

z/VM's Control Program (CP): Useful Things to Know

SHARE
Session 9113
Austin, Texas
March 2009

John Franciscovich
z/VM Development
Endicott, NY
francisj@us.ibm.com

(e-mail me for printable b/w PDF file)

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DB2, e-business logo, ESCON, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/390, System z9, VM/ESA, VSE/ESA, WebSphere, xSeries, z/OS, zSeries, z/VM.

The following are trademarks or registered trademarks of other companies

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

LINUX is a registered trademark of Linux Torvalds in the United States and other countries.

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

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

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

Intel is a registered trademark of Intel Corporation.

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

NOTES:

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

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

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

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

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

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

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

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The 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.

Disclaimer

The information contained in this document is not intended to be an assertion of future action by IBM. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adopt these techniques to their own environment do so at their own risk.

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

Any performance data contained in this presentation was determined in a controlled environment and, therefore, the results which may be obtained in other operating environments may vary significantly. Users of this presentation should verify the applicable data for their specific environment.

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

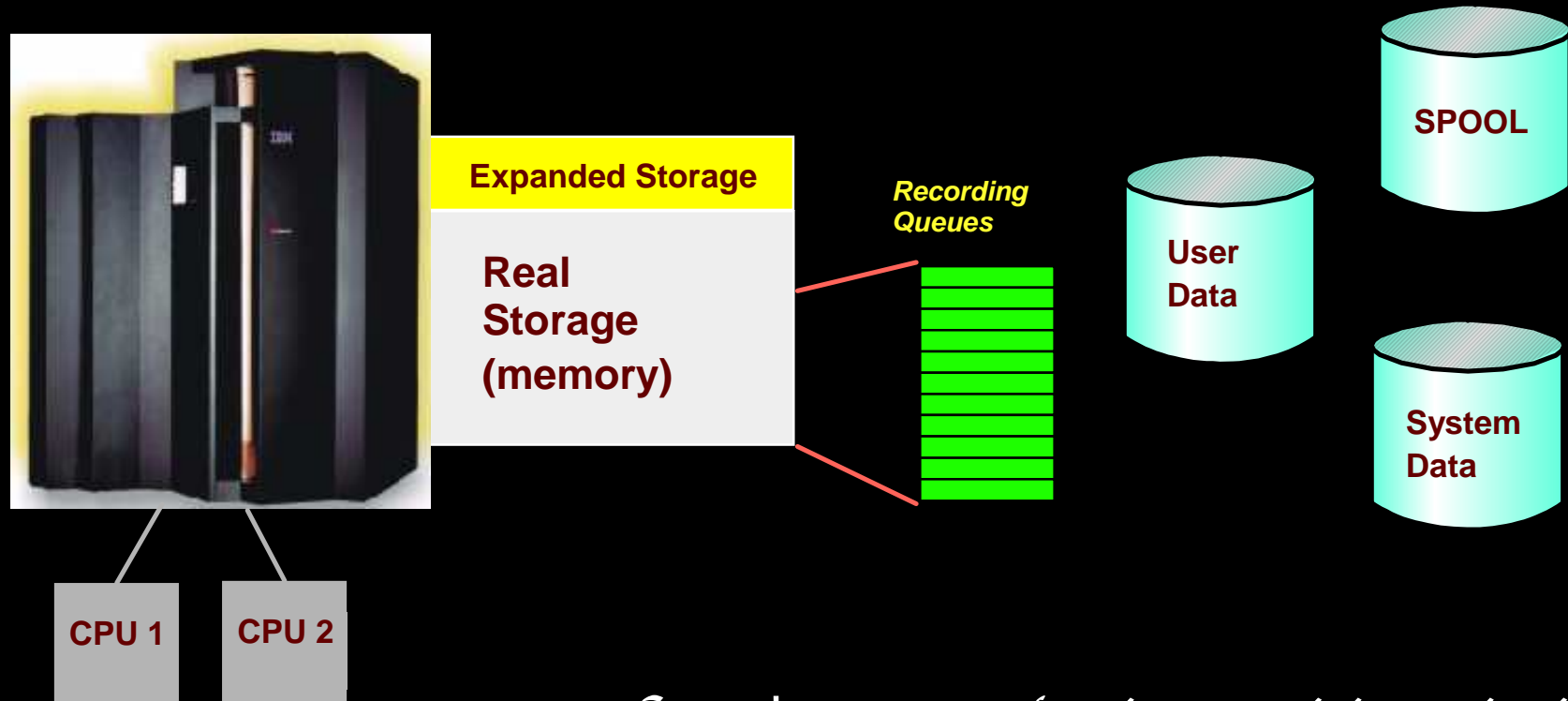
Any feedback that you give IBM regarding this presentation will be treated as non-confidential information. IBM reserves the right to use this information in any form.

Topics

- *Overview of z/VM's CP Facilities and Functions*
- *Starting (IPLing) CP*
 - What you need
 - Saving and Restoring information
- *Defining and Creating Virtual Machines*
- *Virtual Machine Connectivity and Networking*
 - Virtual Machine Communication
 - Virtual Networking
- *Interacting with CP*
- *Collecting Diagnostic Data*

