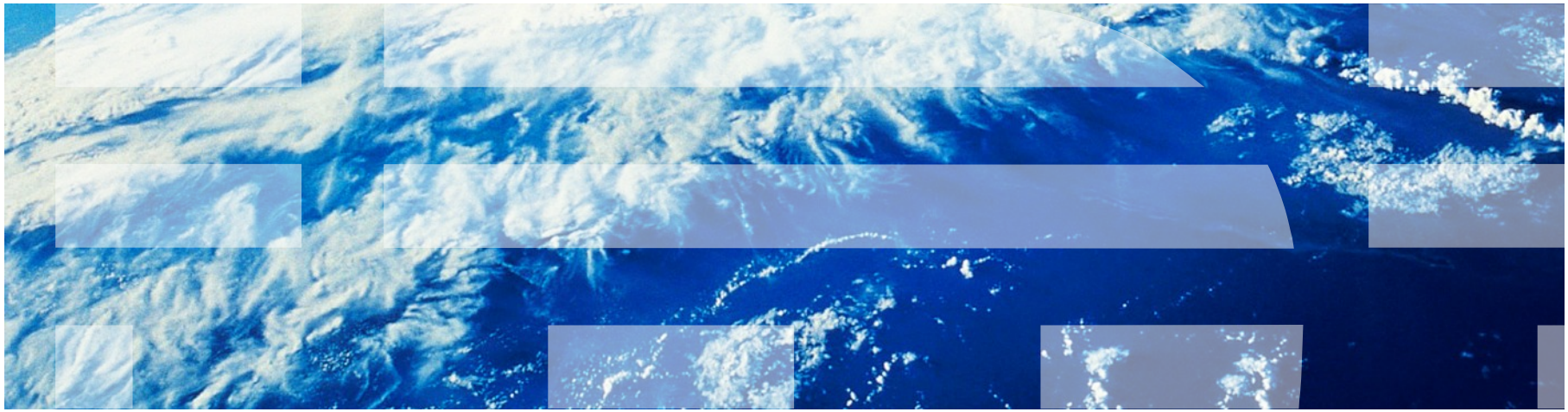
John Franciscovich
francisj@us.ibm.com
June 22, 2010

The z/VM Control Program – Useful Things to Know

2010 Blooming Basics for z/VM & Linux on System z



Trademarks

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

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml:

*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.
Java 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, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* 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.

Disclaimer

The information contained in this document has not been submitted to any formal IBM test and is distributed on an "AS IS" basis without any warranty either express or implied. 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 adapt these techniques to their own environments do so at their own risk.

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

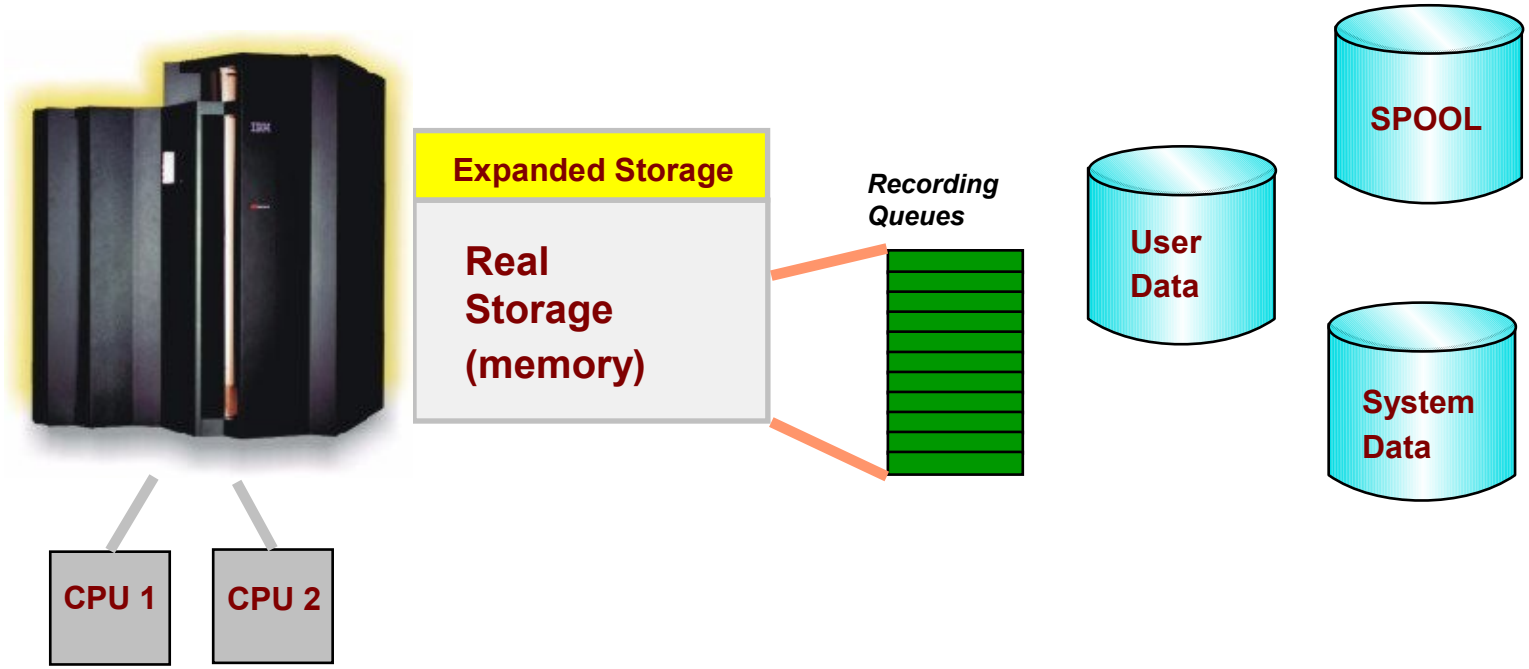
It is possible that this material may contain reference to, or information about, IBM products (machines and programs), programming, or services that are not announced in your country. Such references or information must not be construed to mean that IBM intends to announce such IBM products, programming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

