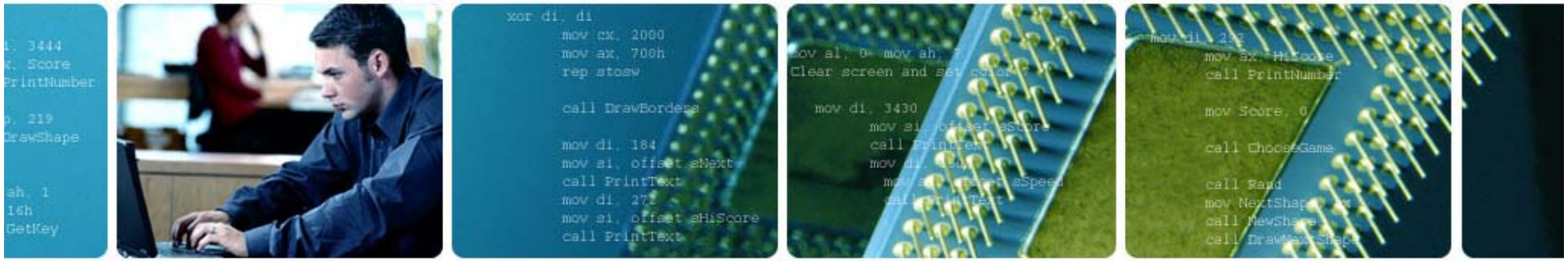
Introduction to the new mainframe

Chapter 5: Working with data sets



Chapter 5 objectives

Be able to:

- Explain what a data set is
- Describe data set naming conventions and record formats
- List some access methods for managing data and programs
- Explain what catalogs and VTOCs are used for
- Create, delete, and modify data sets
- Explain the differences between UNIX file systems and z/OS data sets
- Describe the z/OS UNIX file systems' use of data sets.

Key terms in this chapter

- block size
- catalog
- data set
- high level qualifier or HLQ
- library
- logical record length or LRECL
- member
- PDS and PDSE
- record format or RECFM
- system managed storage or SMS
- virtual storage access method or VSAM
- volume table of contents or VTOC

What is a data set?

A data set is a collection of logically related data records stored on one disk storage volume or a set of volumes.

A data set can be:

- **a source program**
- **a library of macros**
- **a file of data records used by a processing program.**

You can print a data set or display it on a terminal. The logical record is the basic unit of information used by a program running on z/OS.

How data is stored in a z/OS system

- Data is stored on a direct access storage device (DASD), magnetic tape volume, or optical media.
- You can store and retrieve records either directly or sequentially.
- You use DASD volumes for storing data and executable programs, including the operating system itself, and for temporary working storage.
- You can use one DASD volume for many different data sets, and reallocate or reuse space on the volume.

Data management in z/OS

Data management involves all of the following tasks:

- **allocation, placement, monitoring, migration, backup, recall, recovery, and deletion.**

Storage management is done either manually or through automated processes (or through a combination or both).

In z/OS, Data Facility: System-Managed Storage (DFSMS) is used to automate storage management for data sets.

What an access method is

- **Defines the technique used to store and retrieve data.**
- **Includes system-provided programs and utilities to define and process data sets.**
- **Commonly used access methods include the following:**
  - VSAM, QSAM, BSAM, BDAM, and BPAM.

DASD: Use and terminology

Direct Access Storage Device **(DASD) is another name for a disk drive.**

DASD volumes are used for storing data and executable programs.

Data sets in a z/OS system are organized on DASD volumes.

- **A disk drive contains cylinders**
- **Cylinders contain tracks**
- **Tracks contain data records**

Using a data set

To use a data set, you first *allocate* it. Then, access the data using macros for the access method that you have chosen.

Various ways to allocate a data set:

- **ISPF data set panel, option 3.2**
- **Access Method Services**
- **TSO ALLOCATE command**
- **job control language (JCL)**

Allocating space on DASD volumes

How space is specified:

- explicitly (SPACE parameter)
- implicitly (SMS data class)

