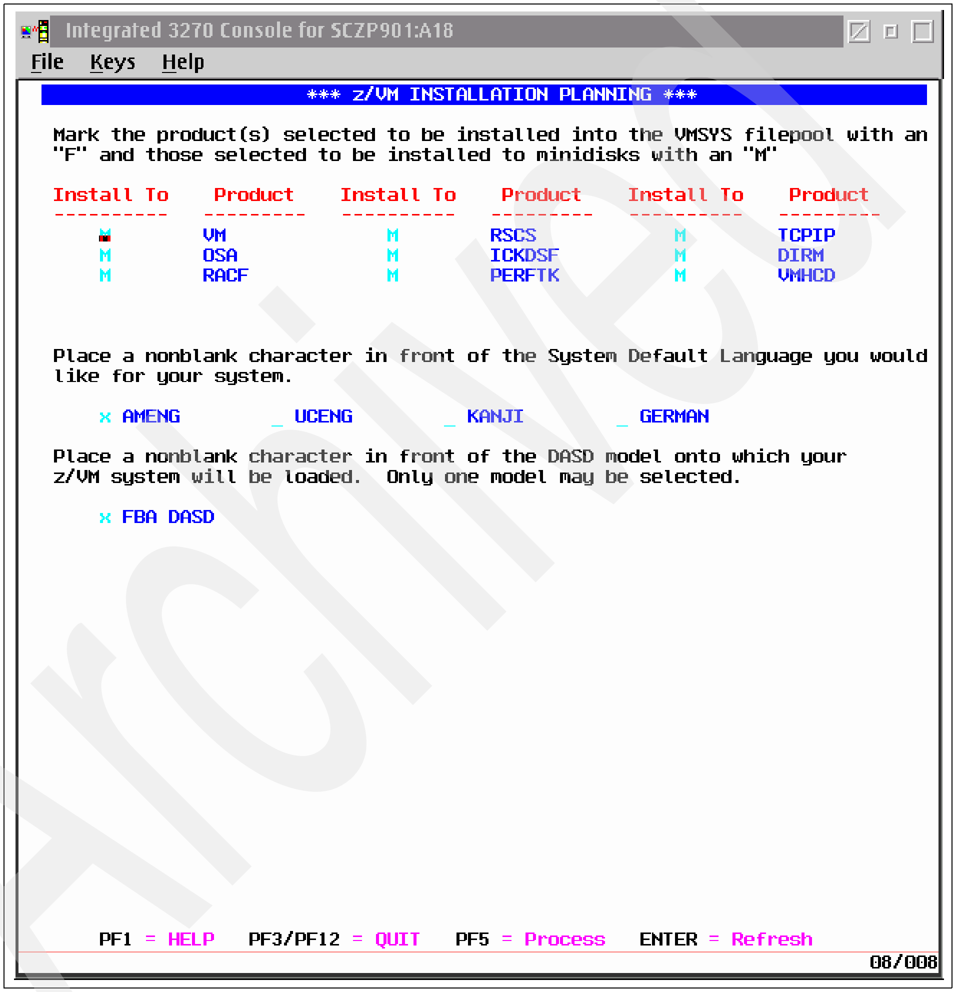
lun The LUN address of the SCSI disk

1.4.3 Installation steps

To install z/VM 5.1 on SCSI disk, use the SCSI System DVD installation set and follow the steps outlined in:

- “Establish an Integrated 3270 Console session” on page 2
- “Access the primary Support Element” on page 3
- “Load the z/VM 5.1 RAMDISK” on page 6

Once the z/VM 5.1 RAMDISK is IPLed, issue the **INSTPLAN FBA** command. The INSTPLAN panel shown in Figure 1-20 is displayed.



```
                        *** z/VM INSTALLATION PLANNING ***

 Mark the product(s) selected to be installed into the VMSYS filepool with an
 "F" and those selected to be installed to minidisks with an "M"

 Install To     Product      Install To     Product      Install To     Product
 ----------    ---------     ----------    ---------     ----------    ---------
     M         VM                M         RSCS              M         TCPIP
     M         OSA               M         ICKDSF            M         DIRM
     M         RACF              M         PERFTK            M         VMHCD


 Place a nonblank character in front of the System Default Language you would
 like for your system.

     x AMENG          _ UCENG          _ KANJI          _ GERMAN

 Place a nonblank character in front of the DASD model onto which your
 z/VM system will be loaded.  Only one model may be selected.

     x FBA DASD

       PF1 = HELP    PF3/PF12 = QUIT    PF5 = Process    ENTER = Refresh
                                                                        08/008
```

Figure 1-20 The INSTPLAN dialog for SCSI installation

Choose the components to install, the system default language, and select the **FBA DASD** option. Use **PF5** to process your selections.

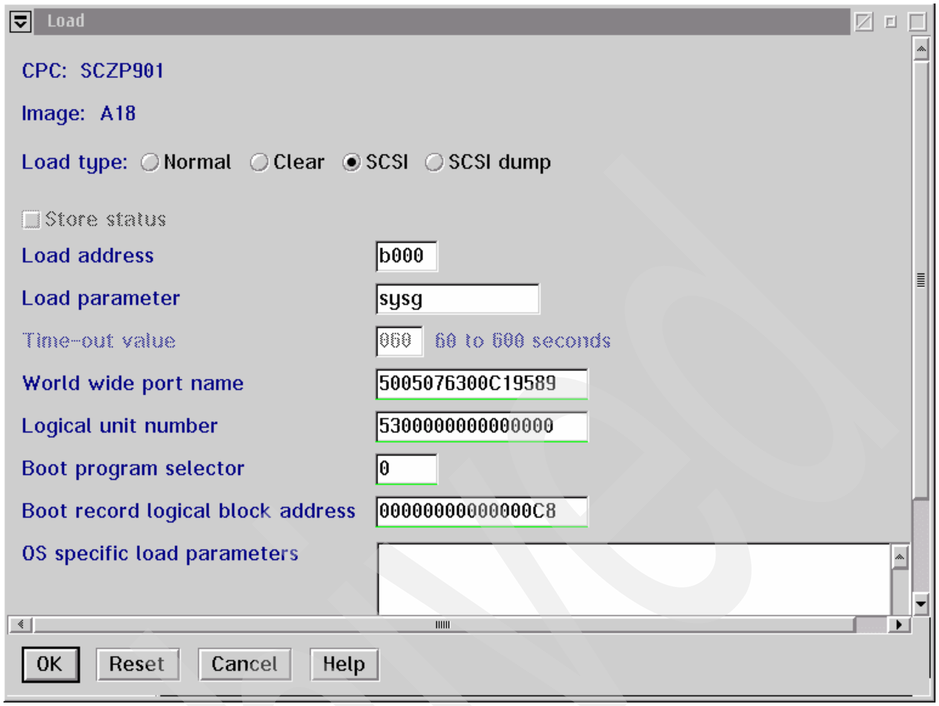
At this point, define the emulated FBA devices for the system volumes on SCSI disk. In Figure 1-21, we use the SET EDEVICE command to define emulated devices for our installation. Once the emulated devices are defined, VARY ONLINE the real FCP devices, and ATTACH the emulated devices to MAINT.

```
SET EDEVICE 5300 TYPE FBA ATTR 2105 FCP_DEV B000 WWPN 5005076300C19589 LUN 5300000000000000
11:47:00 EDEV 5300 was created.
Ready; T=0.01/0.03 11:47:00
SET EDEVICE 5301 TYPE FBA ATTR 2105 FCP_DEV B001 WWPN 5005076300C19589 LUN 5301000000000000
11:47:14 EDEV 5301 was created.
Ready; T=0.01/0.03 11:47:14
SET EDEVICE 5302 TYPE FBA ATTR 2105 FCP_DEV B002 WWPN 5005076300C19589 LUN 5302000000000000
11:47:32 EDEV 5302 was created.
Ready; T=0.01/0.03 11:47:32
SET EDEVICE 5303 TYPE FBA ATTR 2105 FCP_DEV B003 WWPN 5005076300C19589 LUN 5303000000000000
11:47:46 EDEV 5303 was created.
Ready; T=0.01/0.03 11:47:46
VARY ONLINE 5300-5303
11:50:12 5300 varied online
11:50:14 5301 varied online
11:50:16 5302 varied online
11:50:18 5303 varied online
11:50:28 4 device(s) specified; 4 device(s) successfully varied online
Ready; T=0.01/0.01 11:50:28
ATTACH 5300-5303 *
```

Figure 1-21 Define FBA emulated devices for the SCSI disks

At this point, the system image is ready to be installed. Execute the INSTDVD command (for details on the INSTDVD installation panel, see 1.2.5, “Load the system image” on page 12).

Once the system image loaded and the System RSU is applied, z/VM 5.1 is ready for IPL. Navigate to the Load panel on the HMC. The dialog is shown in Figure 1-22 on page 27.



Load

CPC: SCZP901

Image: A18

Load type: ○ Normal ○ Clear ● SCSI ○ SCSI dump

□ Store status

Load address b000

Load parameter sysg

Time-out value 060 60 to 600 seconds

World wide port name 5005076300C19589

Logical unit number 5300000000000000

Boot program selector 0

Boot record logical block address 00000000000000C8

OS specific load parameters

Figure 1-22 The SCSI IPL panel

In this panel:

- Choose the **SCSI** option for **Load type**.
- Provide the **WWPN** used to access the 510RES volume (5005076300C19589 in this example).
- Provide the **LUN** of the 510RES volume (5300000000000000 is this example).

Press **OK** to continue the Stand Alone Program Loader shown in Figure 1-14 on page 15.

> **Important:** When installing on FBA emulated disk, you must provide the **PDVOL=*edev*** IPL parameter to the Stand Alone Program Loader (in addition to the **CONS=SYSG** IPL parameter). The ***edev*** parameter specifies the emulated device address of the 510RES volume. In our example, the complete IPL parameter string to supply is **CONS=SYSG PDVOL=5300**.

1.5 TCP/IP configuration

Once the z/VM 5.1 system, you can configure TCP/IP. In this section, we look at two methods which can simplify TCP/IP configuration:

- The IPWIZARD command
- The IFCONFIG command

For a complete description on TCP/IP configuration, consult *z/VM: TCP/IP Planning and Customization*, SC24-6019.