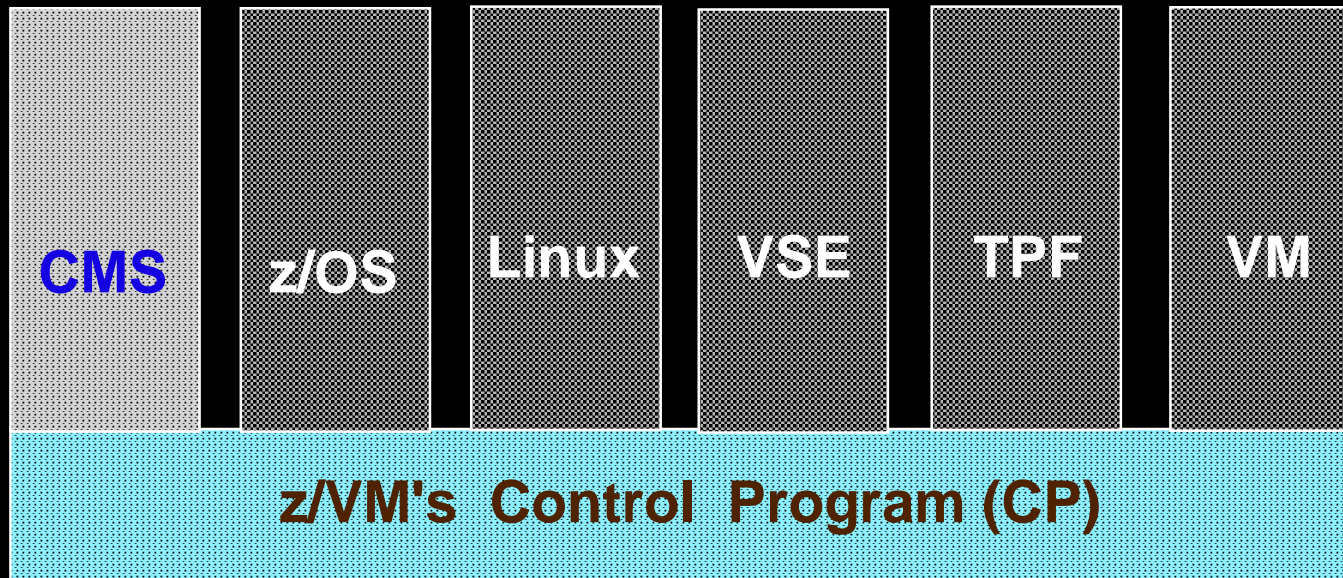
Overview of CP

CP - z/VM's System Control Program



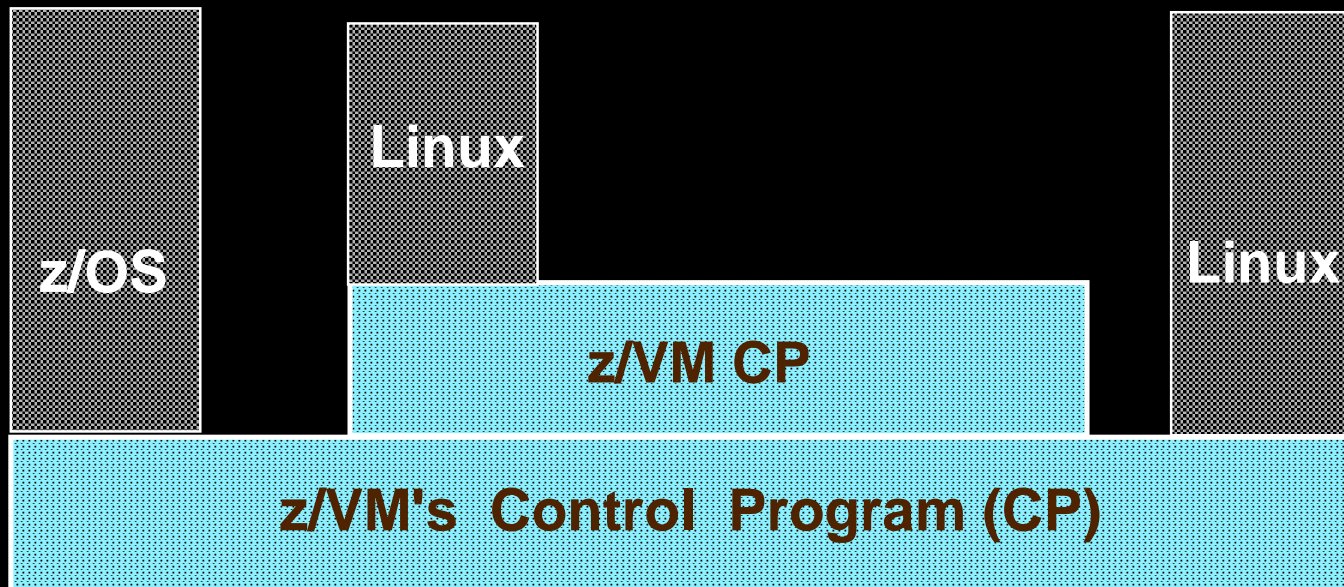
- ▶ Controls resources of environment it is running in
 - LPAR
 - Virtual machine
- ▶ Manages memory and devices
- ▶ Records usage and system event data
- ▶ Provides error recovery facilities

CP - z/VM's System Control Program...



- ***Manages virtual machines***
 - ▶ ESA/390 and z/Architecture
 - ▶ Guest operating systems
 - ▶ Interactive users
 - *CMS is a special single user operating system that is part of z/VM*
- ***Shares real resources among virtual machines***
- ***Supports connectivity among virtual machines***
 - ▶ Virtual networking
 - ▶ Data sharing and exchanging information

CP - z/VM's System Control Program...



Supports multiple layers of virtualization

- ▶ z/VM can run as a guest in a virtual machine
- ▶ Guest z/VM system may host its own guest operating systems

CP Device Support

Real Devices (RDEVs)

- Sensed by CP at IPL time
 - ▶ *Can also be defined to CP in system config file or dynamically*
- Attached or dedicated to a single virtual machine for its exclusive use
- Virtualized and shared among several virtual machines
- Used by CP for system functions

Virtual Devices (VDEVs)

- Appear to virtual machine as a real device
- Defined
 - ▶ *In virtual machine's directory*
 - ▶ *Dynamically after virtual machine is active*
- Either virtualized or simulated
 - ▶ *Virtualized - presents an image of a real device to virtual machines*
 - ▶ *Simulated - defined to virtual machine without an associated real device*

CP Disk Space ("CP Owned")

CP "owns" disk space for system functions

- **PERM**
 - ▶ *Checkpoint and Warmstart areas*
 - ▶ *User minidisks (do not have to be CP Owned)*
 - ▶ *Could contain CP Module*
- **PARM**
 - ▶ *CMS Minidisk containing system configuration files*
 - ▶ *Usually contains CP Module*
- **DRCT**
 - ▶ *User directory (created with DIRECTXA Utility)*
- **PAGE**
 - ▶ *System paging*
- **SPOL**
 - ▶ *Spool files, including DUMP files and System Data files*
- **TDSK**
 - ▶ *Temporary disk space available to users*

CP Disk Space ("CP Owned")...

- *CP disk space is defined in CP_Owned configuration file statement*

```
CP_Owned Slot 1 JF IRES
CP_Owned Slot 2 SPOOL0
CP_Owned Slot 3 MDSP0
CP_Owned Slot 4 RESERVED
```

- ▶ *May be added dynamically to a running system if RESERVED slots are defined*

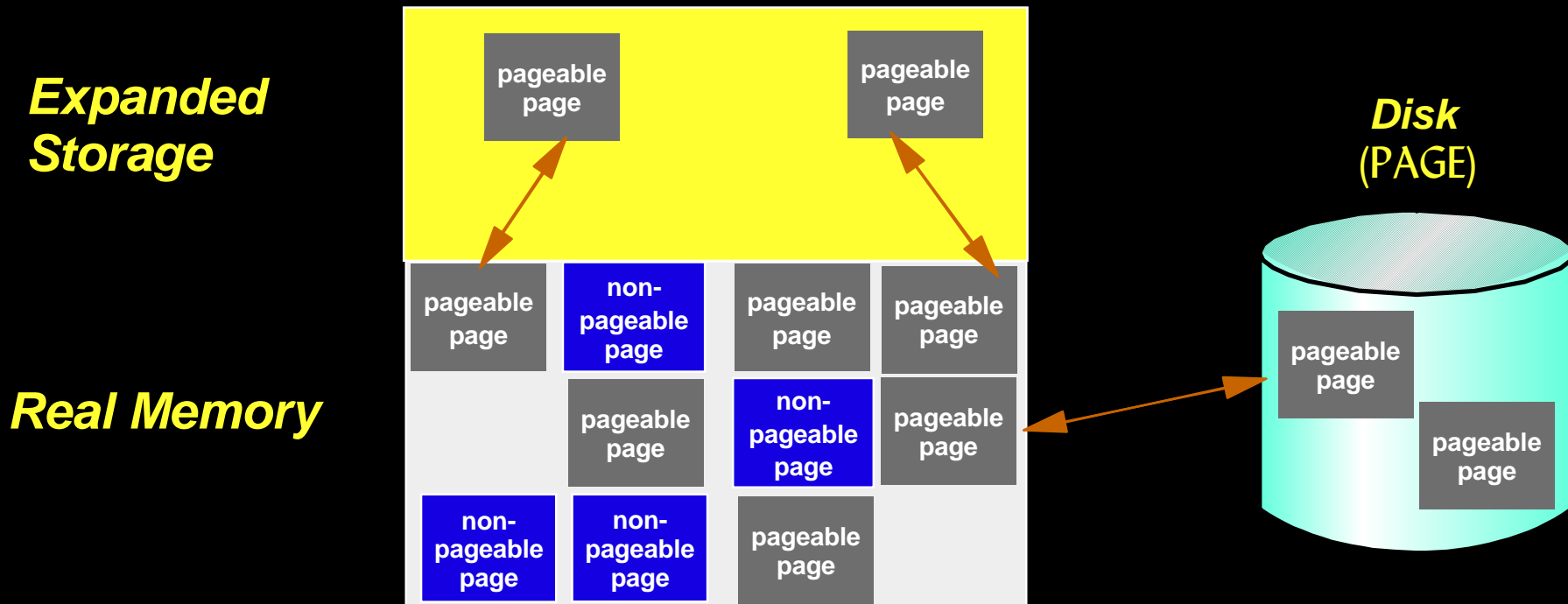
- *CPFMTXA Utility formats and allocates types of CP disk space*
- *QUERY CPOWNED command shows list of CP owned disk volumes*

query cpowned

```
Slot Vol-ID Rdev Type Status
  1 JF IRES 0A40 Own Online and attached
  2 SPOOL0 0780 Own Online and attached
  3 MDSP0 0880 Own Online and attached
  4 ..... ---- ----- Reserved
```