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

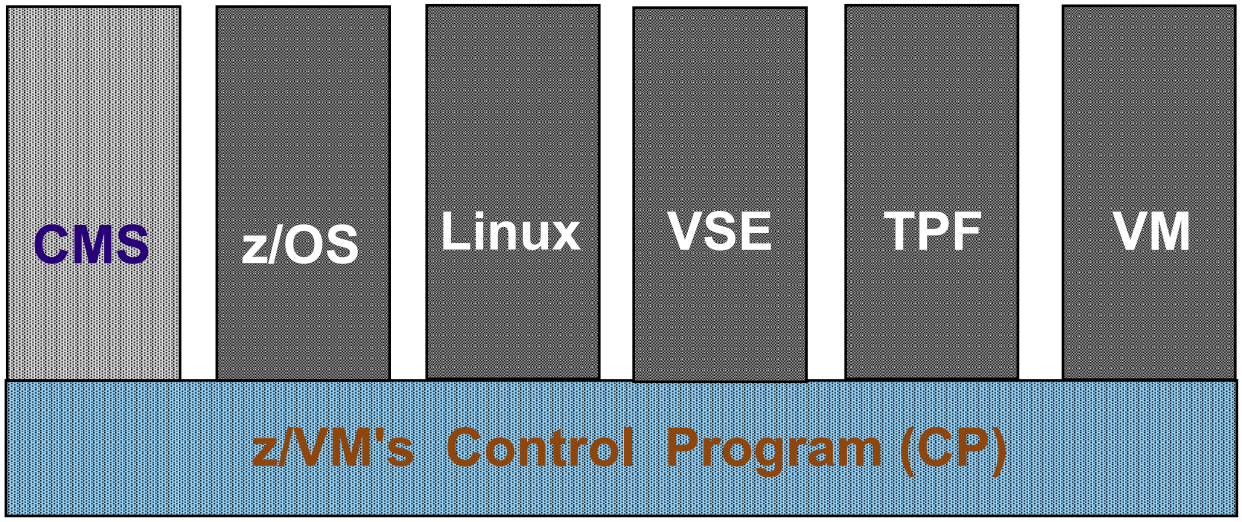
Overview of CP

CP – z/VM's System Control Program



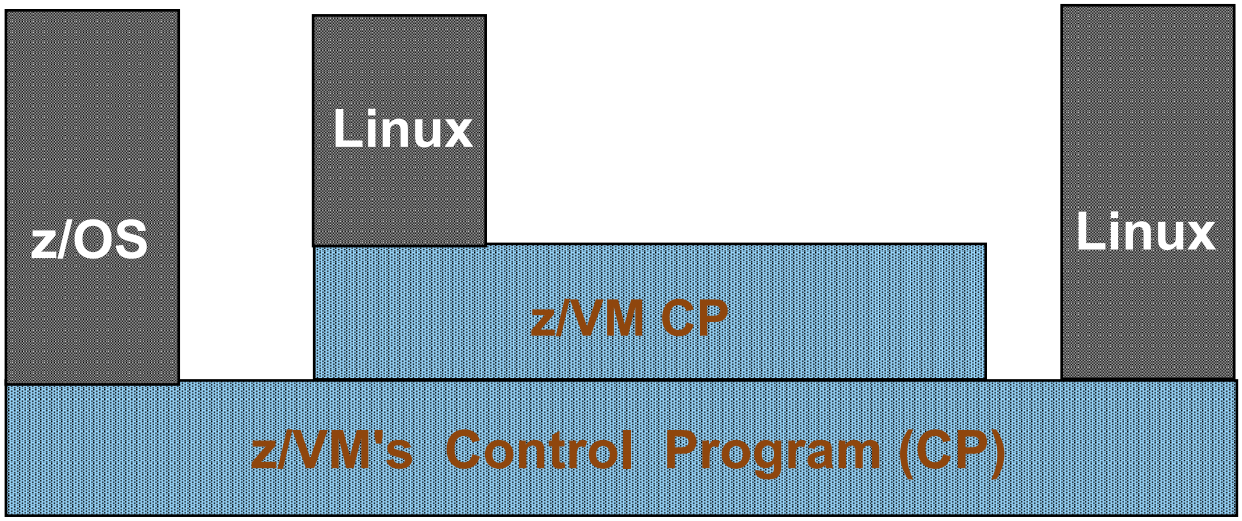
- Controls resources of environment it is running in
  - *LPAR*
  - *Virtual Machine*
- Manages memory and devices
- Records usage and system event data
- Provides error recovery facilities

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
- Provides 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 the CP_Owned configuration file statement

```
CP_Owned  Slot  1  JF1RES
CP_Owned  Slot  2  SPOOL0
CP_Owned  Slot  3  MDSP0
CP_Owned  Slot  4  RESERVED
```

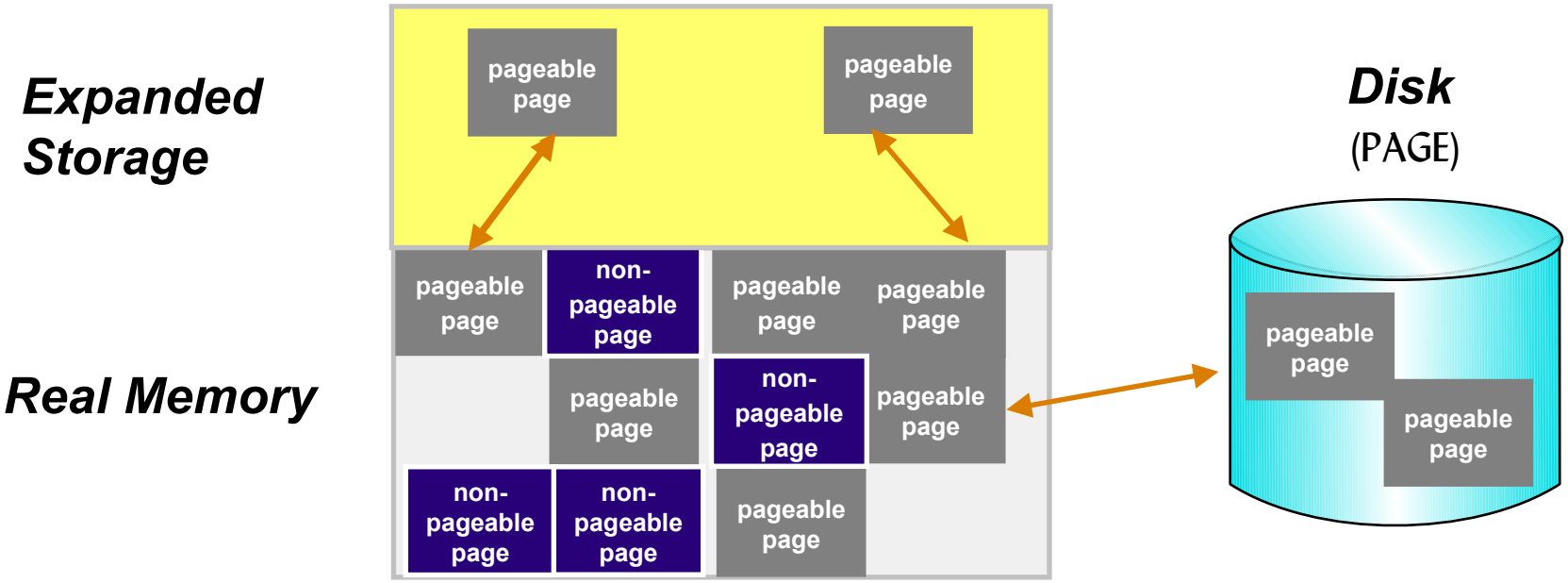
  - *May be added dynamically to a running system*