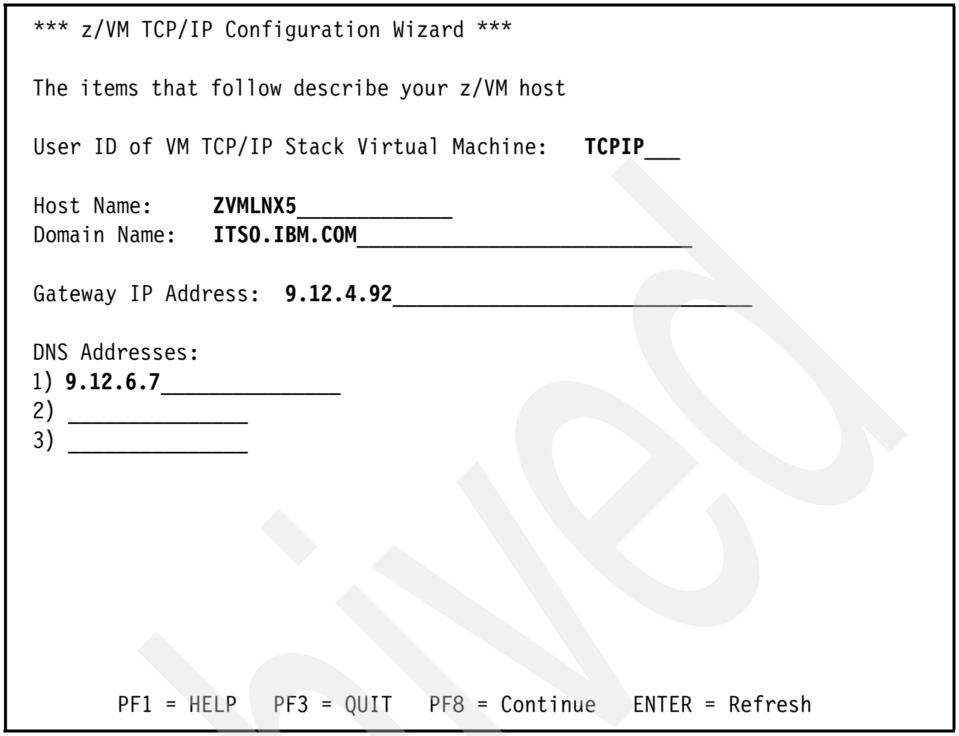
Tip: For a step-by-step description on how to configure a virtual switch (VSWITCH), see *z/VM: Getting Started with Linux on zSeries*, SC24-6096.

1.5.1 Configuring TCP/IP with the IPWIZARD command

The IPWIZARD command is a menu-driven tool to configure the PROFILE TCP/IP file. To access the command, logon as MAINT and access the 193 minidisk, and execute the IPWIZARD command:

```
ACCESS 193 E
IPWIZARD
```

The initial IPWIZARD menu is shown in Figure 1-23 on page 29.



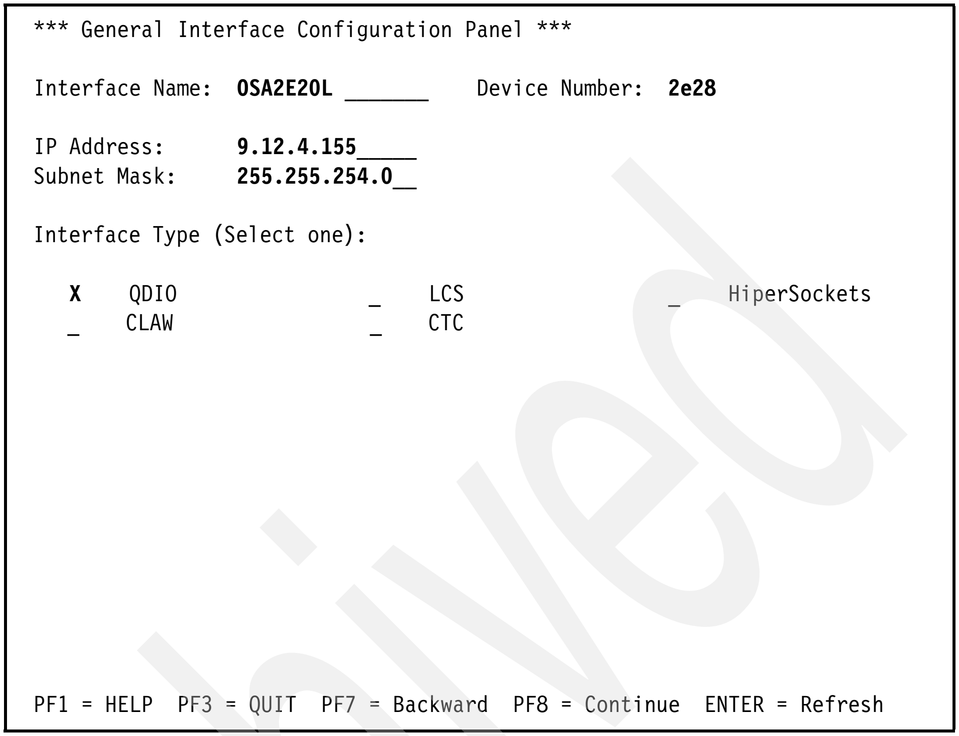
```
*** z/VM TCP/IP Configuration Wizard ***

The items that follow describe your z/VM host

User ID of VM TCP/IP Stack Virtual Machine:   TCPIP___

Host Name:     ZVMLNX5_____________
Domain Name:   ITSO.IBM.COM____________________________

Gateway IP Address:  9.12.4.92______________________________

DNS Addresses:
1) 9.12.6.7_______________
2) _______________
3) _______________



         PF1 = HELP   PF3 = QUIT   PF8 = Continue   ENTER = Refresh
```

Figure 1-23 The IPWIZARD initial menu

Provide the information that describes the z/VM host:

- User ID of VM TCP/IP Stack
  By default, the TCP/IP stack is managed by the TCP/IP user.
- Host Name
  The host name in this example is ZVMLNX5.
- Domain Name
  The host belongs to the ITSO.IBM.COM domain.
- Gateway IP Address
  This is the default route.
- DNS Addresses
  Up to three Domain Name Service (DNS) servers may be specified.

Press **PF8** to continue to the General Interface Configuration Panel shown in Figure 1-24 on page 30.



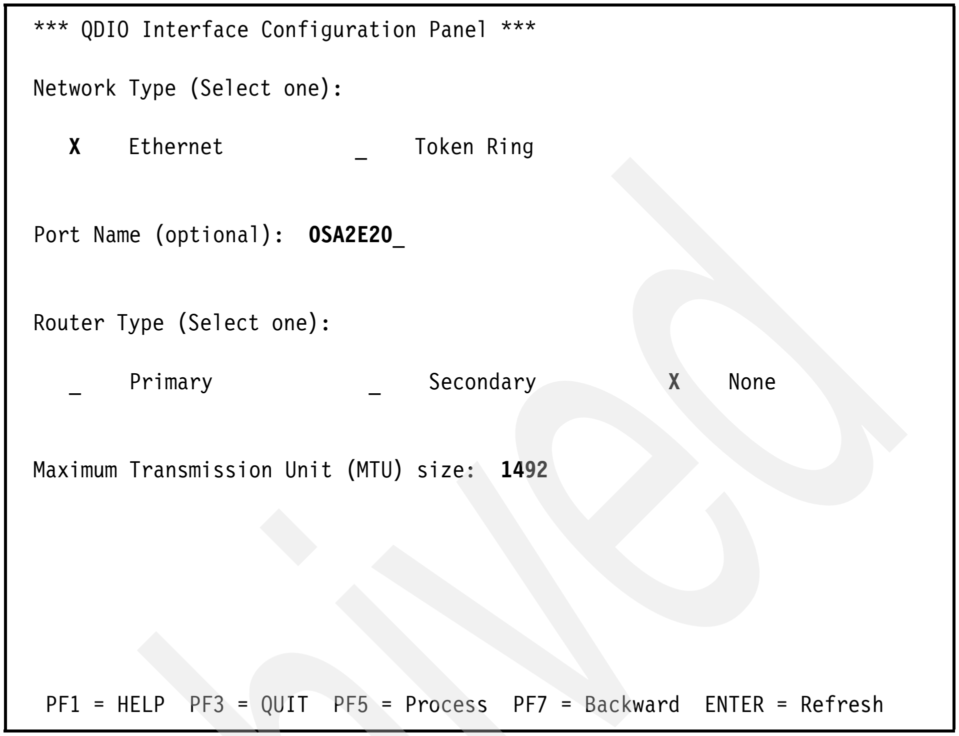
```
*** General Interface Configuration Panel ***

Interface Name:  OSA2E20L _______    Device Number:  2e28

IP Address:      9.12.4.155_____
Subnet Mask:     255.255.254.0__

Interface Type (Select one):

   X    QDIO                    _    LCS                     _    HiperSockets
   _    CLAW                    _    CTC









PF1 = HELP  PF3 = QUIT  PF7 = Backward  PF8 = Continue  ENTER = Refresh
```

Figure 1-24 The IPWIZARD General Interface Configuration Panel

In this menu, we configure a QDIO interface. We provide:

- Interface name
  We use the name OSA2E20IL.
- Device number
  The interface uses the OSA device address 2E28.
- IP address
  The interface is to be assigned IP address 9.12.4.155.
- Subnet mask
  We use mask 255.255.254.0.

Press **PF8** to continue to the QDIO Interface Configuration Panel shown in Figure 1-25 on page 31.



```
*** QDIO Interface Configuration Panel ***

Network Type (Select one):

   X    Ethernet            _    Token Ring


Port Name (optional):  OSA2E20_


Router Type (Select one):

   _    Primary             _    Secondary           X    None


Maximum Transmission Unit (MTU) size:  1492






 PF1 = HELP  PF3 = QUIT  PF5 = Process  PF7 = Backward  ENTER = Refresh
```

Figure 1-25 The IPWIZARD QDIO Interface Configuration Panel

In this menu, we provide details on the QDIO interface configuration. We provide:

- Network type
  This is an ethernet interface.
- Port name
  We provide the port name of the OSA-Express adapter.
- Router Type
  The adapter is neither a primary nor secondary router.
- MTU size
  We use an MTU size of 1492.