- *QUERY ALLOC command shows various views of CP disk usage*

Managing Real Memory Among Virtual Machines



CP optimizes use of real memory for virtual machines

- ▶ Virtual machine memory is pageable
 - ➔ *Demand paged* - only paged out when necessary
- ▶ Paged to
 - ➔ *Expanded storage*
 - ➔ *Disk (CP-Owned PAGE area)*

CP SPOOLing

Simulates real unit record devices

- Virtual unit record devices defined for each virtual machine
 - ▶ *Reader*
 - ▶ *Printer*
 - ▶ *Punch*
- Reads input (reader) files and makes data available
- Writes data into output (punch or printer) files
- Files may be sent to (or read from) associated real devices

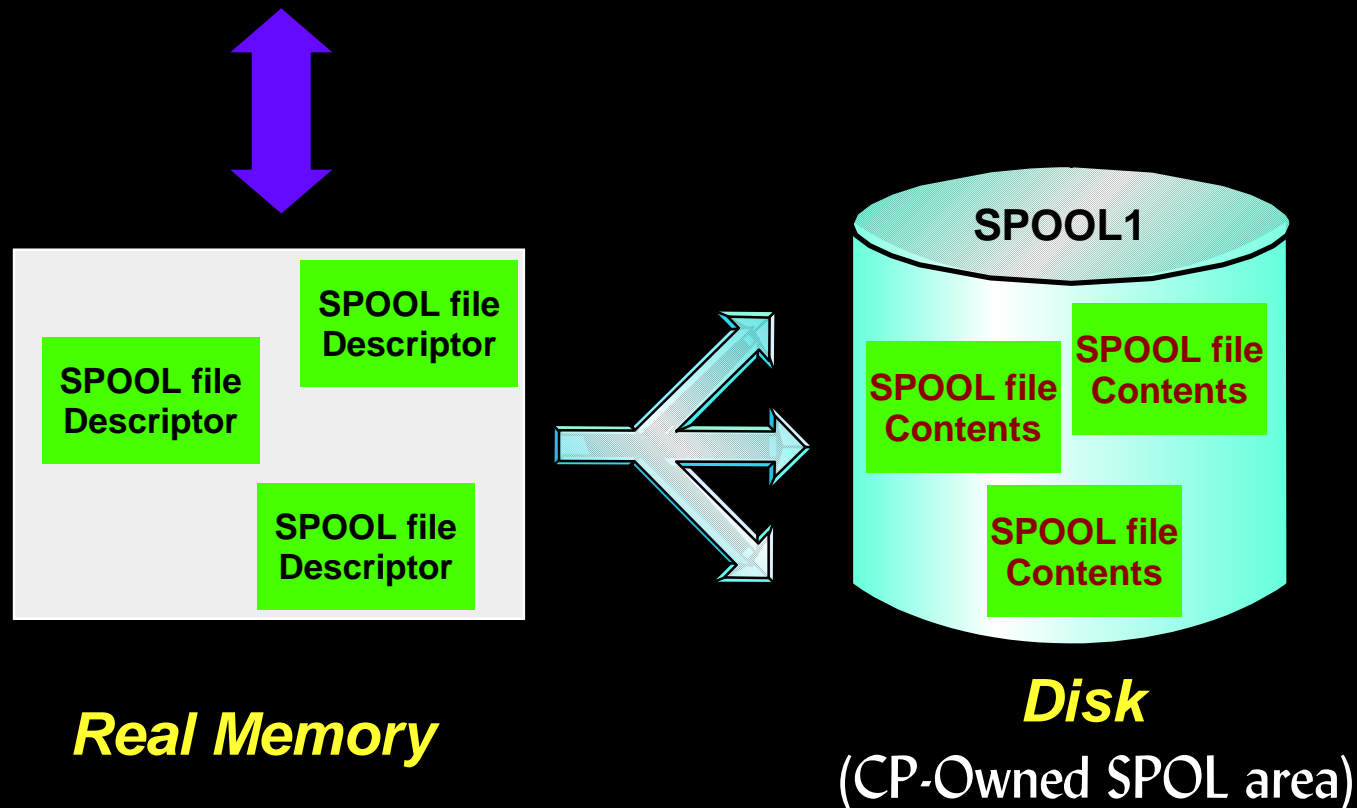
SPOOL files are used for

- Transferring information between virtual machines and systems
- Sending (or receiving) information from associated real devices
- Saving console output
- System and virtual machine dumps
- Specific system functions
- E-mail

CP SPOOLing...

```
q rdr all
```

ORIGINID	FILE	CLASS	RECORDS	CPY	HOLD	DATE	TIME	NAME	TYPE	DIST
OPERATOR	0039	A PUN	00000089	001	NONE	09/02	15:50:06	PROFILE	EXEC	35H:0253
OPERATOR	0037	A RDR	00000006	001	NONE	08/29	15:08:52			OPERATOR
U1	0043	A PUN	00000045	001	NONE	08/03	15:05:53	PROFILE	EXEC	U1



CP SPOOLing - System Data Files

Special SPOOL files used by CP for system functions

- NSS (Named Saved System)
 - ▶ *Named copy of an operating system*
- DCSS (DisContiguous Saved Segment)
 - ▶ *Shared copy of data and/or code*
- IMG (Image Library)
 - ▶ *Definitions such as spacing and character sets used by printers*
- UCR (User Class Restructure)
 - ▶ *Customized privilege class information for commands and diagnose codes*
- NLS (National Language Support)
 - ▶ *Message repositories for translated z/VM messages*
- TRF (System Trace Files)
 - ▶ *System Trace data generated by a virtual machine*
 - ▶ *Created by TRSOURCE and TRSAVE commands*

Named Saved Systems and Saved Segments

NSS (Named Saved System)

- System code saved in a segment
- Can be IPL'd by name (e.g. CMS)
 - ▶ *Single copy on system for N virtual machines instead of N copies*

DCSS (DisContiguous Saved Segment)