- 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    JF1RES      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

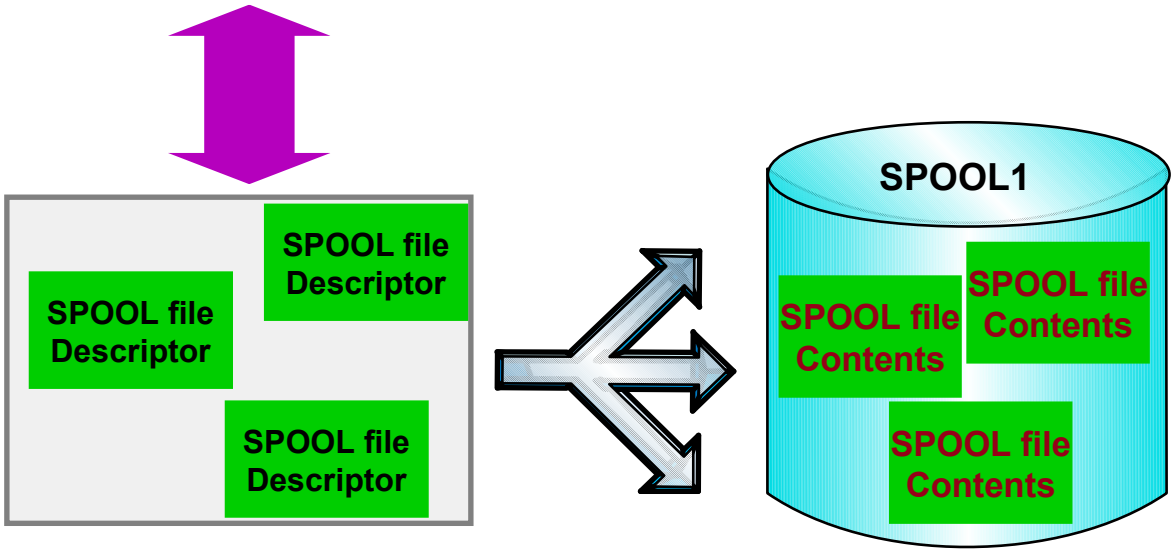
- Virtual unit record devices defined for each virtual machine
  - *Reader*
  - *Printer*
  - *Punch*
- Reads input (reader) files and makes data available
- Writes data into output (printer or punch) files
- Files may be sent to (or read from) associated real devices