Press **PF8** to complete TCP/IP configuration. You are asked whether IPWIZARD should restart TCP/IP; enter **1** to restart TCP/IP. Configuration changes are written to the PROFILE TCP/IP and SYSTEM DTCPARMS files on the 193 minidisk.

1.5.2 Using the IFCONFIG command

Using the IFCONFIG command, you can display the TCP/IP configuration, and temporarily change the configuration without restarting the TCP/IP user.

> **Note:** The IFCONFIG command uses the NETSTAT and OBEYFILE commands. Changes made to the TCP/IP configuration using the IFCONFIG command may be altered by subsequent OBEYFILE commands or when TCP/IP is restarted. For details on the IFCONFIG command, see *z/VM: TCP/IP Planning and Customization*, SC24-6019.

To display the status of all configured network interfaces, execute the IFCONFIG command from the TCPMAINT user as shown in Figure 1-26.

```
IFCONFIG
OSA2E20L inet addr: 9.12.4.155 mask: 255.255.254.0
         UP BROADCAST MULTICAST MTU: 1492
         vdev: 2E28 rdev: 2E28 type: QDIO ETHERNET portname: OSA2E20
         ipv4 router type: NONROUTER ipv6: DISABLED
         cpu: 0 forwarding: ENABLED
         RX bytes: 631722 TX bytes: 2185584
Ready; T=0.02/0.02 19:21:02
```

Figure 1-26 Using the IFCONFIG command

The -SHOW option of the IFCONFIG command displays the TCP/IP server configuration file statements required to change an existing interface. Figure 1-27 illustrates usage of the -SHOW option.

```
IFCONFIG -SHOW
; Generated by <IFCONFIG OSA2E20L -show>
; 18 Aug 2004 19:34:07
DEVICE OSA2E20 CPU 0 OSD 2E28 PORTNAME OSA2E20 NONROUTER
LINK OSA2E20L QDIOETHERNET OSA2E20  MTU 1492
HOME
9.12.4.155  OSA2E20L
GATEWAY
Ready; T=0.02/0.02 19:34:07
```

Figure 1-27 The IPCONFIG -SHOW command

2

Networking enhancements for z/VM 5.1

In this chapter, we discuss networking enhancements available when using a Virtual Switch (VSWITCH). In z/VM 5.1, any quest connecting to a VSWITCH must first be authorized to access the VSWITCH. z/VM 5.1 also introduces Layer 2 functionality to the Virtual Switch.

2.1 Virtual Switch enhancements

The z/VM Virtual Switch (VSWITCH) introduced with z/VM V4.4 connects a Guest LAN to an external network using an OSA-Express port. Up to two additional OSA-Express ports can be specified as backups to the VSWITCH definition. The Linux guests connected to the VSWITCH are on the same subnet as the OSA-Express port or ports and other machines connected to that physical LAN segment.

In z/VM V4.4, VSWITCH operates at Layer 3 (network layer) of the OSI model. This only supports transport of IP packets (it only can be used for TCP/IP applications). Destinations are identified as IP addresses, and MAC addresses are not used. All hosts connected to a Layer 3 VSWITCH share the same OSA-Express MAC address. Outbound packets use the OSA-Express's MAC as the source MAC address. Inbound packets are forwarded by the OSA-Express to the guest based on the destination IP address in the packet.

In z/VM V5.1, the VSWITCH implementation is extended to operate at Layer 2 (data link layer) of the OSI model. In Layer 2 mode, the VSWITCH:

- Uses the MAC destination address to send and receive Ethernet frames, even between the virtual adapters and adapters on the physical portions of the LAN segment.
- Transports Ethernet frames (not IP datagrams) to and from the operating system TCP/IP stack and the physical network.
- Does not offload ARP processing to the OSA-Express adapter; ARP processing performed by the operating system TCP/IP stack.
- Supports MAC level unicast, multicast, and broadcast.

Unlike a Layer 3 VSWITCH, a Layer 2 VSWITCH does not require a router running on the internal Guest LAN segment. This reduces network latency and overall CPU consumption. Removing the router also means that you no longer need specialized skills to configure and administer a VM-based or Linux-based router.

2.2 VSWITCH configuration

Virtual Switch (VSWITCH) is a networking topology originally introduced in z/VM 4.4. The VSWITCH bridges a virtual z/VM Guest LAN to an external physical LAN connected by an OSA-Express adapter. Figure 2-1 on page 35 illustrates a VSWITCH topology.

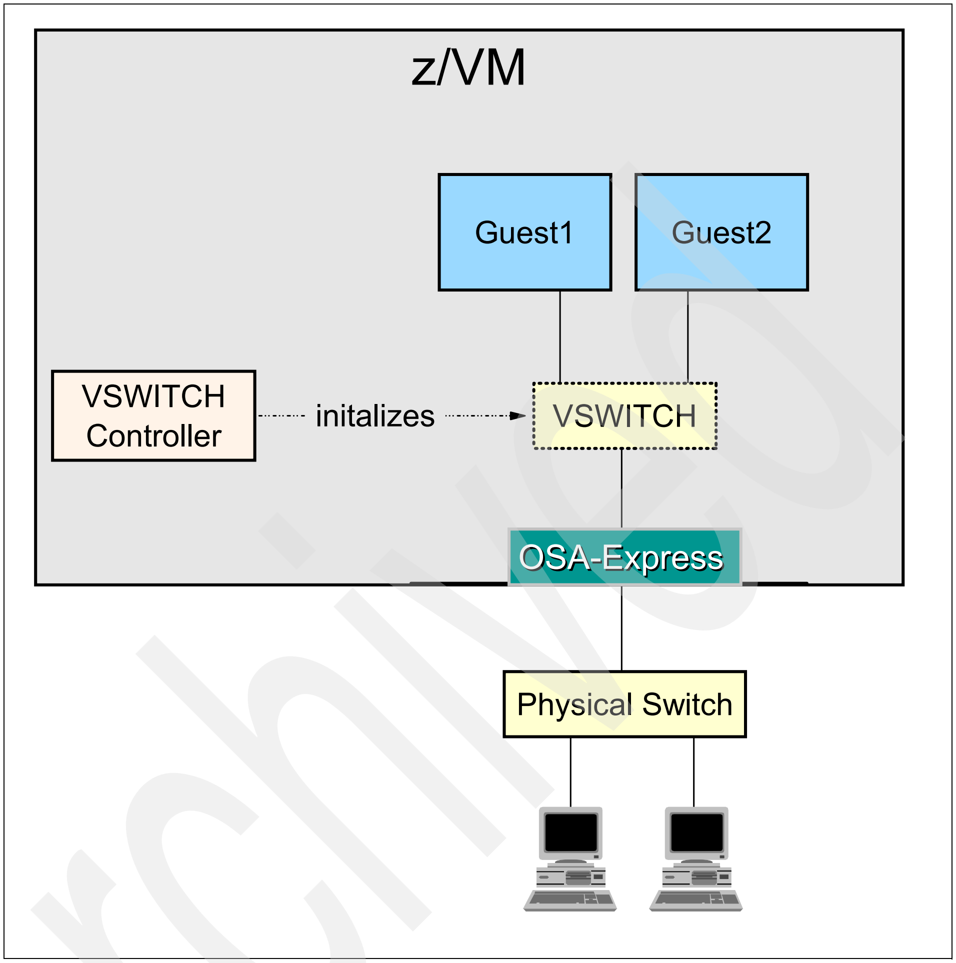


Figure 2-1 VSWITCH topology

2.3 Define the VSWITCH controller

The VSWITCH controller is a VM guest running a TCP/IP stack. The controller initializes the interface to the OSA-Express adapter. In order to use a VSWITCH, at least TCP/IP stack must be configured as a controller.

> **Important:** The VSWITCH controller is only involved in device initialization. Once initialized, all packets pass directly between connection endpoints *without* passing through the VSWITCH controller.

The VSWITCH controller must have the IUCV *VSWITCH statement added to its directory entry. This authorizes the service machine to connect to the *VSWITCH system service. Figure 2-2 shows a sample VSWITCH controller directory entry.

```
USER TCPIP TCPIP 128M 256M ABG
 INCLUDE TCPCMSU
 OPTION QUICKDSP SVMSTAT MAXCONN 1024 DIAG98 APPLMON
 SHARE RELATIVE 3000
 IUCV ALLOW
 IUCV ANY PRIORITY
 IUCV *CCS PRIORITY MSGLIMIT 255
 IUCV *VSWITCH MSGLIMIT 65535
 LINK 5VMTCP10 491 491 RR
 LINK 5VMTCP10 492 492 RR
 LINK TCPMAINT 591 591 RR
 LINK TCPMAINT 592 592 RR
 LINK TCPMAINT 198 198 RR
 MDISK 191 3390 2247 005 510W01  MR RTCPIP   WTCPIP   MTCPIP
```

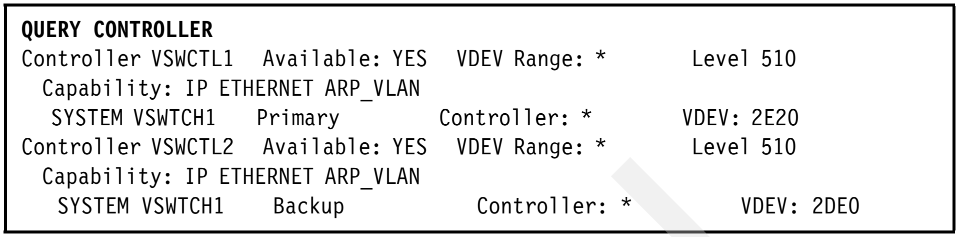
Figure 2-2   Directory entry for a VSWITCH controller

The VSWITCH CONTROLLER ON statement must also appear in the TCP/IP profile for the controller service machine. Figure 2-3 illustrates the VSWITCH CONTROLLER statement in the TCP/IP PROFILE file.

```
; ----------------------------------------------------------------------
; Define whether or not a stack is available to control a CP-defined
; Virtual Switch's connection to a real LAN segment through an
; OSA Express device.  The range of virtual addresses that are to be
; used for such a connection can optionally be specified with the
; VSWITCH statement.
; ----------------------------------------------------------------------
;
VSWITCH CONTROLLER ON
```

Figure 2-3   VSWITCH CONTROLLER entry in TCP/IP PROFILE

As shown in Figure 2-4 on page 37, the QUERY CONTROLLER command reports virtual machines configured as VSWITCH controllers.



```
QUERY CONTROLLER
Controller VSWCTL1  Available: YES  VDEV Range: *      Level 510
  Capability: IP ETHERNET ARP_VLAN
   SYSTEM VSWTCH1   Primary        Controller: *       VDEV: 2E20
Controller VSWCTL2  Available: YES  VDEV Range: *      Level 510
  Capability: IP ETHERNET ARP_VLAN
    SYSTEM VSWTCH1    Backup          Controller: *         VDEV: 2DE0
```

Figure 2-4 The QUERY CONTROLLER command

2.4 Define the VSWITCH

To create a VSWITCH, use the CP DEFINE VSWITCH command (this requires class B privilege). Alternative, the DEFINE VSWITCH statement can be added to the SYSTEM CONFIG file to automatically define the VSWITCH at system IPL.