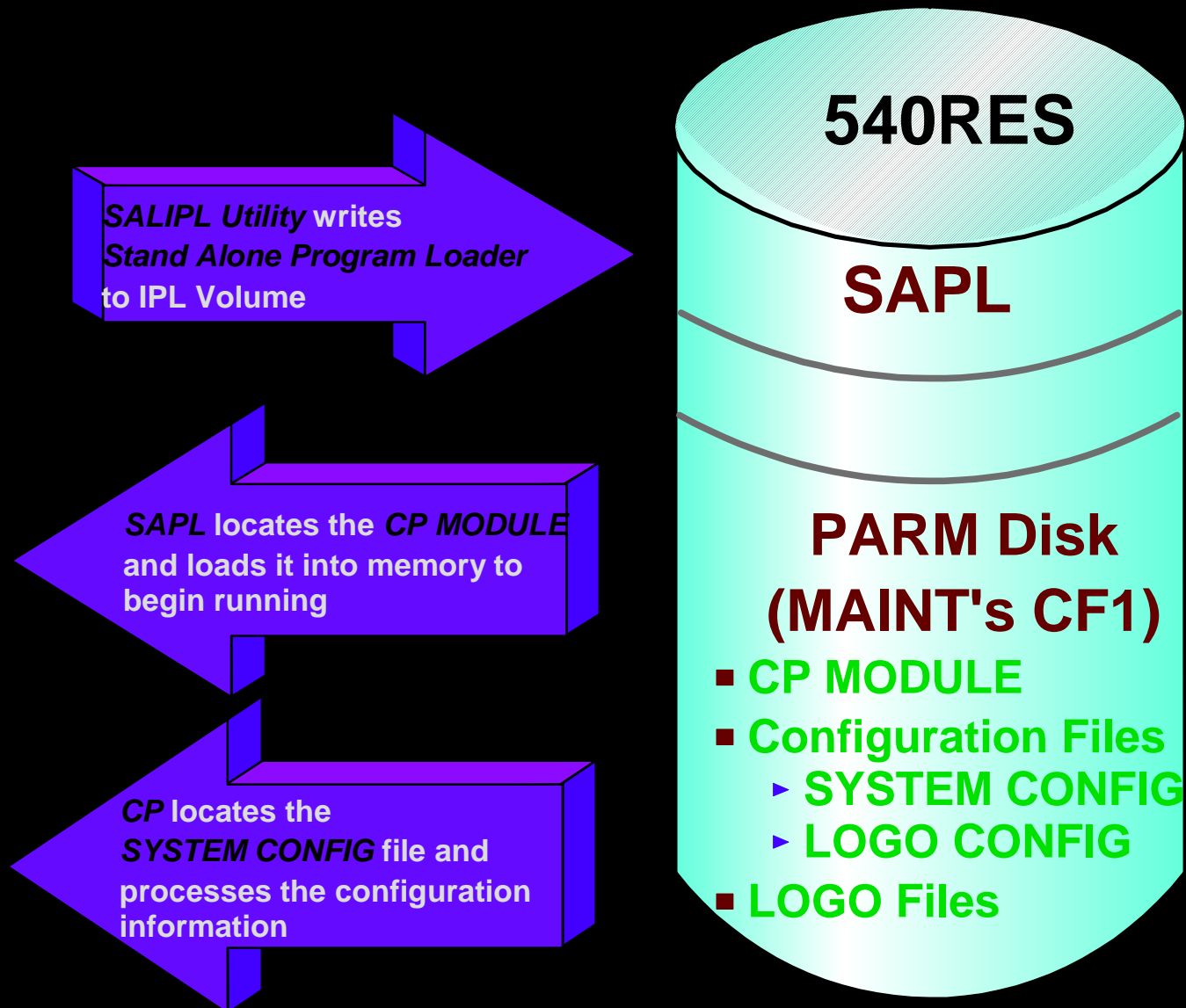
- Used to contain shared
 - ▶ *Data*
 - ▶ *System code*
- Single copy is shared among all virtual machines

Created with DEFSYS and DEFSEG commands

- Skeleton files
- Must be saved with SAVESYS and SAVESEG commands before they can be used by virtual machines

Starting (IPLing) CP

What you Need to IPL CP



Restoring Information during IPL

CP saves system environment and data during SHUTDOWN, including:

- Accounting, EREP, and Symptom records
- Unit record device status
- System log message
- Spool files
- System data files

Type of IPL determines how much saved system information is restored

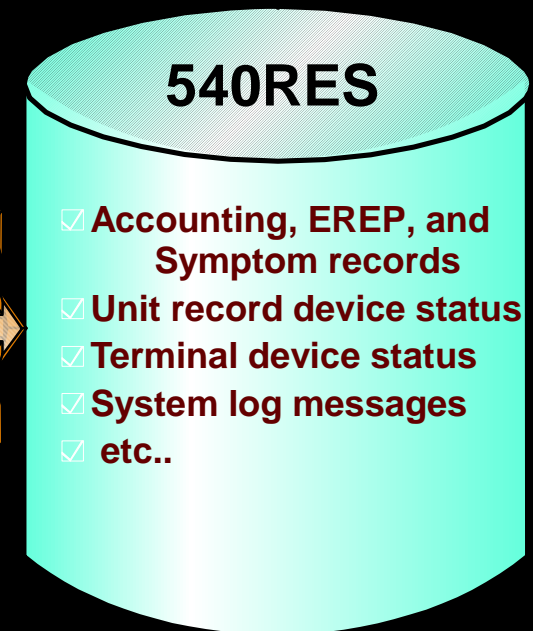
- WARM
 - ▶ *Restores all information saved during SHUTDOWN*
- FORCE
 - ▶ *Restores as much saved information as possible*
- COLD
 - ▶ *Only restores system data files*
- CLEAN
 - ▶ *Does not restore any saved information*

Restoring System Data - Checkpoint Area

System data to be restored during an IPL (WARM or FORCE)

- ▶ Located on a CP -owned volume
- ▶ Not necessarily the IPL volume

System_Residence ,
Checkpoint Valid 540RES From Cylinder 21 For 9 ,
Warmstart Valid 540RES From Cylinder 30 For 9

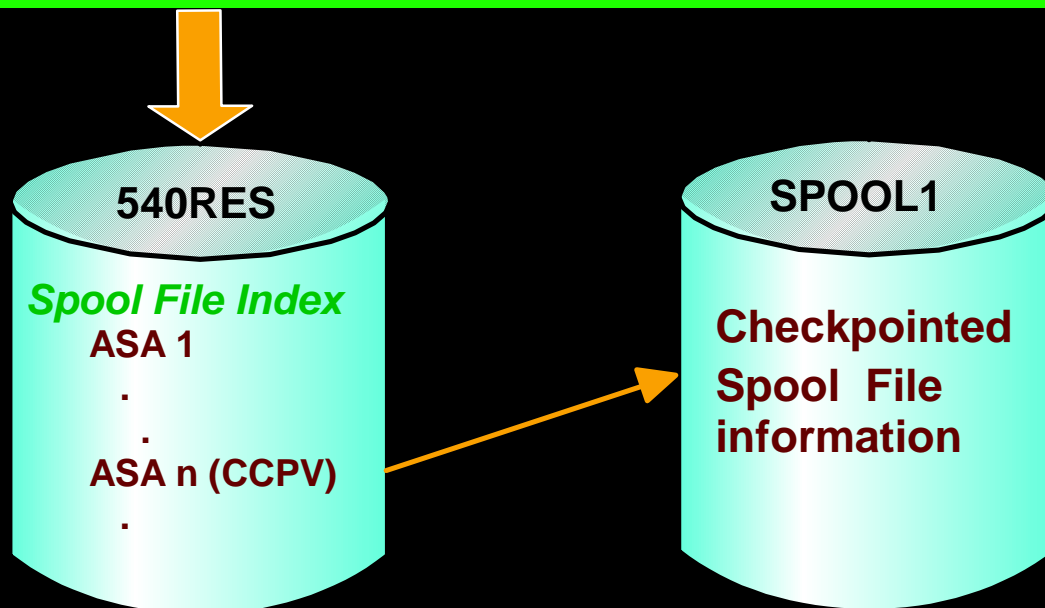


Restoring System Data... Warmstart Area

Spool Files to be restored during a system restart

- ▶ One entry per file : 4-byte Disk (Auxiliary Storage) Address
- ▶ Updated whenever a spool file is created or deleted