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
```



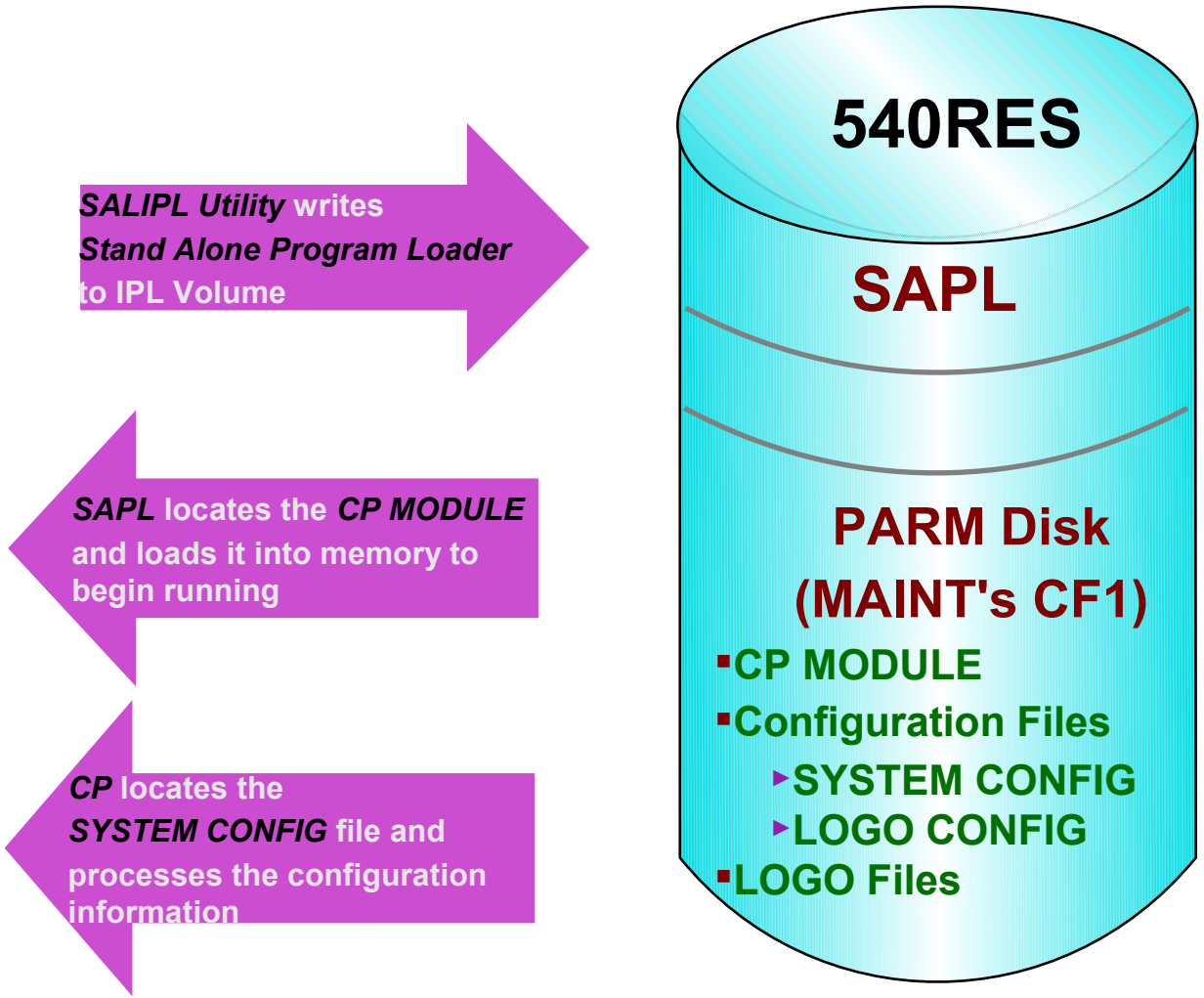
CP SPOOLing – System Data Files

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*
- **NLS (National Language Support)**
  - *Message repositories for translated z/VM messages*
- **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*
- **TRF (System Trace Files)**
  - *System trace data generated by a virtual machine*
  - *Created by TRSOURCE and TRSAVE commands*

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:

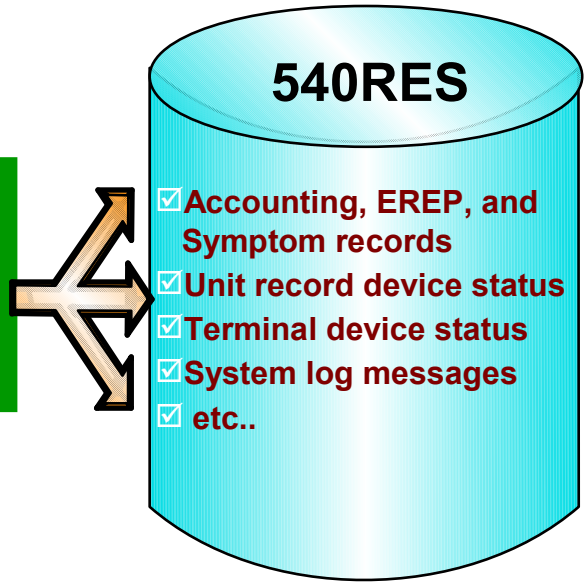
- WARM
  - *Restores all information saved during SHUTDOWN*
- FORCE
  - *Restores as much information as possible*
- COLD