Syntax of the DEFINE VSWITCH statement and command is:

```
DEFINE VSWITCH switchname [ operands ]
```

Where:

switchname Defines the VSWI TCH name.

operands Defines attributes of the VSWITCH.

Command operands are summarized in Table 2-1.

Table 2-1 Common operands of the DEFINE VSWITCH statement

Operands	Description
RDEV *rdev-list*	A real device address to be used to connect the Virtual Switch to a QDIO OSA-Express device. You can specify a maximum of three real device numbers. Each real device address represents a trio of devices. For example, specifying RDEV 111 222 333 means that the first devices, 111-113, are used to provide the connection to the real hardware LAN segment. If there is a problem with the connection, devices 222-224 are used next to provide the connection, and if those devices fail to connect, devices 333-335 are used. This feature provides dynamic recovery for OSA-Express device failures.
CONnect	Indicates that the device identified by the RDEV keyword must be activated, and traffic must flow through the device to the real LAN segment.

Operands	Description
CONTRoller * or CONTRoller *userid*	Identifies the z/VM user ID that controls the OSA-Express device connected at the device address identified by rdev. CONTROLLER * means that CP selects from any of the eligible z/VM TCP/IP stacks. If you specify multiple real devices on the RDEV keyword, specify CONTROLLER *, or allow it to default. The controller functions are then spread across multiple z/VM TCP/IP stacks, providing more flexibility in case of a failure.
IP or ETHernet	Indicates whether the transport mode for the Virtual Switch is ETHERNET or IP. An ETHERNET Virtual Switch operates at the Layer 2 level of the OSI model, and an IP Virtual Switch operates at Layer 3.[a]
VLAN *defvid*	Defines the default VLAN id associated with untagged frames on this switch. The default is VLAN UNAWARE, which indicates that the virtual switch will ignore VLAN tags.
PORTType *type*	Defines the default port type for guests attached to this virtual switch. The *type* can be ACCESS or TRUNK. This operand is not valid if VLAN UNAWARE is specified.
PORTname *portname*	A 1-to-8 character name that identifies the OSA-Express adapter. You can specify a maximum of three port names. Multiple port names are used when different port names are needed for the multiple rdevs specified on the RDEV operand.[b]

[a] For details on Layer 2 Virtual Switches, see “Configuring a Layer 2 VSWITCH in z/VM” on page 42.

[b] For details on port names, see “Supplying OSA port names” on page 39

To define a Layer 2 Virtual Switch named “SW1” at system IPL, add the following statement to the SYSTEM CONFIG file: add the following DEFINE VSWITCH statement to the System:

```
DEFINE VSWITCH SW1 RDEV 2E20 CONTROLLER * ETHERNET
```

This connects the Virtual Switch to OSA devices 2E20-2E22. The VSWITCH is controlled by the default VSWITCH controller. When executed as a command, the Virtual Switch is created immediately:

```
DEFINE VSWITCH SW1 RDEV 2E20 CONTROLLER * ETHERNET
VSWITCH SYSTEM SW1 is created
Ready; T=0.01/0.01 10:47:59
TCPIP   : 10:47:59 DTCOSD360I VSWITCH-OSD link added for SW32E20DEV
TCPIP   : 10:47:59 DTCOSD080I VSWITCH-OSD initializing:
TCPIP   : 10:47:59 DTCPRI385I  Device SW32E20DEV:
```

```
TCPIP   : 10:47:59 DTCPRI386I     Type: VSWITCH-OSD, Status: Not started
TCPIP   : 10:47:59 DTCPRI387I     Envelope queue size: 0
TCPIP   : 10:47:59 DTCPRI388I     Address: 2E20
TCPIP   : 10:47:59 DTCQDI001I QDIO device SW32E20DEV device number 2E22:
TCPIP   : 10:47:59 DTCQDI007I   Enabled for QDIO data transfers
TCPIP   : 10:47:59 DTCOSD341I Obtained MAC address 00096B1A1F38 for
device SW32E
20DEV
HCPSWU2830I VSWITCH SYSTEM SW1 status is ready.
HCPSWU2830I TCPIP is VSWITCH controller.
```

2.4.1 Supplying OSA port names

The portname parameter is required on all S/390 G5 and G6 servers. With z800 and z900 systems, the portname parameter is not required if driver level 3G, EC stream J11204 MCL032 (OSA Level 3.333) is installed. The portname is never required for z890, z990, System z9, or later servers. If required, the portname must be one to eight uppercase characters in length and must match the OSA port name specified by all systems sharing the OSA. If not required, we recommend omitting the parameter. For details on the portname parameter, refer to *OSA-Express MCL Enhancements - October 2003* at:

http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/FLASH10250

2.5 Authorize guests to access the VSWITCH

Before connecting to a Virtual Switch, a guest must first be authorized to the VSWITCH. This can be done by CP security or by an external security manager (ESM). Authorization checking is done when the guest attempts to connect to the Virtual Switch using the COUPLE command.

Important: If an ESM is in place, the security definitions made by the SET VSWITCH command are overridden by the ESM definitions.

Figure 2-5 on page 40 depicts the authorization logic used by the COUPLE command.

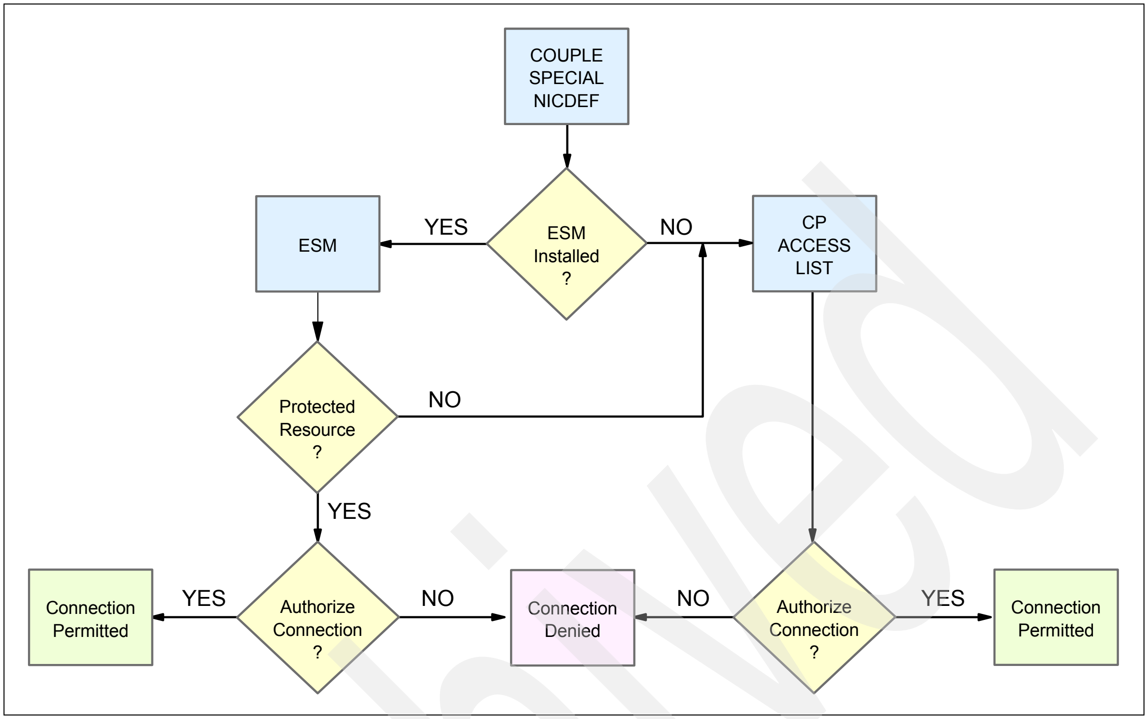


Figure 2-5 Authorization logic for the COUPLE command

2.5.1 Using CP authorization

Without an ESM (such as RACF®), authorization is granted using the SET VSWITCH command with the GRANT option. For example, to authorize guest LNXSU4 access to Virtual Switch SW1, use:

```
SET VSWITCH SW1 GRANT LNXSU4
Command complete
Ready; T=0.01/0.01 11:32:59
```

To remove a guest's access, use the SET VSWITCH command with the REVOKE option:

```
SET VSWITCH SW1 REVOKE LNXSU4
Command complete
Ready; T=0.01/0.01 11:35:1
```

To see virtual machines authorized to access a Virtual Switch, use the QUERY VSWITCH command with the ACCESS option as shown in Figure 2-6 on page 41.

```
Q VSWITCH SW1 ACCESS
VSWITCH SYSTEM SW1      Type: VSWITCH Connected: 2    Maxconn: INFINITE
  PERSISTENT  RESTRICTED    NONROUTER                 Accounting: OFF
  VLAN Unaware
  State: Ready
  IPTimeout: 5         QueueStorage: 8
  Portname: OSA2E20    RDEV: 2E2C Controller: TCPIP    VDEV:  2E2C
    Authorized userids:
      LNXRH1   LNXRH2   LNXRH3   LNXRH4   LNXRH5   LNXSU1
      LNXSU2   LNXSU3   LNXSU4   LNXSU5   SYSTEM
Ready; T=0.01/0.01 10:23:01
```

Figure 2-6 Display of the authorization list for VSWITCH SW1

2.5.2 Using RACF authorization

z/VM 5.1 provides a resource class called VMLAN. This class is used to protect the COUPLE function to VSWITCH and VLAN. To establish RACF security for VSWITCH:

1. Define a profile for the VSWITCH
2. Grant access to specific virtual machines
3. Activate the RACF VMLAN class

> **Attention:** RACF is pre-installed on z/VM version 5.1, but is disabled. Activate RACF only if you have a valid license.