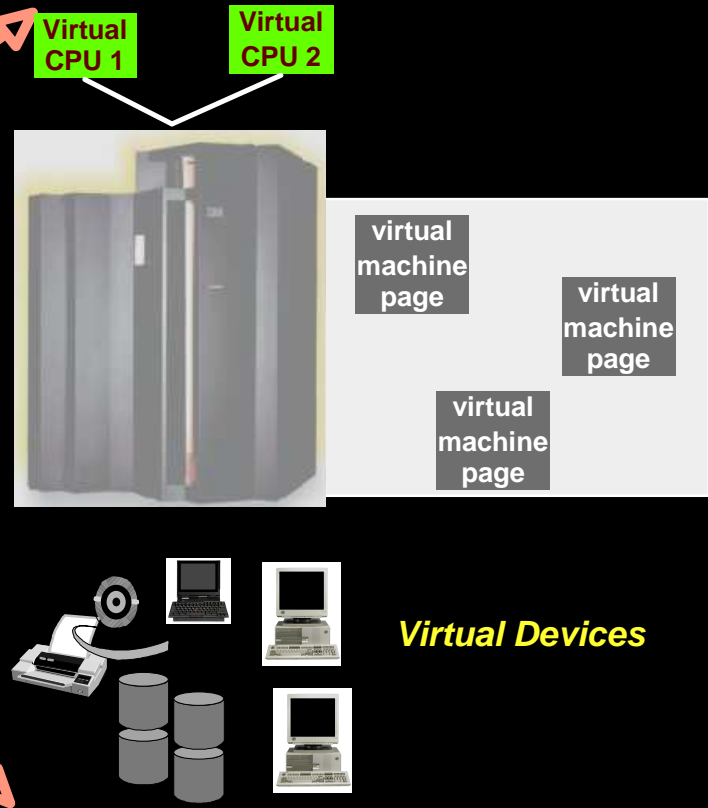
System_Residence ,
Checkpoint Valid 540RES From Cylinder 21 For 9 ,
Warmstart Valid 540RES From Cylinder 30 For 9



Virtual Machines

Defining a Virtual Machine

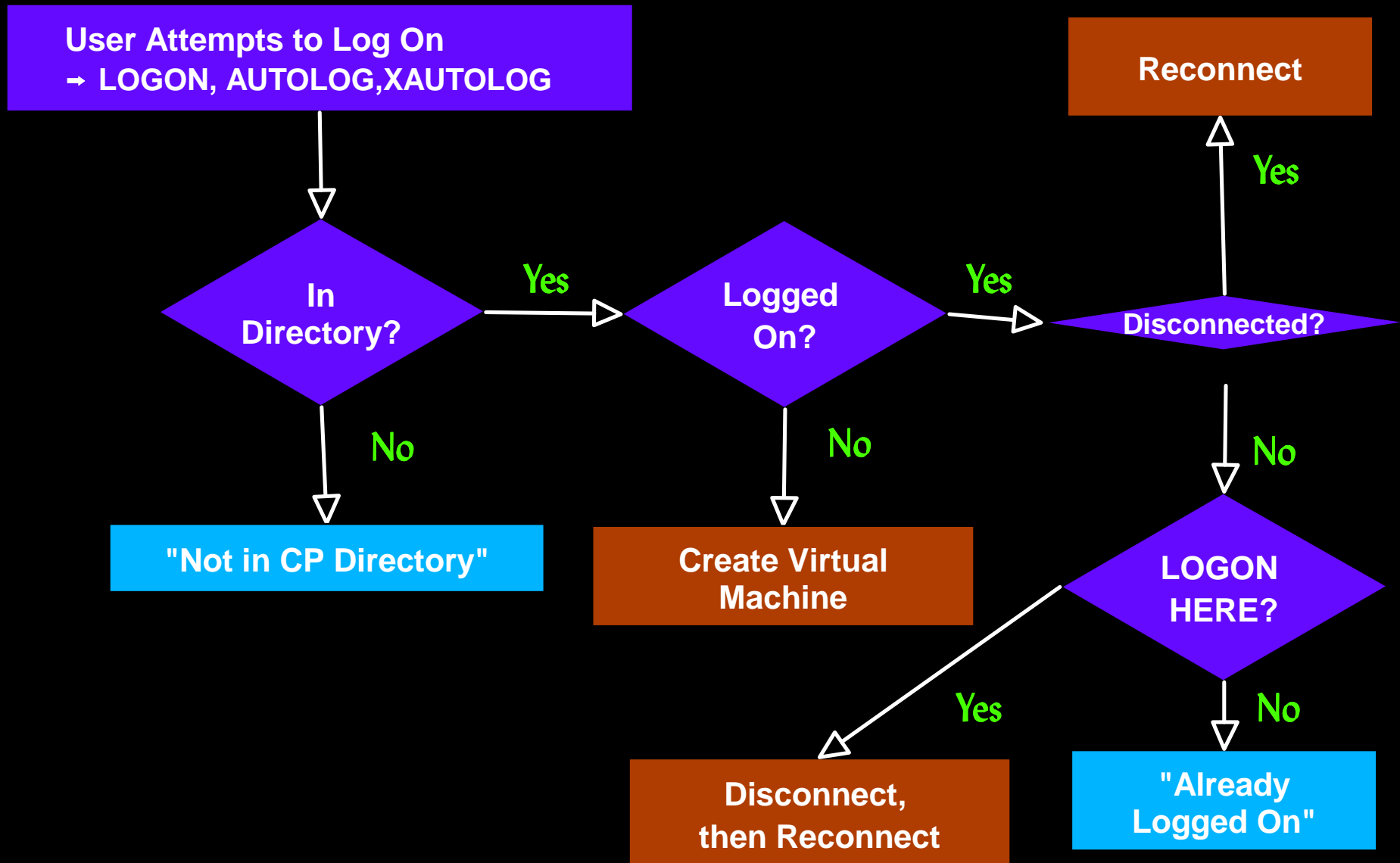
```
USER U1      U1PW 32M 32G G
IPL 190     PARM AUTOCR
MACHINE ESA 2
COMMAND ATTACH 555 * 555
CONSOLE 009 3215 T MAINT
SPOOL 00C 2540 READER A
SPOOL 00D 2540 PUNCH A
SPOOL 00E 1403 A
MDISK 191 3390 000 009 JAF191 MR
MDISK 193 3390 000 017 JAF193 RR
LINK MAINT 190 190 RR
LINK MAINT 19E 19E RR
```



Created when a user logs on

- ▶ Real hardware and features are virtualized
 - Processors
 - Devices
 - Memory
- ▶ Aggregate of virtual objects and memory may be greater than available real resources

Logging on to z/VM (creating a virtual machine)



*Virtual Machine
Connectivity
and
Networking*

Communication between Virtual Machines

IUCV (Inter-User Communication Vehicle)

- provides an efficient data transfer protocol unique to the VM platform

Virtual CTCA

- simulates existence of real Channel-to-Channel devices for each Virtual Machine