  - *Only restores system data files*
- CLEAN
  - *Does not restore any saved information*

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

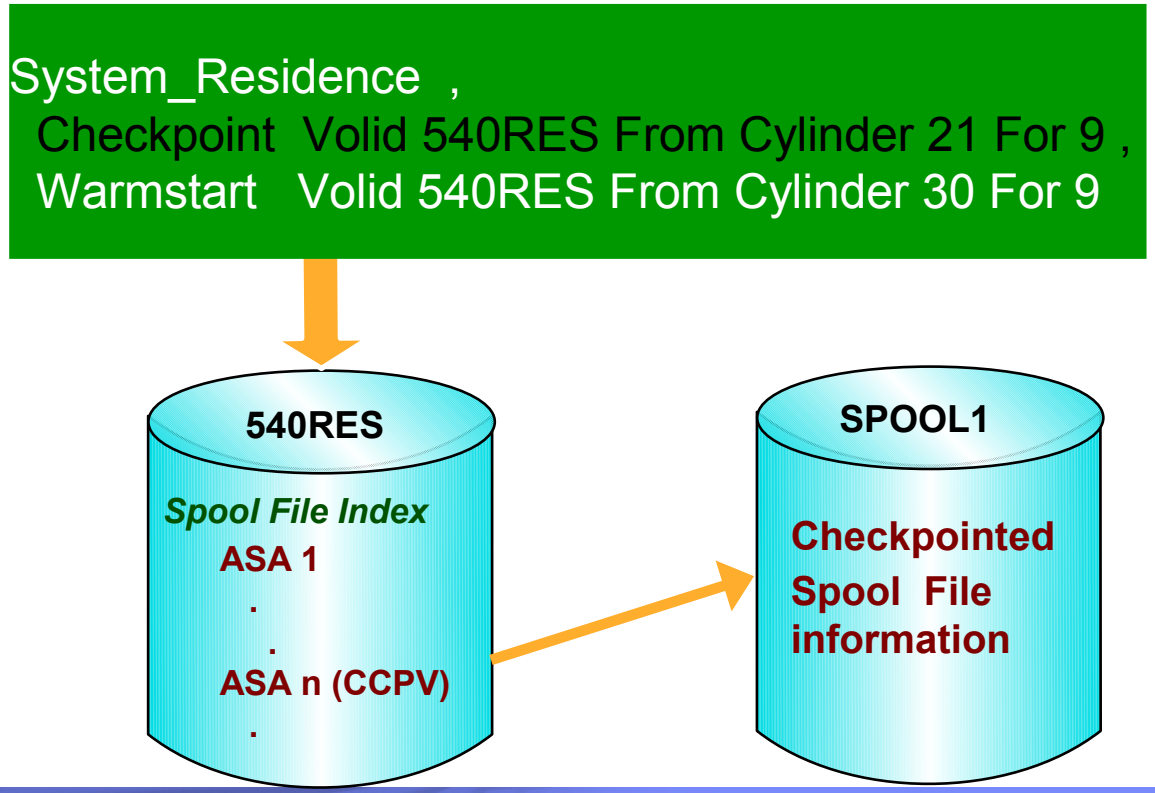
```
System_Residence  ,
  Checkpoint  Volid 540RES From Cylinder 21 For 9 ,
  Warmstart   Volid 540RES From Cylinder 30 For 9
```



Restoring System Data – Warmstart Area

Spool files to be restored during a system restart

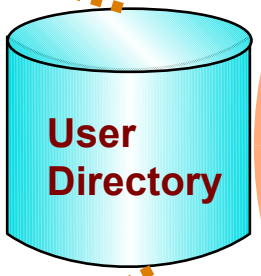
- One entry per file : 4-byte Disk (Auxiliary Storage) Address
- Updated whenever a spool file is created or deleted

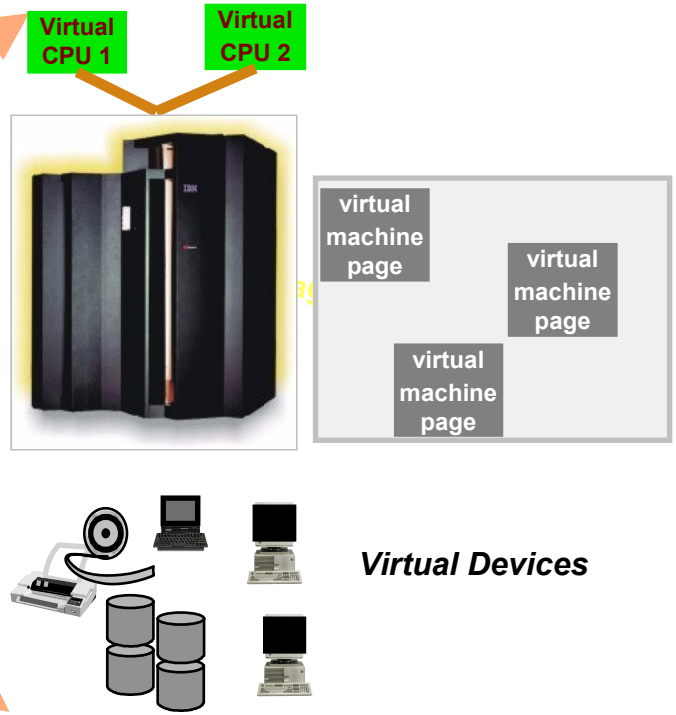


Virtual Machines

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
```