Define a profile for the VSWITCH

First, define a RACF profile for the Virtual Switch:

```
RAC RDEFINE VMLAN SYSTEM.SW1 UACC(NONE)
Ready; T=0.01/0.01 15:18:43
```

A profile for a specific VLAN can be defined for the Virtual Switch:

```
RAC RDEFINE VMLAN SYSTEM.SW1.0001 UACC(NONE)
Ready; T=0.01/0.01 15:18:43
```

In this example, we define a RACF profile for VSWITCH SW1 and VLAN 0001.

Grant access to specific virtual machines

Next, allow specific virtual machines to access the Virtual Switch:

```
RAC PERMIT SYSTEM.SW1 CLASS(VMLAN) ACCESS(UPDATE) ID(LNXSU4)
Ready; T=0.01/0.01 13:34:47
```

This example grants access to user LNXSU4 to VSWITCH SW. Once authorized, the virtual machine can use the CP COUPLE command to connect to the VSWITCH. If an unauthorized user issues the COUPLE command, an error message is generated:

```
COUPLE 3000 SYSTEM SW1
RPIMGR032E YOU ARE NOT AUTHORIZED TO COUPLE TO SYSTEM.SW1
HCPCPL6011E You are not authorized to COUPLE to SYSTEM SW1
Ready(06011); T=0.01/0.01 11:13:58
```

To revoke access to a VSWITCH, use the RACF PERMIT command with the DELETE option:

```
RAC PERMIT SYSTEM.SW1 CLASS(VMLAN) ACCESS(UPDATE) ID(LNXSU4) DELETE
Ready; T=0.01/0.01 13:34:47
```

Activate the RACF VMLAN class

Once the RACF profile for the VSWITCH is defined, activate the RACF class VMLAN:

```
RAC SETROPTS CLASSACT(VMLAN)
Ready; T=0.01/0.01 13:54:2
```

> **Attention:** Before you activate the VMLAN class, you must have defined a profile for each VSWITCH. Otherwise guests will be unable to connect to the VSWITCH.

2.6 Connect the guest to the VSWITCH

Once authorized, a guest must first define a virtual QDIO NIC using the DEFINE NIC command:

```
CP DEF NIC 3000 QDIO
```

Next, connect the virtual NIC to Virtual Switch using the COUPLE command:

```
COUPLE 3000 SYSTEM SW1
NIC 3000 is connected to VSWITCH SYSTEM SW1
Ready; T=0.01/0.01 11:49:46
```

2.7 Configuring a Layer 2 VSWITCH in z/VM

In Ethernet mode, a VSWITCH operates at Layer 2. Each guest connected to the Virtual Switch is assigned a unique MAC address. Assignment of the guest's MAC address is performed by z/VM under the control of the LAN administrator:

- The VMLAN statement in the SYSTEM CONFIG file provides the MACPREFIX and MACIDRANGE operands to control generation of the system-wide range of MAC addresses.
- The NICDEF statement in the user directory entry provides the MACID operand to control the specific MAC address assigned to a guest.

> **Restriction:** Port sharing is only supported between Virtual Switches of the same transport mode. Attempting to communicate between a Layer 2 VSWITCH and a Layer 3 VSWITCH sharing the same OSA-Express adapter results in a network time-out. To resolve this, ensure the Layer 2 VSWITCH and Layer 3 VSWITCH use separate OSA-Express adapters.

2.7.1 The VMLAN statement

The MACPREFIX operand of VMLAN statement allows the administrator to specify a three byte manufacturer ID prefix for all MAC addresses generated on the z/VM system. The MACIDRANGE operand controls the range of values used when generating unique MAC address. The syntax to control MAC generation is:

```
VMLAN MACIDRange SYSTEM xxxx-xxxx [ USER xxxx-xxxx ]
```

Operands to the statement are explained in Table 2-2 on page 44.

Table 2-2 Operands to VMLAN statement to control MAC generation

Operand	Description
MACPREFIX *macprefix*	Specifies the three byte prefix (manufacturer ID) used when generating locally administered MAC addresses on the system. This must be six hexadecimal digits within the range of 020000 through 02FFFF (inclusive). In combination with the MAC ID used on the NICDEF directory statement, the MACPREFIX allows unique identification of virtual adapters within a network. If MACPREFIX is not specified, the default is 020000 (02-00-00).
MACIDRANGE SYSTEM *xxxxxx-xxxxxx* USER *xxxxxx-xxxxxx*	
SYSTEM *xxxxxx-xxxxxx*	The range of identifiers (up to six hexadecimal digits each) to be used by CP when generating the unique identifier part (last six hexadecimal digits) of a virtual adapter MAC address. If a SYSTEM MACIDRANGE is not specified, CP creates unique identifiers in any range (000001-FFFFFF).
USER *xxxxxx-xxxxxx*	The subset of the SYSTEM range of identifiers reserved for user definition of MACIDs in the NICDEF directory statement. When specified, CP does not assign MACIDs within this USER range during creation of virtual adapters defined dynamically (DEFINE NIC) or with the NICDEF (or SPECIAL) directory statement without the MACID operand. In these cases, CP generates a unique identifier for the adapter outside of the USER range. Any MACID values specified on a NICDEF directory statement must be within the USER range or the virtual adapter is not defined during LOGON processing. If a USER MACIDRANGE is not specified, CP creates unique identifiers within the SYSTEM MACIDRANGE.

Note: If you run multiple z/VM systems on the same CEC, you should change the MACPREFIX of each system to avoid MAC address duplication.

As shown in Figure 2-7 on page 45, the QUERY VMLAN command reports the configured range of system-wide MAC address prefixes.

```
QUERY VMLAN
VMLAN maintenance level:
  Latest Service: Base
VMLAN MAC address assignment:
  MACADDR Prefix: 02EEEE
  MACIDRANGE SYSTEM: 100000-1FFFFF
             USER:   000000-000000
VMLAN default accounting status:
  SYSTEM Accounting: OFF      USER Accounting: OFF
VMLAN general activity:
  PERSISTENT Limit: INFINITE   Current: 1
  TRANSIENT  Limit: INFINITE   Current: 0
```

Figure 2-7 The QUERY VMLAN command

2.7.2 Assigning a specific MAC address

Use the MACID operand of the NICDEF statement to assign a specific MAC address to a guest. During LOGON, three byte MACID is appended to the system three byte MACPREFIX to form a unique MAC Address for the NIC. If omitted, CP generates a unique MAC from the range specified in the VMLAN statement.

2.8 Configuring a Layer 2 VSWITCH in Linux

Layer 2 VSWITCH support is provided by the Linux qeth device driver. An additional “layer2” keywork is used to enable Layer 2 support. In SELS8, add the layer2 keyword to the /etc/chandev.conf file. The keyword is placed on the entry for the interface to the Layer 2 VSWITCH:

```
noauto;qeth0,0xc204,0xc205,0xc206;add_parms,0x10,0xc204,0xc206,layer2
```

> **Note:** The qeth device also accepts the new “no_layer2” keyword (which the interface at Layer 3). If neither keyword is specified, the qeth driver operates at Layer 3 (IP).

When the interface is configured, reboot the Linux guest. To verify the interface is correctly configured, use the using the **ifconfig** command:

```
# ifconfig -a
eth0      Link encap:Ethernet  HWaddr 02:EE:EE:10:00:00
          inet6 addr: fe80::2ee:ee00:10:0/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:4 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:100
          RX bytes:0 (0.0 b)  TX bytes:416 (416.0 b)
          Interrupt:7
```

When configured for Layer 2, the assigned MAC address appears in the HWaddr field.

3

Cryptographic hardware support

In this chapter, we discuss using cryptographic hardware acceleration for Secure Sockets Layer (SSL) transactions with Apache. Topics include:

- Configuring the LPAR to use cryptographic devices
- Configuring the z/VM system
- Configuring SSL for Apache

3.1 Configuring the LPAR to use cryptographic devices

Cryptographic coprocessors can be accessed by Linux from an LPAR or as a z/VM guest. The LPAR running Linux or z/VM must be configured to enable the cryptographic hardware. Linux access the cryptographic hardware using the z90crypt device driver. The driver supports the following cryptographic devices:

- PCI Cryptographic Coprocessor (PCICC)
- PCI Cryptographic Accelerator (PCICA)
- PCI-X Cryptographic Coprocessor (PCIXCC)
- Cryptographic Express2 Coprocessor (CEX2C)
- Cryptographic Express2 Accelerator (CEX2A)

The device driver uses hardware to enhance performance during SSL handshaking. It supports only clear-key functions (cryptographic operations used during the handshake) for:

- Public/private key encryption and decryption.
- RSA exponentiation

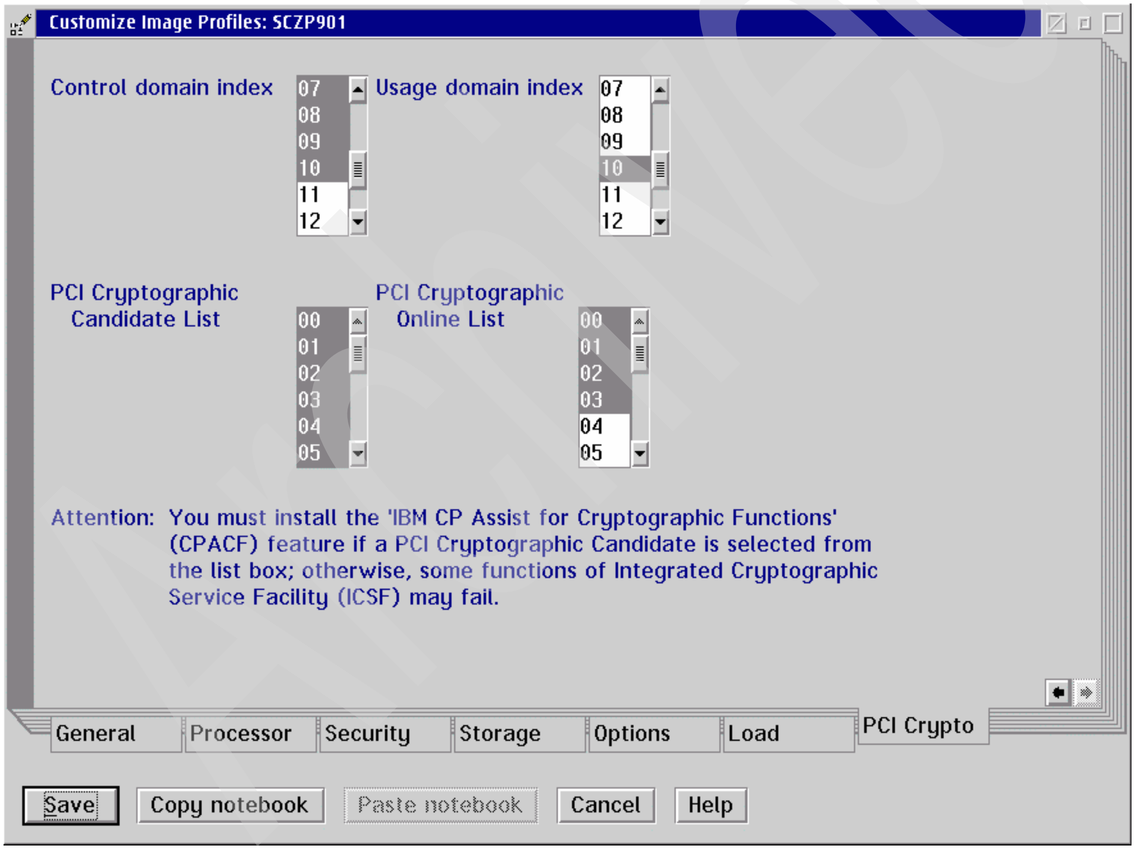
Before configuring the LPAR, be aware of some restrictions:

- PCICC devices are not supported on the IBM @server® zSeries 990.
- If a PCICA device is installed, z/VM hides any PCICC and PCIXCC devices from a guest.
- PCIXCC devices are detected but ignored when running with z90crypt versions earlier than 1.2.1.
- PCIXCC is only available for z990 as "PCIX Cryptographic Coprocessor (PCIXCC) feature (#0868)".
- Installation of the CP Assist for Cryptographic Functions (CPACF) DES/TDES enablement (feature code 3863) is required to enable use of PCIXCC and PCICA features. Feature code 3863 enables DES and TDES algorithms on the CPACF (the SHA-1algorithm is always enabled).
- The z990 supports up to two PCICA features per I/O cage. This allows for a maximum of six PCICA features, or twelve PCICA coprocessors per z990 server.
- The maximum number of PCIXCC features (or cryptographic coprocessors) per I/O cage is four; the maximum number of PCIXCC features per z990 is also four.
- The total number of cryptographic features may not exceed eight per z990 for any combination of PCIXCC and PCICA features. In addition, any combination of PCIXCC, PCICA, OSA-Express and FICON-Express features may not exceed 20 features per I/O cage, or 60 features per z990 server.

- On the z990, the PCIXCC and PCICA features do not use CHPIDs from the Logical Channel Subsystem pool, but are assigned as follows:
  - One PCHID is assigned per PCIXCC feature.
  - Two PCHIDs are assigned per PCICA feature.

3.1.1 LPAR definition

The z990 server only operates in LPAR mode. For each logical partition that accesses a cryptographic coprocessor (either PCICA or PCIXCC), you must customize the partition image profile. This is done from the Hardware Management Console and the Support Element. First, start an HMC session and enter the LPAR definitions panel, as shown in Figure 3-1.



Customize Image Profiles: SCZP901

Control domain index: 07, 08, 09, 10, 11, 12

Usage domain index: 07, 08, 09, 10, 11, 12

PCI Cryptographic Candidate List: 00, 01, 02, 03, 04, 05

PCI Cryptographic Online List: 00, 01, 02, 03, 04, 05

Attention: You must install the 'IBM CP Assist for Cryptographic Functions' (CPACF) feature if a PCI Cryptographic Candidate is selected from the list box; otherwise, some functions of Integrated Cryptographic Service Facility (ICSF) may fail.

Figure 3-1 Customize LPAR

The **PCI Cryptographic Online List** identifies cryptographic coprocessors numbers brought online automatically during logical partition activation. Available coprocessor numbers are shown in the **PCI Cryptographic Candidate List**. Be sure to select the PCIXCC cards in the **PCI Cryptographic Online List**.

Once the coprocessors are selected, deactivate and activate the partition. Then IPL the system.

3.2 Configuring the z/VM system

z/VM automatically detects installed cryptography hardware. To determine available cryptography hardware, use the **QUERY CRYPTO** command:

```
Q CRYPTO
No CAM or DAC Crypto Facilities are installed
Crypto Adjunct Processor Instructions are installed
Ready; T=0.01/0.01 10:34:59
```

For detailed information on installed Adjunct Processors (AP), use the **QUERY CRYPTO APQS** command. In the following example, we see two available PCIXCC queues

```
Q CRYPTO AP
AP 02 PCIXCC Queue 14 is installed
AP 03 PCIXCC Queue 14 is installed
Ready; T=0.01/0.01 11:03:54
```

Restriction: If both PCICA and PCIXCC hardware are installed, z/VM hides the PCIXCC cards:

```
Q CRYPTO AP
AP 00 PCICA Queue 00 is installed
AP 01 PCICA Queue 01 is installed
AP 02 PCIXCC Queue 14 is superseded by PCICA
AP 03 PCIXCC Queue 14 is superseded by PCICA
```

To enable a z/VM guest access to a virtual AP, add the CRYPTO APVIRT statement to the user's directory entry as shown in Figure 3-2 on page 51.

```
USER LNXSU4 XXXXXX 768M 2G G
 ACCOUNT 1 SYSPROG
 MACH ESA
 IPL 190
 CRYPTO    APVIRT
 CONSOLE 0009 3215 T MAINT
 SPECIAL 3000 QDIO 3 SYSTEM SW1
 SPOOL 000C 2540 READER *
 SPOOL 000D 2540 PUNCH A
 SPOOL 000E 1403 A
 LINK MAINT    0190 0190 RR
 LINK MAINT    019D 019D RR
 LINK MAINT    019E 019E RR
 LINK TCPMAINT 0592 0592 RR
 MDISK 0191 3390 61 20 DK15D2 MR
 MDISK 0201 3390 1 200 LX1512 MR
 MDISK 0202 3390 201 3138 LX1512 MR
 MDISK 0300 FB-512 V-DISK 300000 WV
```

Figure 3-2 Sample directory entry for using cryptography

To display the status of the adjunct processor, issue the **`QUERY VIRTUAL CRYPTO`** as shown in Figure 3-3.

```
Q V CRYPTO
No CAM or DAC Crypto Facilities defined
AP 37 PCIXCC Queue 13 shared
Ready; T=0.01/0.01 11:30:54
```

Figure 3-3 Query installed virtual cryptographic processors

3.3 Configuring SSL for Apache

The Apache Web server can use hardware cryptographic devices to improve performance during SSL handshake negotiation. It utilizes OpenSSL to access the cryptographic hardware. The software to support cryptographic hardware is shown in Figure 3-4 on page 52.

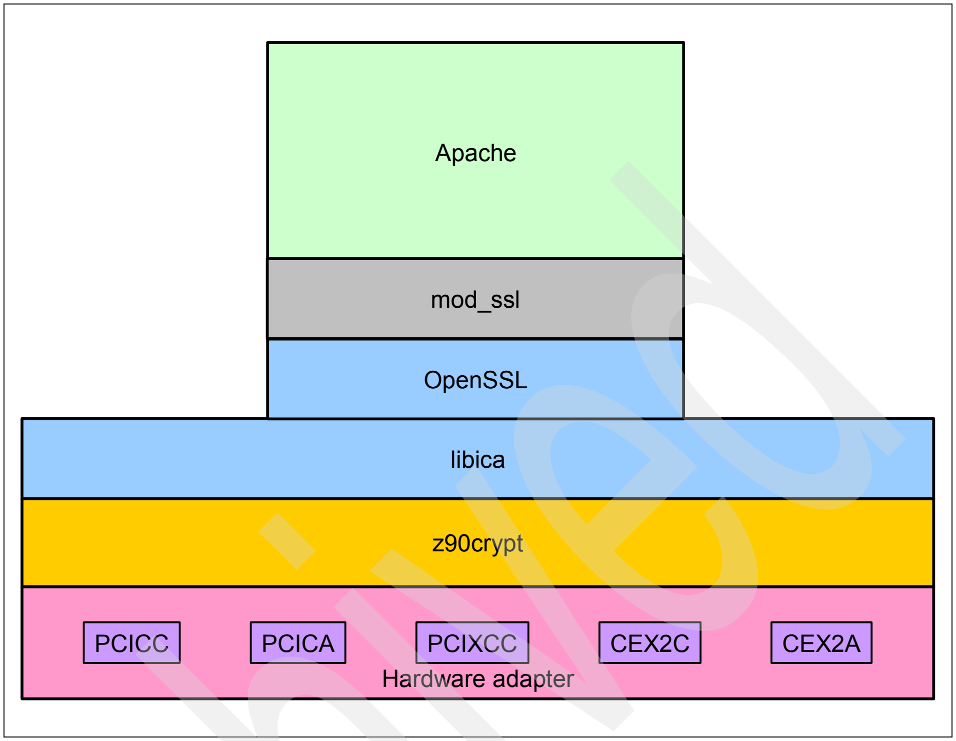


Figure 3-4 Using cryptographic hardware with Apache

The Apache mod_ssl module provides the interface to OpenSSL. The OpenSSL interface engine uses the libica shared library to access the device through the z90crypt device driver. With SLES8, the required RPM packages include:

- The mod_ssl RPM package
- The openssl RPM package
- The libica RM package

To enable hardware cryptographic acceleration, we:

- Load the hardware device driver
- Verify OpenSSL supports IBM cryptographic hardware
- Generate certificate and keys for SSL
- Configure the Apache Web server

3.3.1 Load the hardware device driver

T0 enable the PCIXCC card for Linux, first load the z90crypt device driver using the **rcz90crypt** command:

```
# rcz90crypt start
```

To unload the device driver is loaded, use the stop option:

```
# rcz90crypt stop
```

To see if the device driver is loaded, supply the status option:

```
# rcz90crypt status
Checking for module z90crypt: z90crypt 28928   0  (unused) running
```

To automatically load the device driver at system boot, use the **insserv** command:

```
# insserv z90crypt
# chkconfig z90crypt
z90crypt  on
```

Once loaded, device status can be checked using the /proc/driver/z90crypt interface as shown in Figure 3-5 on page 54.

```
# cat /proc/driver/z90crypt

z90crypt version: 1.2.2
Cryptographic domain: 4
Total device count: 1
PCICA count: 0
PCICC count: 0
PCIXCC count: 1
requestq count: 0
pendingq count: 0
Total open handles: 1


Online devices: 1 means PCICA, 2 means PCICC, 3 means PCIXCC
    0000000000000000 0000000000000000 0000000000000000 0000003000000000


Waiting work element counts
    0000000000000000 0000000000000000 0000000000000000 0000000000000000


Per-device successfully completed request counts
    00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
    00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
    00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
    00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
    00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
    00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
    00000000 00000000 00000000 00000000 00000000 00000000 00000002 00000000
    00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
```

Figure 3-5 Status page of the z90crypt driver

Output indicates the driver version. The “`Total open handles:`” value indicates the number of applications using the device (zero indicates no application is currently using the z90crypt driver).

> **Attention:** PCIXCC device support requires at least Version 1.2.1 of the z90crypt driver. Otherwise, the card is detected but is unusable by applications.

3.3.2 Verify OpenSSL supports IBM cryptographic hardware

The OpenSSL engine interface links to a shared library to access a specific hardware cryptographic device. The IBM Cryptographic Accelerator (ICA) library provides the shared library (libica) to interface to System z9 and zSeries cryptographic adapters. OpenSSL is installed as part of the default SLES8 system. Once the z90crypt device driver is loaded, the With SLES8 SP1, the OpenSSL interface engine is enable to use IBM cryptographic hardware. After the z90crypt device driver is installed, verify that the OpenSSL engine interface supports IBM cryptographic devices using the **openssl** command as shown in Figure 3-6.

```
# openssl speed rsa1024 -engine ibmca -elapsed
engine "ibmca" set.
You have chosen to measure elapsed time instead of user CPU time.
To get the most accurate results, try to run this
program when this computer is idle.
Doing 1024 bit private rsa's for 10s: 543 1024 bit private RSA's in 10.01s
Doing 1024 bit public rsa's for 10s: 7948 1024 bit public RSA's in 10.00s
OpenSSL 0.9.6g [engine] 9 Aug 2002
built on: Tue Oct  7 13:54:36 UTC 2003
options:bn(64,32) md2(int) rc4(ptr,int) des(idx,cisc,4,long) blowfish(idx)
compiler: gcc -fPIC -DTHREADS -DDSO_DLFCN -DHAVE_DLFCN_H -DB_ENDIAN -DNO_RC5 -DNO_IDEA -O2
-fsigned-char -fomit-frame-pointer -DTERMIO -Wall
                  sign    verify    sign/s verify/s
rsa 1024 bits   0.0184s   0.0013s     54.2    794.9
```

Figure 3-6 Verification that IBM hardware is supported in OpenSSL

The string “engine ‘ibmca’ set.” indicates IBM cryptographic hardware is supported.

3.3.3 Generate certificate and keys for SSL

For a secure connection, we need to generate a certificate and the public/private key pair as shown in Figure 3-7 on page 56.

```
# cd /etc/httpd/ssl.crt/
# openssl req -new > new.cert.csr
Using configuration from /etc/ssl/openssl.cnf
Generating a 1024 bit RSA private key
.............++++++
...++++++
writing new private key to 'privkey.pem'
Enter PEM pass phrase:
Verifying password - Enter PEM pass phrase:
-----
You are about to be asked to enter information that will be incorporated
into your certificate request.
What you are about to enter is what is called a Distinguished Name or a DN.
There are quite a few fields but you can leave some blank
For some fields there will be a default value,
If you enter '.', the field will be left blank.
-----
Country Name (2 letter code) [AU]:US
State or Province Name (full name) [Some-State]:New York
Locality Name (eg, city) []:Poughkeepsie
Organization Name (eg, company) [Internet Widgits Pty Ltd]:IBM Corp.
Organizational Unit Name (eg, section) []:ITSO
Common Name (eg, YOUR name) []:Lutz Kuehner
Email Address []:Lutz.Kuehner@de.ibm.com

Please enter the following 'extra' attributes
to be sent with your certificate request
A challenge password []:itso
An optional company name []:
```

Figure 3-7 Generating a private key

Once created, remove the passphrase from the key from the certificate as shown in Figure 3-8.

```
# openssl rsa -in privkey.pem -out new.cert.key
read RSA key
Enter PEM pass phrase:
writing RSA key
```

Figure 3-8 Remove pass phrase

Next, convert the key to a certificate as shown in Figure 3-9 on page 57.

```
# openssl x509 -in new.cert.csr -out new.cert.cert -req -signkey new.cert.key -days 365
Signature ok
subject=/C=US/ST=New York/L=Poughkeepsie/O=IBM Corp./OU=ITSO/CN=Lutz
Kuehner/Email=Lutz.Kuehner@de.ibm.com
Getting Private key
```

Figure 3-9 Converting a key to a certificate

When complete, the contents of the current directory is similar to that shown in Figure 3-10.

```
-rw-r--r--    1 root     root          977 Aug 12 12:52 new.cert.cert
-rw-r--r--    1 root     root          749 Aug 12 12:46 new.cert.csr
-rw-r--r--    1 root     root          891 Aug 12 12:50 new.cert.key
-rw-r--r--    1 root     root          963 Aug 12 12:46 privkey.pem
```

Figure 3-10 Stored files after creation

The new.cert.cert file is the generated certificate Apache is to use; the new.cert.key file is the private key for the certificate.

3.3.4 Configure the Apache Web server

The Apache Web server is installed in our system with RPM package apache-1.3.26-36. First, we must modify the /etc/httpd/httpd.conf Web server configuration file. Two modifications are required:

- **Identify the location of the certificate to the Web server.**

  Find the <VirtualHost _default_:443> directive in the /etc/httpd/httpd.conf file. Set the SSLCertificateFile directive to the certificate generated in3.3.3, “Generate certificate and keys for SSL” on page 55. Set the SSLCertificateKeyFile directive to the certificate’s private key file. An example is shown below:

  ```
  <VirtualHost _default_:443>
  .
  .
  .
  SSLCertificateFile    /etc/httpd/ssl.crt/new.cert.cert
  SSLCertificateKeyFile /etc/httpd/ssl.key/new.cert.key
  ```

- **Configure OpenSSL to use the libica engine interface.**

  To enable OpenSSL to use the libica engine interface, set the SSLCryptoDevice directive (in the <IfModule mod_ssl.c> section) to “ibmca”. An example is shown below:

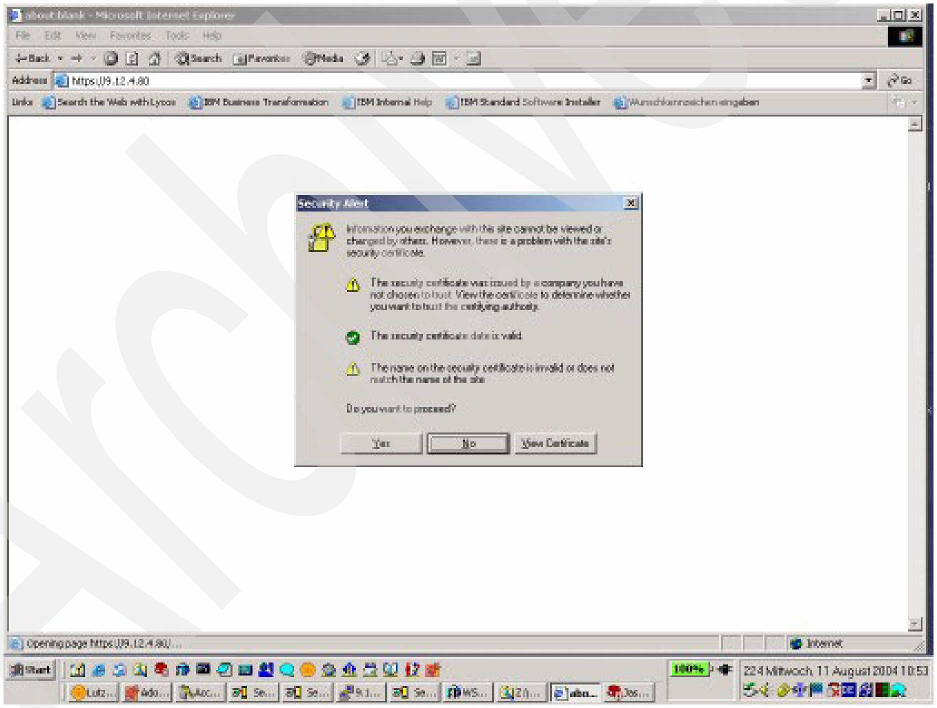
```
<IfModule mod_ssl.c>
.
.
.
SSLCryptoDevice ibmca
```

To start the Apache Web server with SSL support, use the apachectl command:

```
apachectl startssl
/usr/sbin/apachectl startssl: httpd started
```

3.3.5 Accessing the Web server

To access the Apache Web server, point your browser to the URL https://*hostname*. At this point, you are prompted to accept the certificate as shown in Figure 3-11.


about:blank - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address https://9.12.4.80
Security Alert
Yes
No
View Certificate
Opening page https://9.12.4.80/...
Internet
Start
224 Mittwoch, 11 August 2004 10:53

Information you exchange with this site cannot be viewed or changed by others. However, there is a problem with the site's security certificate.