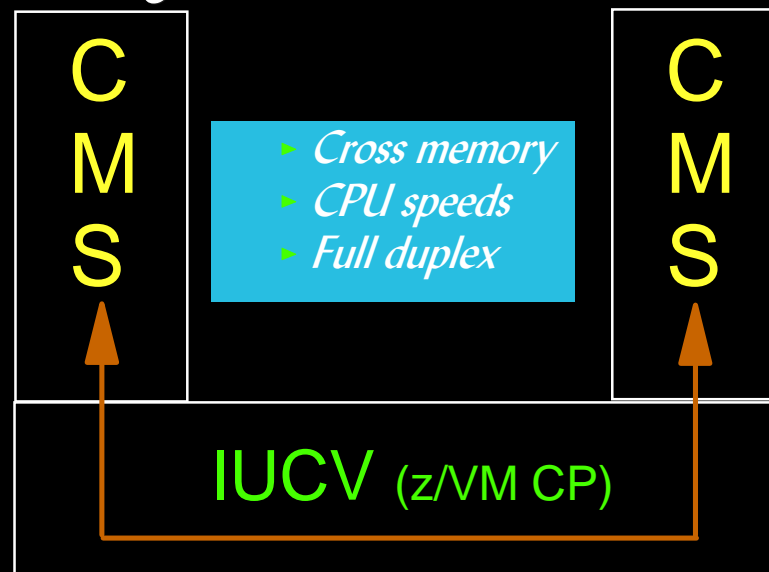
Virtual NIC

- simulates existence of real Network Interface Cards for each Virtual Machine

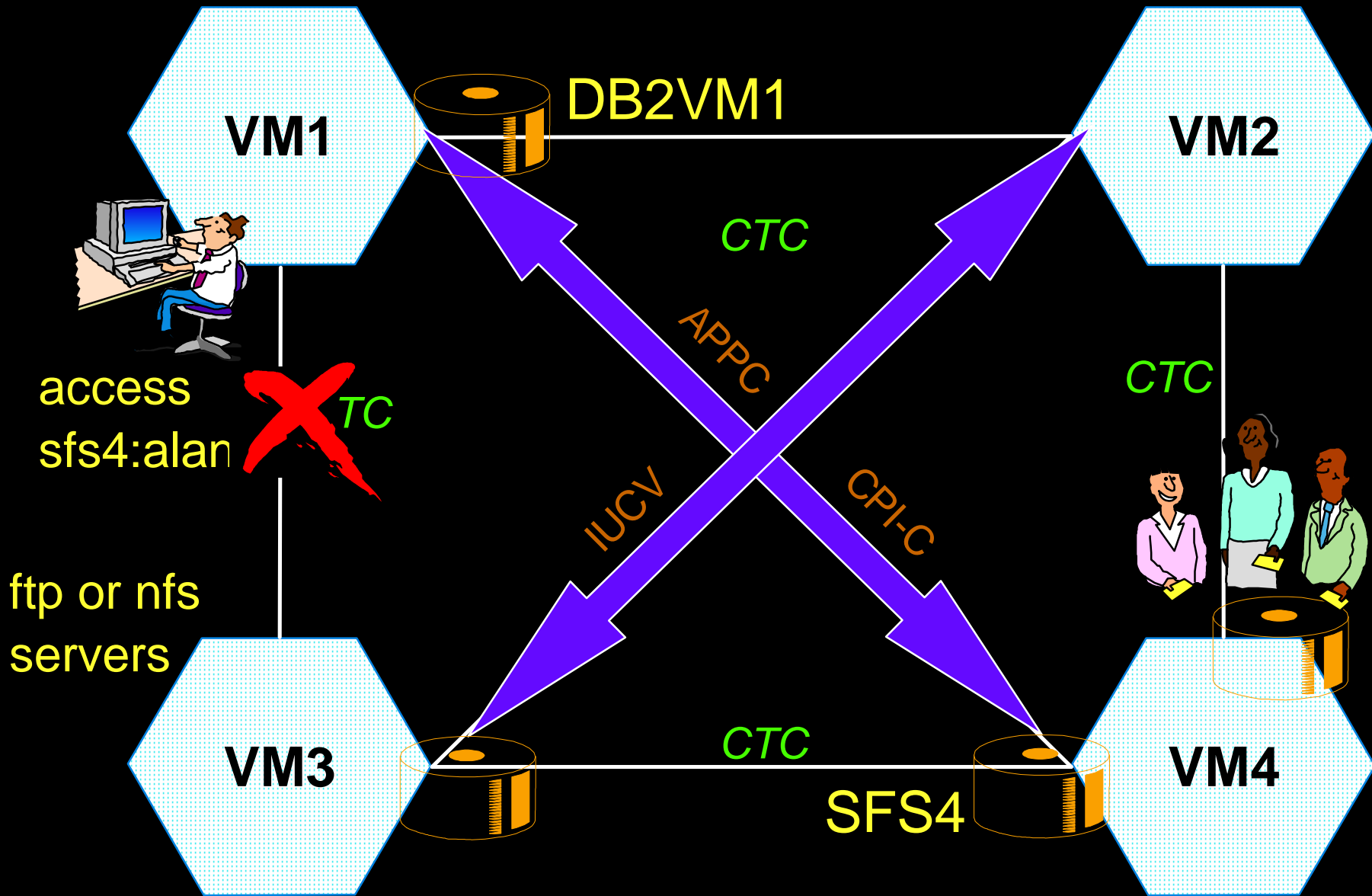
IUCV Communication

Inter-User Communication Vehicle (IUCV)

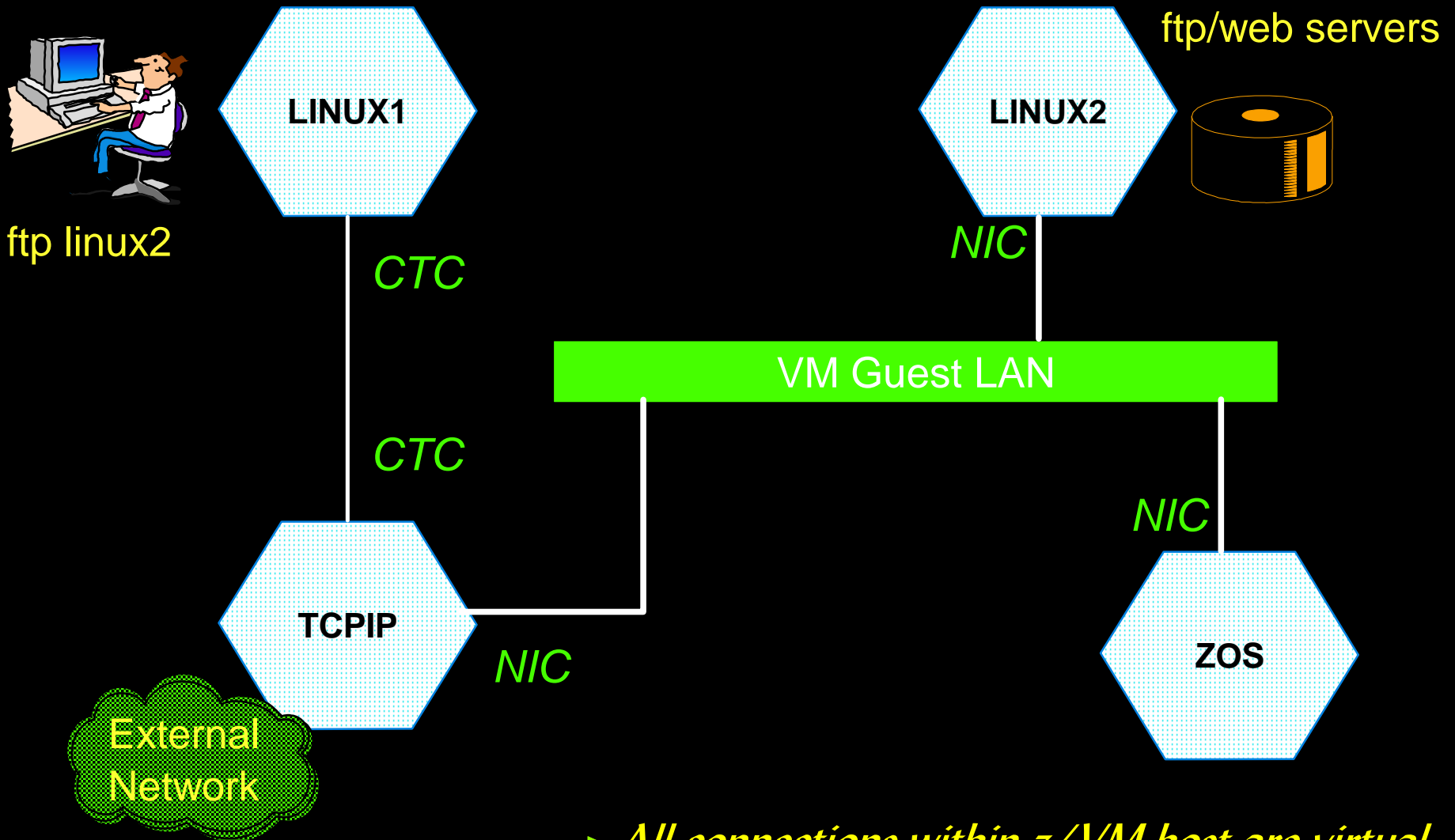
- Allows communication between an application and other virtual machines or CP system services
- Simultaneous communication over multiple connections allowed for each virtual machine
- Transparent communication between virtual machines on different systems via ISFC (*Inter-System Facility for Communications*)
- Point-to-Point networking between Linux and z/VM TCP/IP