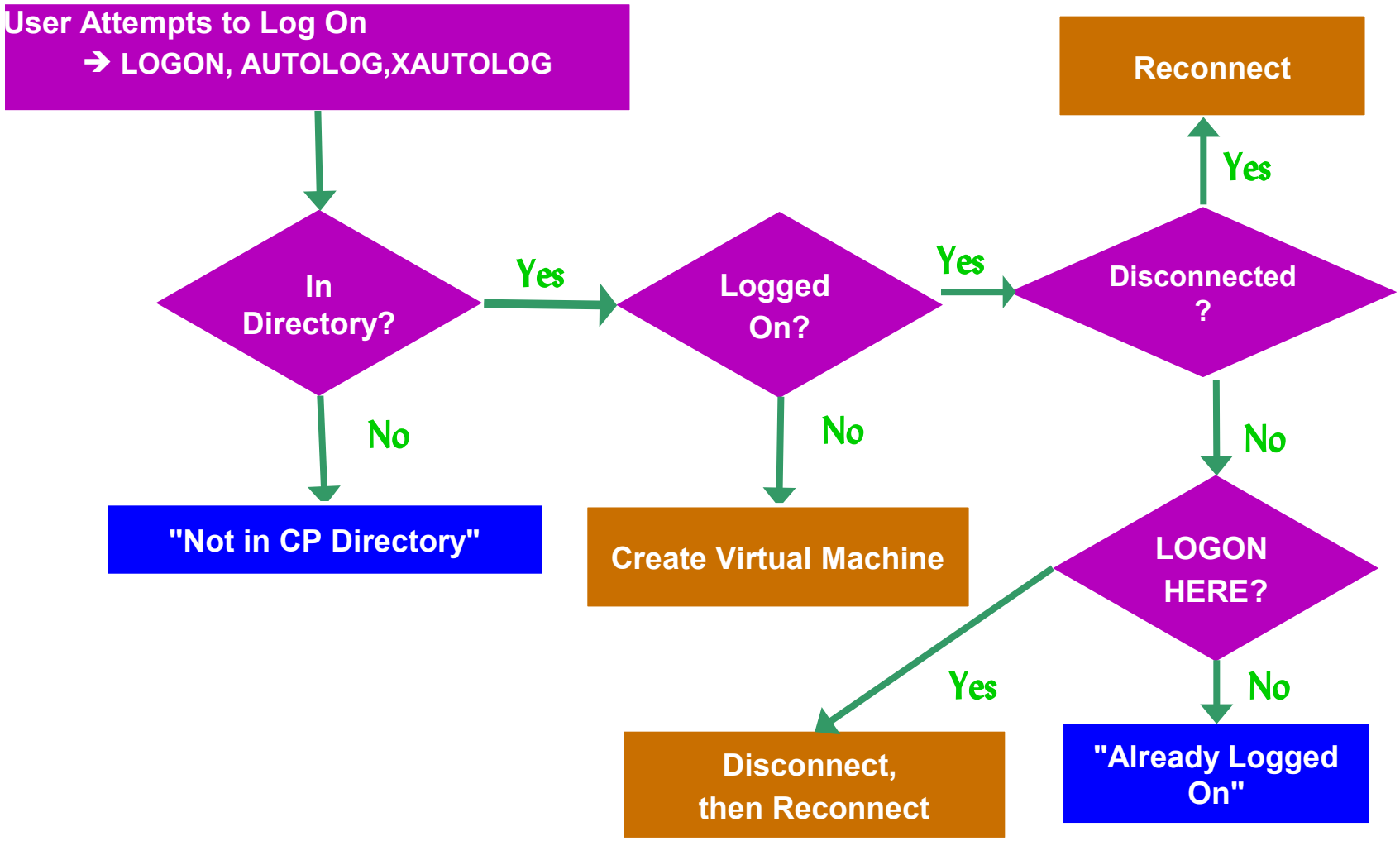


Virtual Devices

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

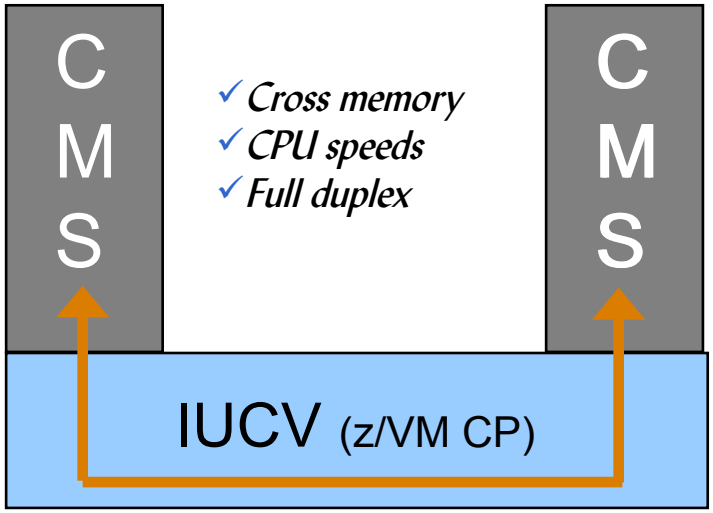
Virtual NIC

- Simulates existence of real Network Interface Cards for each virtual machine

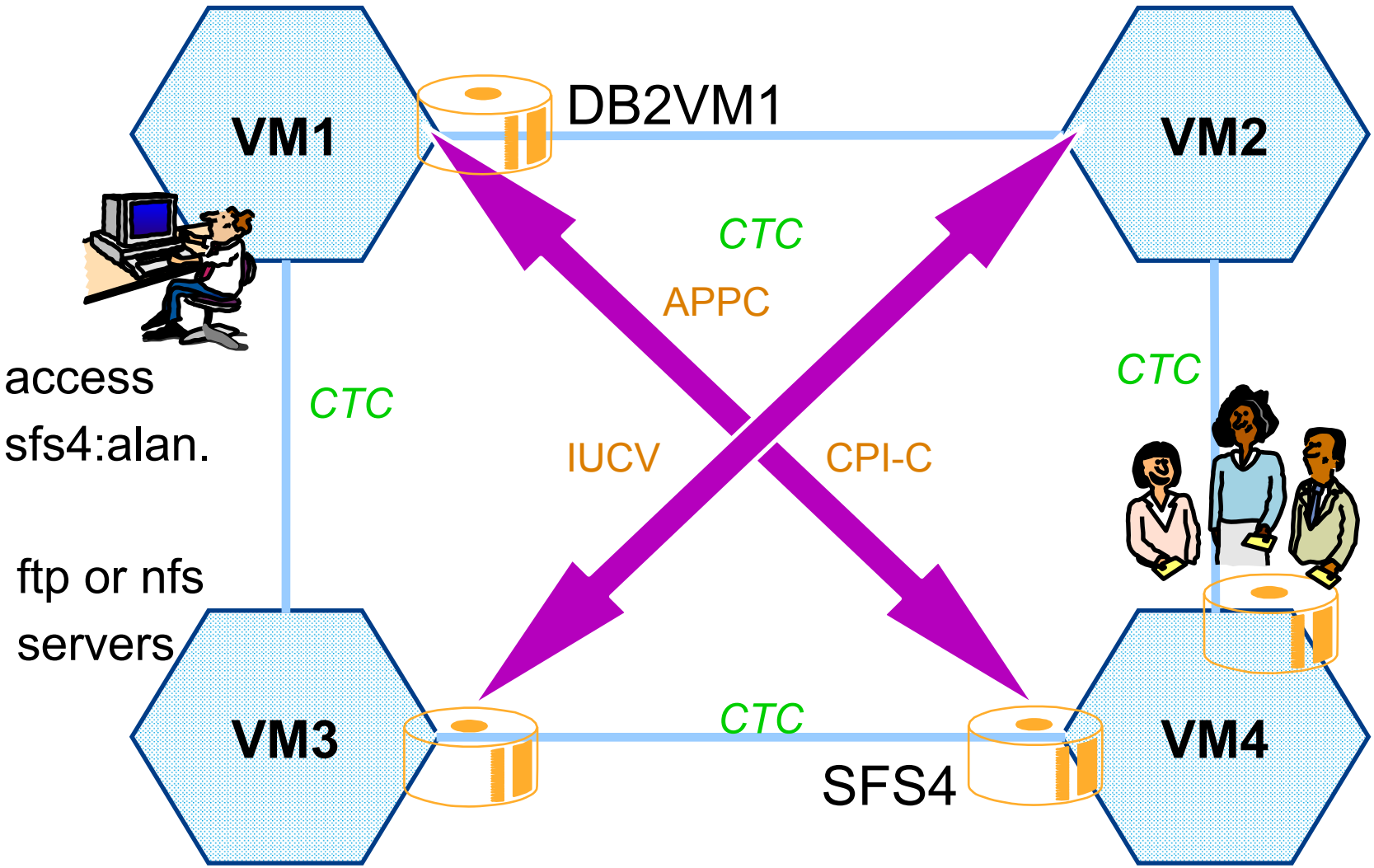
IUCV Communication

Inter-User Communication Vehicle (IUCV)