The security certificate was issued by a company you have not chosen to trust. View the certificate to determine whether you want to trust the certifying authority.

The security certificate date is valid.

The name on the security certificate is invalid or does not match the name of the site

Do you want to proceed?

Figure 3-11 Prompt for certificate acceptance

Once the certificate is accepted, you should see the yellow lock sign on the status line as shown in Figure 3-12 on page 59. This indicates a secure SSL connection to the Web server has been established.

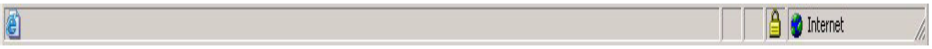


Figure 3-12 Status line for SSL connection

A

z/VM and Linux for zSeries in an On Demand environment

This appendix provides information about IBM Tivoli® Intelligent ThinkDynamic Orchestrator (ITITO) — one of the components of Virtualization Engine™ that allows z/VM and Linux for zSeries virtual machines or LPARs to participate in an On Demand environment.

ITITO as a managing server is not yet supported under Linux for zSeries, so the management server must be installed on a separate platform (such as an IBM @server pSeries® or xSeries® machine). The necessary agents and drivers will enable Linux for zSeries images to be managed by ITITO from another platform.

For technical information about ITITO including vocabulary and definitions, refer to *Developing Workflows and Automation Packages for IBM Tivoli Intelligent ThinkDynamic Orchestrator*, SG24-6057 and *Provisioning On Demand Introducing IBM Tivoli Intelligent ThinkDynamic Orchestrator*, SG24-8888.

IBM Tivoli Intelligent ThinkDynamic Orchestrator

In today's competitive, dynamic and fluctuating business environment, new on demand technologies are required to reduce total cost of ownership, improve the return on IT investments, and support company growth. System and datacenter utilization rates are actually very low, and these under-utilized assets can be tapped to respond to workload demands that can spike with only so much warning as an unexpected news release. Today it is no longer necessary to incur in excess server capacity and cost for just-in-case situations. IBM Tivoli Intelligent Orchestrator helps you deal with these challenges.

It helps you to improve return of IT assets and increase server utilization. It helps boost server-to-administrator ratios by automatically triggering the provisioning, configuration and deployment of a solution into production. This automated process supports servers, operating systems, storage, middleware, applications and network devices. By utilizing existing hardware, software and network devices without rewiring, you can minimize implementation times and achieve a faster return on investment. And new enhancements, such as resource reservation and scheduling support, enhanced reporting, storage provisioning and an expanded library of automation workflows enable this product to help improve your data center operations better than ever.

IBM Tivoli Intelligent Orchestrator provides a powerful solution that can also help improve service levels by constantly monitoring resources and requirements for anticipated peak workloads and then triggering the appropriate response in accordance with business priorities.

In sum, Tivoli Intelligent Orchestrator:

- Protects your existing investments, lowers implementation costs and delivers a rapid return on investment by improving the utilization of existing hardware, software, storage and network devices without rewiring or changing the network architecture
- Automates repetitive, manual tasks performed by system, network, and storage administrators, thereby saving time and money. It boosts administrator productivity by automatically triggering the execution of the steps necessary to provision, configure and deploy a complete solution into productive use
- Increases IT resource utilization tied directly to business results. Orchestration allows companies to manipulate their IT environment in real time - according to defined business policies - to achieve desired business goals. Orchestration does this by sensing the increase or decrease in IT resource demand and automatically taking action to re-allocate resources accordingly throughout the entire system, allowing multiple applications to be

efficiently run according to business priorities on a common, dynamic, intelligently managed IT infrastructure.

- Anticipates, plans and dynamically provides server capacity to meet peak business needs on demand.

IBM Tivoli Intelligent Orchestrator extends the benefits of the IBM Tivoli Provisioning Manager. It intelligently and dynamically issues instructions to Tivoli Provisioning Manager which then uses automation workflows to maintain server availability and meet required service levels in accordance with business priorities. It provides the why, where and when of a complete orchestration solution.

1. The why - By monitoring the applications under its control, IBM Tivoli Intelligent Orchestrator can sense degrading performance and determine why actions need to be taken.
2. The where - Because solutions are monitored closely, IBM Tivoli Intelligent Orchestrator can determine where (which application) a resource is needed and instruct the IBM Tivoli Provisioning Manager to deploy a server, install the necessary software, provision the proper storage and configure the network. This enables an application to maintain service levels while improving IT resource utilization.
3. The when - Utilizing its capacity management capabilities, IBM Tivoli Intelligent Orchestrator can predict when resources will become available or needed. It will start the provisioning process on demand to help match IT resources with an application's growing or decreasing workload.

IBM Tivoli Intelligent Orchestrator provides a powerful solution that can:

- Gather information about the performance of your application clusters and build a workload model that can predict impending resource requirements
- Manage resources across your application clusters to optimize business-aligned service delivery
- Automate the deployment of computing resources for each application environment
- Provide applications with priority access to data center resources based on class of service

IBM Tivoli Intelligent Orchestrator provides the following capabilities:

- Utility Computing enablement

  Enterprises with mainframe systems have been charging business units for MIPS of consumed computing capacity for many years. This cost allocation model was lost in the move to distributed systems where service level management required large, dedicated server deployments that sit largely idle waiting for infrequent peaks in demand (just-in-case). With the dynamic resource provisioning enabled by IBM Tivoli Intelligent Orchestrator, there is opportunity to move back to a pay-per-use model for computing power. This new form of utility computing can take the form of a cost allocation model to charge departments within an enterprise, or a billing model used by outsourcers to bill their enterprise customers.

- Improved server utilization

  IBM Tivoli Intelligent Orchestrator can improve server utilization and lower the total cost of ownership of your existing distributed systems. By establishing shared pools of resources and dynamically provisioning as needed, fewer but more highly utilized servers are required. Applications can continue to run as isolated, dedicated infrastructures. This allows a company to start quickly with little modification to their existing infrastructure or training costs.

- Application server support

  IBM Tivoli Intelligent Orchestrator can orchestrate the provisioning of e-business application environments to provide capacity on demand. By understanding the capacity requirements of each application and comparing it to the committed service level, it determines how much capacity is required. IBM Tivoli Intelligent Orchestrator can then trigger the provisioning of additional servers to add capacity or reclaim servers to reduce capacity.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

IBM Redbooks

For information on ordering these publications, see "How to get IBM Redbooks" on page 66. Note that some of the documents referenced here may be available in softcopy only.

- *Linux on IBM @server zSeries and S/390: VSWITCH and VLAN Features of z/VM 4.4*, REDP-3719
- *Networking Overview for Linux on zSeries*, REDP-3901
- *Linux for zSeries: Fibre Channel Protocol Implementation Guide*, SG24-6344
- *zSeries Crypto Guide Update*, SG24-6870
- *Developing Workflows and Automation Packages for IBM Tivoli Intelligent ThinkDynamic Orchestrator*, SG24-6057
- *Provisioning On Demand Introducing IBM Tivoli Intelligent ThinkDynamic Orchestrator*, SG24-8888

Other publications

These publications are also relevant as further information sources:

- *z/VM: CP Command and Utility Reference*, SC24-6008
- *z/VM: CP Planning and Administration*, SC24-6043
- *z/VM: General Information*, GC24-5991
- *z/VM: TCP/IP Planning and Customization*, SC24-6019
- *z/VM: TCP/IP Messages and Codes*, GC24-6022
- *z/VM: Guide for Automated Installation and Service*, GC24-6099
- *z/VM: Getting Started with Linux on zSeries*, SC24-6096
- *z/VM: Connectivity*, SC24-6080
- *z/OS and z/VM: Hardware Configuration Manager User's Guide*, SC33-7989

- *z/VM: I/O Configuration*, SC24-6100

Online resources

These Web sites and URLs are also relevant as further information sources:

- IBM @server: Linux on System z9 and zSeries

  http://www.ibm.com/servers/eserver/zseries/os/linux/

- IBM: z/VM Operating System

  http://www.vm.ibm.com/

- IBM @server zSeries cryptography for highly secure transactions

  http://www.ibm.com/servers/eserver/zseries/security/cryptography.html

- OSA-Express MCL Enhancements - October 2003

  http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/FLASH10250

How to get IBM Redbooks

You can search for, view, or download Redbooks, Redpapers, Hints and Tips, draft publications and Additional materials, as well as order hardcopy Redbooks or CD-ROMs, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

(0.1"spine)
0.1"<->0.169"
53<->89 pages

Running Linux on IBM System z9 and zSeries under z/VM

Native SCSI support for z/VM

Networking enhancements for Linux guests

Layer 2 network support

This IBM Redbook discusses running Linux under z/VM on IBM System z9 and zSeries platforms. We describe enhancements available in z/VM Version 5.1. The intended audience for this book is IT administrators responsible for installing and configuring z/VM 5.1 systems running Linux guests.

Using Fibre Channel Protocol (FCP), z/VM 5/1 can be installed on and operate from SCSI disks. Configured as emulated Fixed Block Architecture (FBA) disks, z/VM 5.1 can use FCP-attached disks for its system paging, spooling, directory, and minidisks.

z/VM 5.1 adds new functions for Virtual Switches (VSWITCH). For increased network security, guests must have authorization before connecting to a VSWITCH. z/VM 5.1 introduces VSWITCH Layer 2 support. Operating at Layer 2, a VSWITCH delivers and receives network traffic in Ethernet frames. This provides the ability to handle non-IP protocols such as SNA, NetBIOS, and IPX. In addition, Layer 2 support reduces network latency and CPU overhead.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-6311-00 ISBN 0738496405