Communications Services (CS) Collection - ISFC

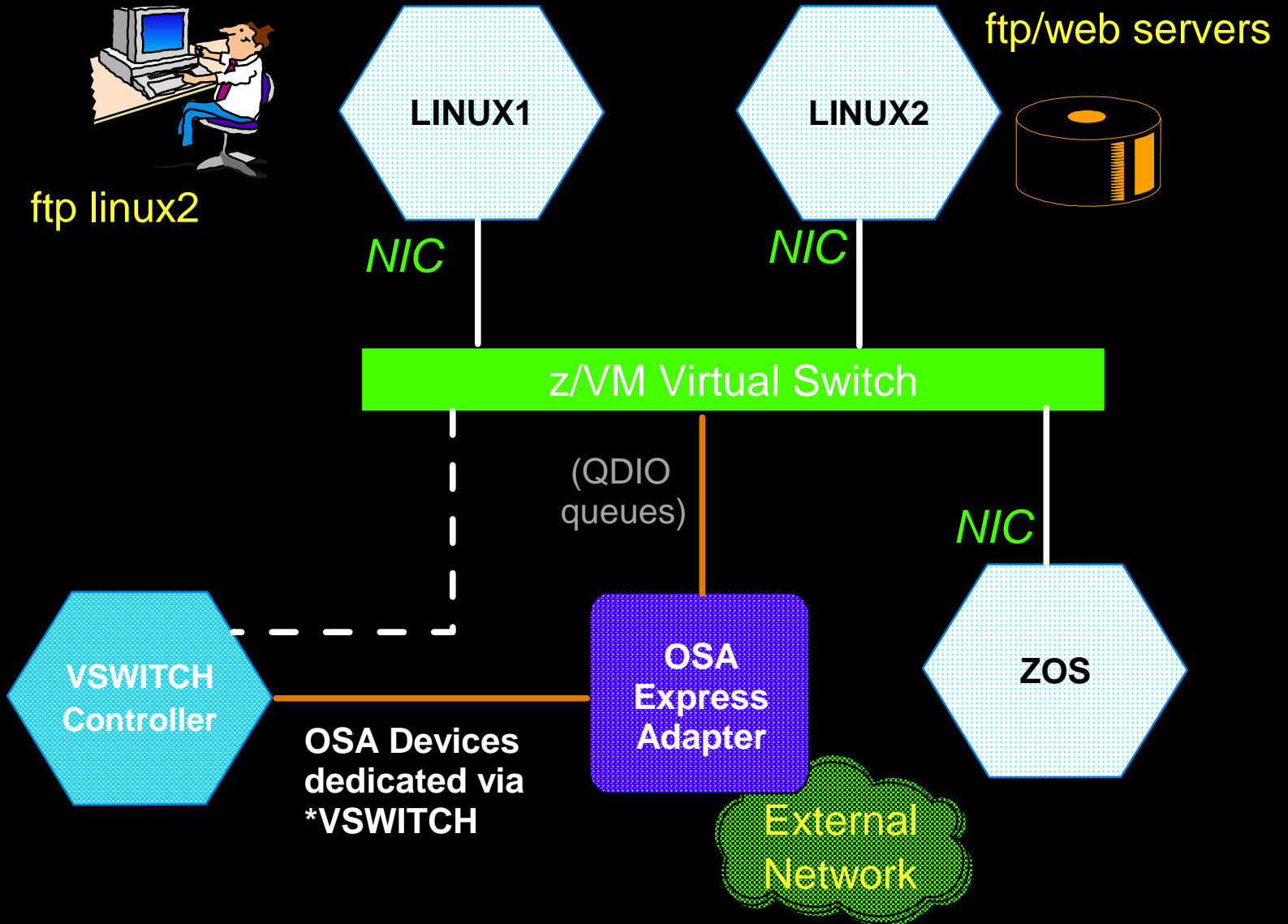


A Virtual Network (Guest LAN)



- ▶ *All connections within z/VM host are virtual*
 - no physical connections are necessary

A Virtual Network (z/VM Virtual Switch)

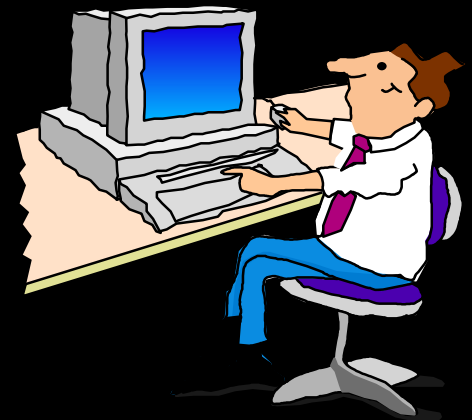


Interacting with CP

CP Commands

Used for a variety of purposes, including:

- System Operator functions
- System status
- DEFINE/SET/QUERY
 - ▶ *system and virtual machine characteristics*
 - ▶ *real/virtual device settings*
 - ▶ *system and user data*
- Assigning/releasing system resources
- Moving data and files between users
- Communicating between virtual machines



CP Commands...

- *COMMAND* directory statement
- *FOR* command
 - ▶ *allows a user to issue CP commands on behalf of another user*
 - ◆ issuer must have SECUSER authority or Class C privileges
 - ◆ target must have appropriate authorization for subject command
 - ▶ *command responses sent to issuer*
 - ◆ no indication to target user
 - ▶ *Example: from user OPERATOR:*

```
FOR u I CMD q v stor
```

```
Ready;
```

```
11:59:21 UI      : STORAGE = 32M
```

```
11:59:21 UI      : HCPFOR069I Command Complete. CP return code = 0000.
```

Privilege Classes

Each user (virtual machine) has one or more privilege classes

- most are only Class G

Identifies:

- which CP commands the user is allowed to issue
- scope of influence of commands

Privilege classes may be modified

- user
- command

Class	Type of User and Function
A	<i>System Operator</i> : responsible for availability of system and resources
B	<i>System Resource Operator</i> : controls real resources of system, except for those controlled by the system operator and spooling operator
C	<i>System Programmer</i> : Changes system-wide parameters
D	<i>Spooling Operator</i> : Controls spool files and system's real reader, printer, and punch equipment
E	<i>System Analyst</i> : Examines and saves system operation data
F	<i>Service Representative</i> : Reserved for IBM use
G	<i>General User</i> : Controls functions associated with a particular virtual machine
Any	Commands available to any user regardless of the user's privilege class

CP Programming Interfaces

Provide application programs with access to

- CP Services
- Data created by CP to be processed by applications
- Certain CP data areas

Types of programming interfaces

- Diagnose Codes
- CP System Services
- IUCV and APPC/VM macros
- Address space macros (ESA/XC virtual machines)



Customizing CP

CP Exit support

- Allows non-disruptive additions and deletions of customized CP routines
 - ▶ *CP Commands*
 - ▶ *Diagnose Codes*
 - ▶ *Message Repositories*
 - ▶ *Exit Routines (user modifications to CP)*
- No need to shutdown and IPL to apply user code
 - ▶ *Modifications applied with commands and configuration file statements*
- Minimizes rework to user code due to IBM source code changes

*Collecting
Diagnostic Data*

Diagnostic Data

Several types of data created by CP can help diagnose problems

- Console messages and logs
- Dumps
 - ▶ *System (CP)*
 - ▶ *Virtual Machine*
- TRACE Data
- Performance Data
 - ▶ *Reports from Performance Tools*
 - ▶ *INDICATE commands*
 - ▶ *MONITOR data*

Diagnostic Data . . .