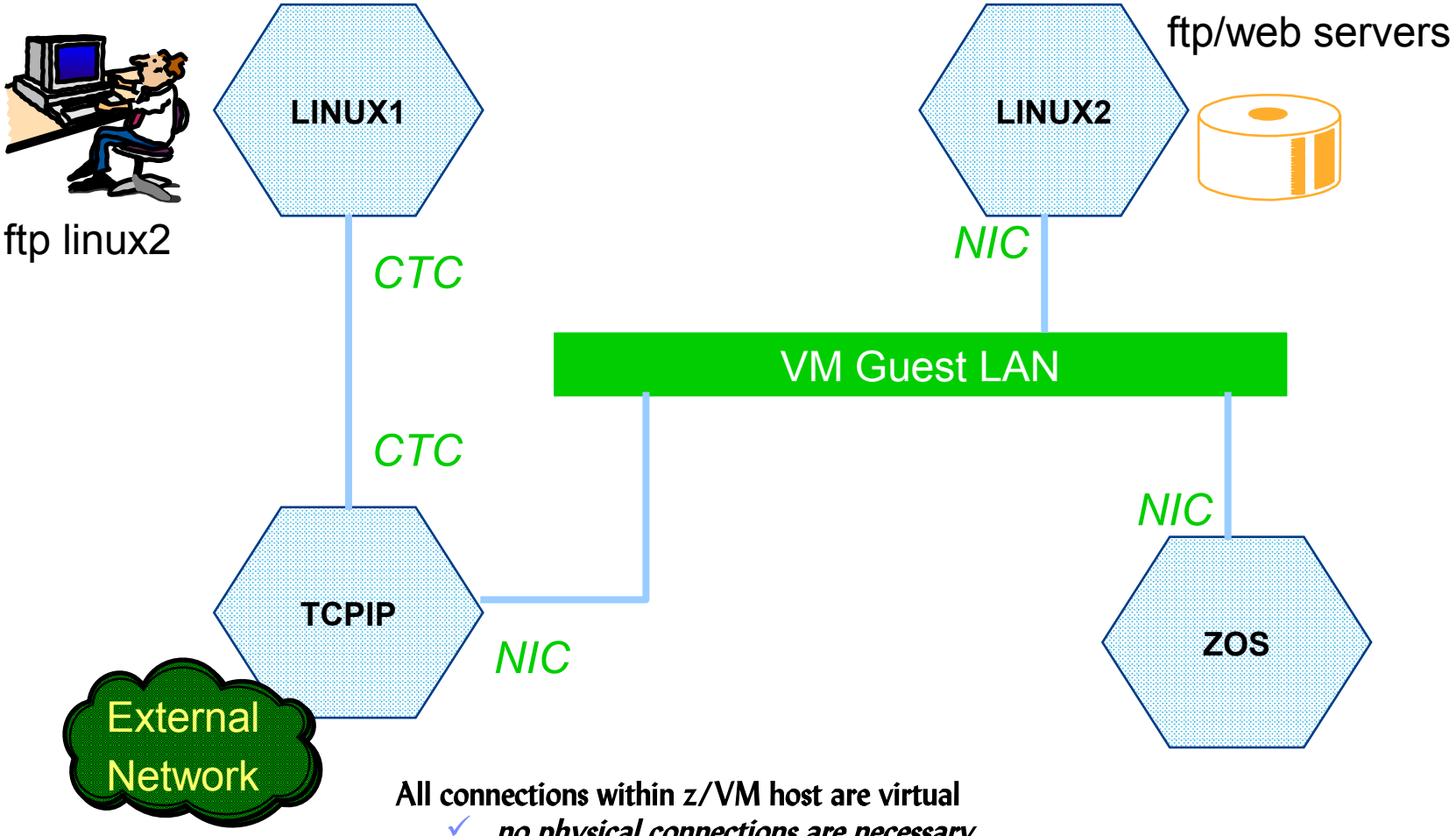
- Allows communication between an application and:
  - *Other virtual machines*
  - *CP system services*
- Simultaneous communication over multiple connections allowed for each virtual machine
- Transparent communication between virtual machines on different systems via ISFC (Inter-System Facility for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



A VM Collection



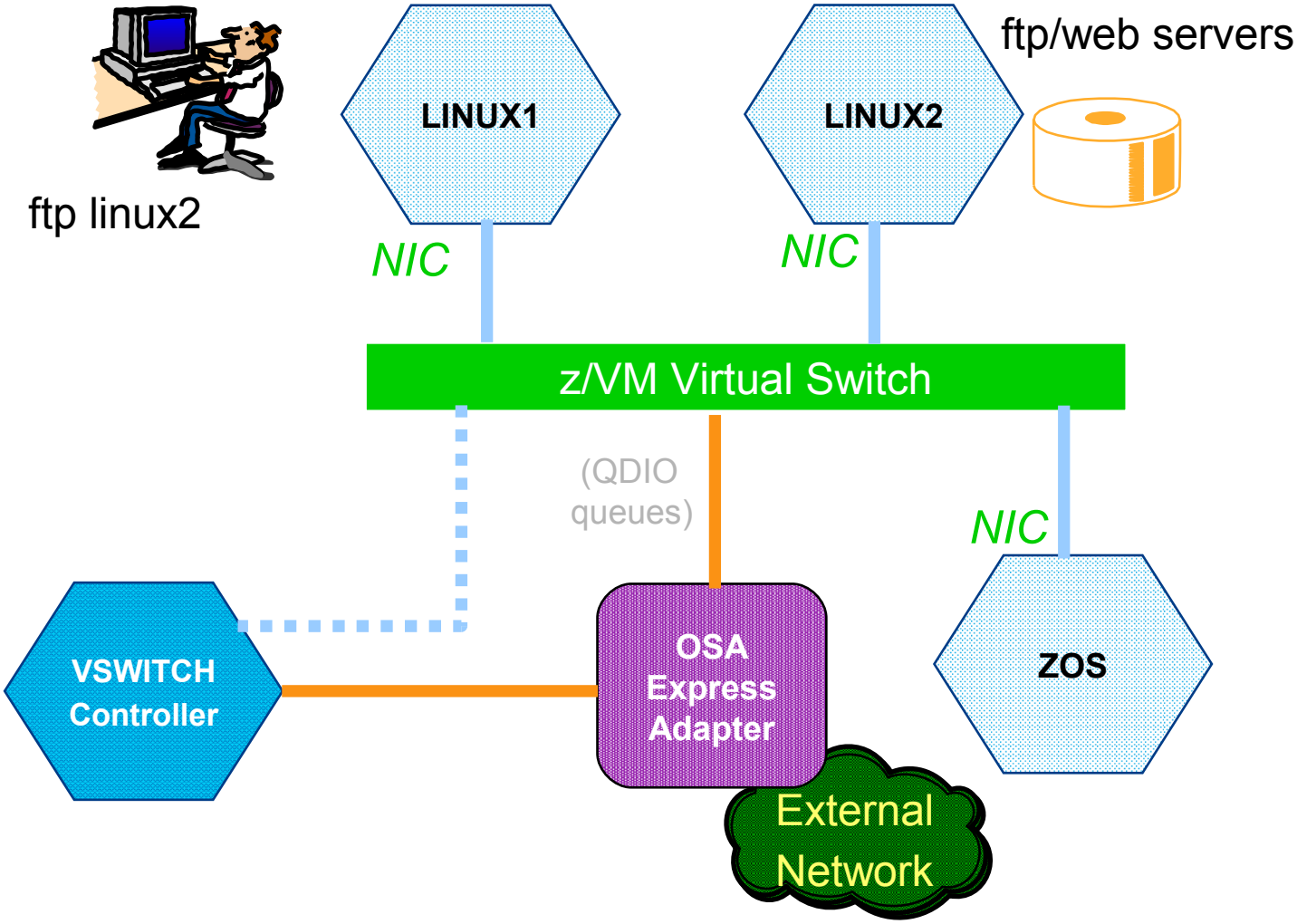
A Virtual Network (Guest LAN)