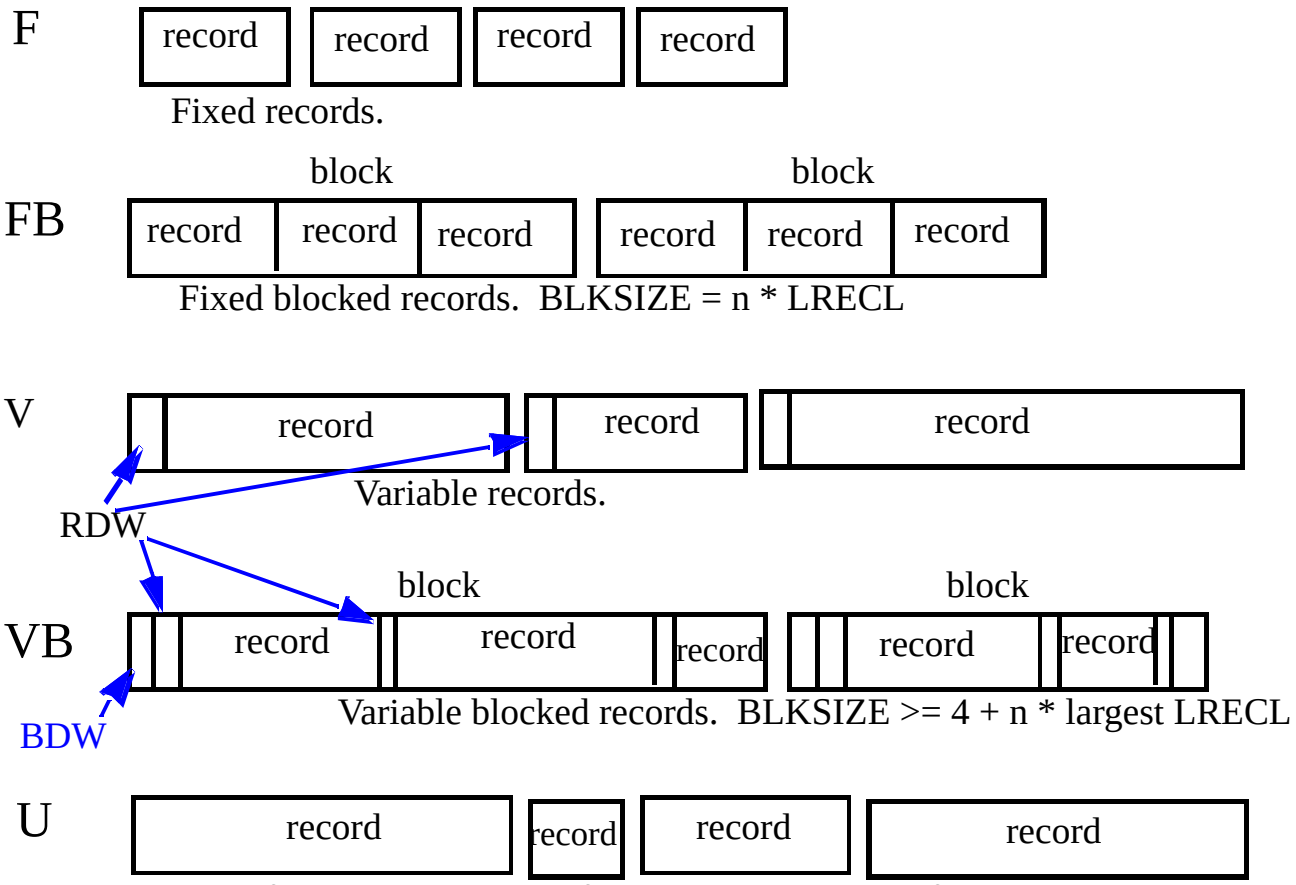
Logical records and blocks:

- Smallest amount of data to be processed
- Grouped in physical records named blocks

Data set extents:

- Space for a disk data set is assigned in extents

Data set record formats



Record and block descriptors words are each 4 bytes long

Types of data sets

We discuss three types in this class:

- **Sequential, partitioned, and VSAM**

A sequential data set is a collection of records written and read in sequential order from beginning to end.

A partitioned data set (PDS) is a collection of sequential data sets, called members.

- **Consists of a directory and one or more members.**
- **Also called a library.**

A PDSE is a partitioned data set extended.