Commands may be used to collect additional information

- QUERY
- LOCATE
 - ▶ *Host Logical and/or Host Real addresses, depending on resource being located*
- DISPLAY
 - ▶ *Specify Host Logical or Host Real addresses to be displayed*
- etc...

Console Messages and Logs

Most applications and system functions write messages to the virtual machine's console

- System messages are displayed on the operator's console

Console information can be easily saved for review

- SPOOL CONSOLE START command
 - ▶ *Begin collecting console data*
 - ▶ *Direct console file to desired virtual machine*
- SPOOL CONSOLE STOP/CLOSE command
 - ▶ *Stop collecting console data*
 - ▶ *Close the file so it may be saved and reviewed*
- RECEIVE file to disk or PEEK it in RDR
 - ▶ *Use "(FOR *" if PEEKing file*

CP Dumps

Written to SPOOL or tape

- Determined by SET DUMP command
 - ▶ *SET DUMP DASD for SPOOL*

Hard Abend

- Contains all of CP-owned memory

Soft Abend

- Does not cause system termination
- Contains
 - ▶ *VMDBK of the active virtual machine at time of abend*
 - ▶ *CP Trace Table for processor where error occurred*

SNAPDUMP

- Contains same information as Hard Abend dump
- Does not terminate the system

Other information common to all types of CP dumps

More Dumps

VMDUMP (Virtual Machine Dump)

- Created with VMDUMP command
 - ▶ *Unformatted dump*
 - ◆ 4K pages of virtual machine's memory
 - ▶ *Placed in virtual reader*
 - ▶ *DUMpload command used to load into CMS file*

Stand-Alone Dump

- Same format asabend dump
 - ▶ *Writes dump of all of main memory*
- Created when stand-alone dump utility is IPLed
 - ▶ *Utility created by HCPSADMP EXEC*
 - ◆ placed on volume that can be IPLed to start Stand-Alone Dump
- Always written to tape



Processing CP Dumps

CP Dumps are generally sent to OPERATNS reader (RDR)

- DUMpload command processes dumps from RDR (or tape) to disk

The VM Dump Tool is used to analyze dumps

- CP Abend, SNAPDUMP, or Stand-Alone dumps
- Issue VMDUMPTL command

```
z/VM Version 5 Release 2.0, service level 0000 (CP 64-BIT)
Generated at 09/12/06 22:07:57.000000, IPLd at 09/12/06 23:07:29.766710
Date 09/12/06 Time 23:10:38.459806
```

```
CPUID = FF319B9E 20948000
```

```
CPU address is 0000      Prefix register is 00038000      (failing)
03D20660 23:10:34 Call from HCPGRF+1892 to HCPIOSRQ sav 01146C00
                    parm 01151000
```

```
Summary of CP exits
```

```
  0 Pre-defined exits found
  0 Dynamic exits found
  0 Diagnose exits found
```

```
SVC002 (hard abend) A restart interrupt occurred. For a first level
system, a restart interrupt occurs when the primary system
operator selects the restart function on the hardware console.
For a second level system, a restart interrupt occurs when the
"SYSTEM RESTART" command is entered on the first level console.
```

Tracing

General CP Tracing

- CP builds trace tables for each CPU during initialization
- All occurrences of traceable system events are recorded

VMDUMPTL Display of CP Trace Table

```
>>> trace merge for 100 one
03D20660 CPU 0000 Call from HCPGRF+1892 to HCPIOSRQ sav 01146C00
03D20640 CPU 0000 Obtain 38 dw (GSD) at 010C7D28 by HCPGRF+112E for OPERA
03D20620 CPU 0000 Obtain 16 dw (RCW) at 0120D6E8 by HCPGRS+448 for OPERAT
03D20600 CPU 0000 Unstack IORBK at 01207008 user OPERATOR VMDBK 01151000
03D205E0 CPU 0000 Exit to dispatcher from HCPIOL+12E userid SYSTEM
03D205C0 CPU 0000 Release 59 dw (IOR) at 01207E08 by HCPIFI+A64 for SYSTE
03D205A0 CPU 0000 Return LR<SR to HCPIFI+912 fr HCPGER+2B6 sav 01146C00
03D20580 CPU 0000 Return to HCPGER+2A0 fr HCPERP+1906 sav 01146000
03D20560 CPU 0000 Return to HCPERP+18CE fr HCPGRE+FE sav 01142C00
03D20540 CPU 0000 Return to HCPGRE+EE fr HCPINV+12C sav 01146400
03D20520 CPU 0000 Return to HCPINV+2E0 fr HCPUSL+B8 sav 0116D600
03D20500 CPU 0000 Add user OPERATOR VMDELIST 01 VMDWSSPR/HOTWS 00000736
03D204C0 CPU 0000 /Monitor event at HCPSCI+94
03D20480 CPU 0000 /Monitor event at HCPSCH+42A
03D20460 CPU 0000 Stack IORBK at 01207008 user OPERATOR by HCPUSL+AA
03D20440 CPU 0000 Obtain 59 dw (IOR) at 01207008 by HCPUSL+6E for SYSTEM
03D20420 CPU 0000 Call from HCPINV+2E0 to HCPUSLAT sav 0116D600
03D20400 CPU 0000 Return to HCPINV+216 fr HCPRBK+A8 sav 0116D600
03D203E0 CPU 0000 Call from HCPINV+216 to HCPRBKDA sav 0116D600
03D203C0 CPU 0000 Release 2 dw (???) at 011705D0 by HCPINV+120 for SYSTEM
03D203A0 CPU 0000 Release 2 dw (???) at 011705F8 by HCPINV+1F8 for SYSTEM
03D20380 CPU 0000 Obtain 2 dw (???) at 011705F8 by HCPINV+1B8 for SYSTEM
.....
```

Tracing...

TRACE Command

- Monitors events in virtual machines
 - ▶ Execution of instructions
 - ▶ Memory Alteration
 - ▶ Register Alteration
 - ▶ I/O Activity

Data, I/O, and Guest Tracing

- TRSOURCE and TRSAVE commands
- Data written to system Trace File (TRF)

```
CP TRSOURCE ID TRAP1 SET TRSAMPLE TYPE DATA LOC HCPSPX + C42 41200074
```

```
CP TRSOURCE ID TRAP1 SET TRSAMPLE TYPE DATA DL G0:15=REGS
```

```
CP TRSOURCE ID TRAP1 SET TRSAMPLE TYPE DATA DL G5.D0=SPFBK
```

```
CP TRSAVE FOR ID TRAP1 DASD TO * SIZE 256 KEEP 4
```

```
CP TRSOURCE ENABLE SET TRSAMPLE
```

```
CP TRSOURCE DISABLE SET TRSAMPLE
```

```
QUERY TRF ALL
```

```
TRACERED x x x x CMS TRSDATA OUTPUT A
```

```
where x = spoolid(s) of TRF file(s)
```

Summary

Summary

VM's Control Program (CP):

- Efficiently manages the environment it is running in
 - *LPAR*
 - *Virtual Machine*
- Manages processors, memory, and devices among virtual machines
 - *Efficiently shares available resources to meet virtual machine requirements*
 - *Virtualizes resources for use by virtual machines*
- Preserves and restores data across system IPLs
- Provides virtual networking and connectivity capability
- Records Diagnostic Information
 - *Several types of data*
 - *Many ways to collect it*

Additional References

z/VM Library

- <http://www.vm.ibm.com/library/>

IBMVM Listserver

- <http://listserv.uark.edu/scripts/wa.exe?AO=ibmvm>