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 and virtual device settings*
  - *System and user data*
- Assigning/releasing system resources
- Moving data and files between users (virtual machines)
- Communicating between virtual machines

CP Commands…

COMMAND directory statement

- CP command may be specified in virtual machine's directory entry
- Executed after virtual machine logs on but before IPL

FOR command

- Allows a user to issue CP commands on behalf of another user
  - *Issuer must have privilege class C or SECUSER authority for target user*
- Command responses are sent to the issuer
  - *No indication to target user*

- Example - from user OPERATOR:

```
FOR u1 CMD q v stor

Ready;

11:59:21 U1   : STORAGE = 32M
11:59:21 U1   : HCPFOR069I Command Complete.  CP return code = 0000.
```

Privilege Classes

- Each user (virtual machine) has one or more privilege classes
  - *Most are only Class G*
- Identify:
  - *Which CP commands the user is allowed to issue*
  - *Each command's scope of influence*
- May be modified for:
  - *Users*
  - *Commands*

Class	Type of User and Function
A	***System Operator:*** responsible for availability of system and resources
B	***System Resource Operator:*** controls real resources of system, except for those controlled by the system operator and spooling operator
C	***System Programmer:*** Changes system-wide parameters
D	***Spooling Operator:*** Controls spool files and system's real reader, printer, and punch devices
E	***System Analyst:*** Examines and saves system operation data
F	***Service Representative:*** Reserved for IBM use
G	***General User:*** 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
- Etc.

Customizing CP

CP Exit support

- Allows non-disruptive additions and deletions of customized CP code
  - *CP commands*
  - *Diagnose codes*
  - *Message repositories*
  - *Exit routines (user modifications to CP)*
- Modifications are applied with commands or configuration file statements
  - *No need to shutdown and re-IPL to apply user code*
- 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*

Commands may be used to collect additional data

- QUERY
- LOCATE
- DISPLAY
- 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 the user's virtual Reader (RDR)
  - *Use "(FOR *" if PEEKing the file*

CP Dumps

Written to SPOOL or tape

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

Hard Abend

- Contains all of CP owned memory

Soft Abend

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

SNAPDUMP

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

Other information is common to all types of CP dumps

More Dumps

VMDUMP (Virtual Machine Dump)

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

Stand-Alone Dump

- Same format as abend dump
  - *Dumps all of main memory*
- Created when stand-alone dump utility is IPLed
  - *Utility is created by HCPSADMP EXEC*
  - *Can be IPLed to start Stand-Alone Dump*
- Always written to tape

Processing CP Dumps

CP dumps are generally sent to user 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 6 Release 1.0, service level 0000 (CP 64-BIT)
Generated at 02/28/10 22:26:27.000000, IPLd at 02/28/10 22:27:16.976383
Date 02/28/10 Time 22:27:49.293349

CPUID = FF129F30 20978000

CPU address is 0000    Prefix register is 00038000    (failing)
07CE0C20 22:27:42  Call fr HCPGRF+1942 to HCPIOSRQ cpebk 00EE9C00 iac Primary
                   parm 00EC4000

Summary of CP exits
      0 Pre-defined exits found
      0 Dynamic exits found
      0 Diagnose exits found

SVC002 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
07CE0C20 00 22:27:42.035363  Call fr HCPGRF+1942 to HCPIOSRQ cpebk 00EE9C00
07CE0C00 00 22:27:42.035362  Obtain 38 dw (GSD) at 00F4ED28 by HCPGRF+11E0
07CE0BE0 00 22:27:42.035361  Obtain 16 dw (RCW) at 00F78378 by HCPGRS+44A
07CE0BC0 00 22:27:42.035358  Unstack IORBK/TRQBK at 00F18250 vmdbk 00EC4000
07CE0BA0 00 22:27:42.035352  Exit to dispatcher fr HCPIOL+130 vmdbk 00002000
07CE0B80 00 22:27:42.035351  Release 65 dw (IOR) at 00F29DB8 by HCPIFI+ABE
07CE0B60 00 22:27:42.035350  Rtrn to HCPIFI+962 fr HCPGER+2E0 cpebk 00EE9C00
07CE0B40 00 22:27:42.035350  Rtrn to HCPGER+2C8 fr HCPERP+1938 cpebk 00EFC800
07CE0B20 00 22:27:42.035350  Rtrn to HCPERP+18FE fr HCPGREFS cpebk 00F8E600
07CE0B00 00 22:27:42.035349  Rtrn to HCPGRE+EE fr HCPINV+12E cpebk 00EAD600
07CE0AE0 00 22:27:42.035349  Rtrn to HCPINV+2E0 fr HCPUSL+BA cpebk 00EACE00
07CE0AC0 00 22:27:42.035343  Add User vmdbk 00EC4001 OPERATOR
07CE0A80 00 22:27:42.035340  Monitor Call at HCPSCI+94
07CE0A40 00 22:27:42.035336  Monitor Call at HCPSCH+432
07CE0A20 00 22:27:42.035326  Stack IOR/TRQ at 00F18250 by HCPUSL+AA
07CE0A00 00 22:27:42.035323  Obtain 65 dw (IOR) at 00F18250 by HCPUSL+70
07CE09E0 00 22:27:42.035323  Call fr HCPINV+2E0 to HCPUSLAT cpebk 00EACE00
07CE09C0 00 22:27:42.035322  Rtrn to HCPINV+216 fr HCPRBK+AA cpebk 00EACE00
07CE09A0 00 22:27:42.035321  Call fr HCPINV+216 to HCPRBKDA cpebk 00EACE00
07CE0980 00 22:27:42.035321  Release 2 dw (???) at 00EDA738 by HCPINV+122
07CE0960 00 22:27:42.035321  Release 2 dw (???) at 00EDA760 by HCPINV+1FA
07CE0940 00 22:27:42.035320  Obtain 2 dw (???) at 00EDA760 by HCPINV+1BA
......
```

Tracing…

TRACE Command

- Monitors events in virtual machines
  - *Execution of instructions*
  - *Memory alteration*
  - *Register alteration*
  - *I/O activity*
- All occurrences of traceable system events are recorded

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

z/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
- Records diagnostic information
  - Several types of data
  - Many ways to collect it

Additional Resources

z/VM Library

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

IBMVM List server

- http://listserv.uark.edu/scripts/wa.exe?A0=ibmvm