PDS versus PDSE

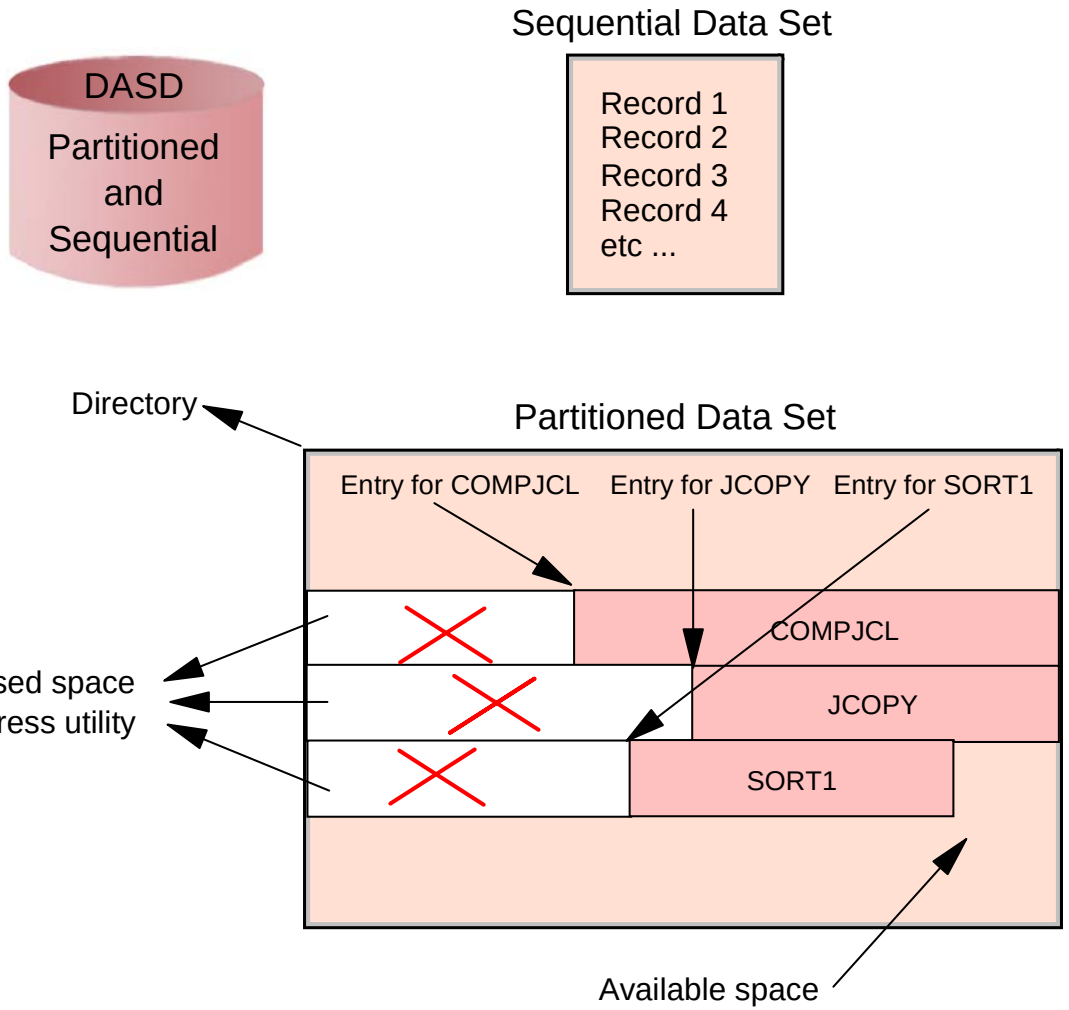
PDS data sets:

- **Simple and efficient way to organize related groups of sequential files.**

PDSE data sets:

- **Similar to a PDS, but advantages include:**
  - Space reclaimed automatically when a member is deleted
  - Flexible size
  - Can be shared
  - Faster directory searches

What is a data set, and how is it stored



VSAM

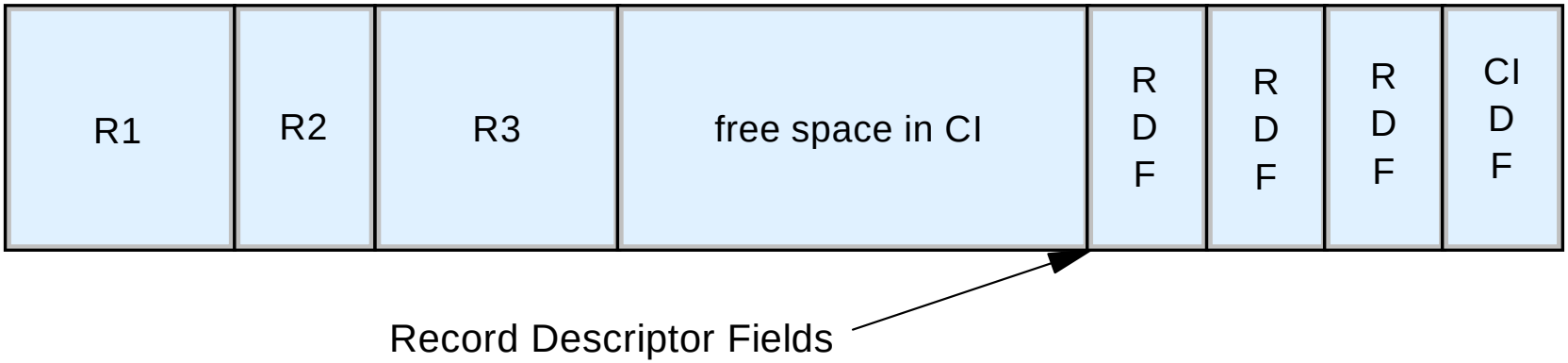
VSAM is *Virtual Storage Access Method*

VSAM provides more complex functions than other disk access methods

VSAM record formats:

- **Key Sequence Data Set (KSDS)**
- **Entry Sequence Data Set (ESDS)**
- **Relative Record Data Set (RRDS)**
- **Linear Data Set (LDS)**

Simple VSAM control interval



How data sets are named

Data set naming convention

- **Unique name**
  - Maximum 44 characters
- **Maximum of 22 name segments: level qualifier**
  - The first name in the left: high level qualifier (HLQ)
  - The last name in the right: low level qualifier (LLQ)
  - Level qualifiers are separated by '.'
- **Each level qualifier:**
  - From 1 up to 8 characters
  - The first must be alphabetical (A-Z) or special (@ # $)
  - The 7 remaining: alphabetical, national, numeric (0-9) or hyphen (-)
  - Upper case only
- **Example: MYID.JCL.FILE2 HLQ: MYID 3 qualifiers**

Member name of partitioned data set

- **8 bytes long**
- **First byte: alphabetical (A-Z) or special (@ # $)**
- **The 7 remaining: alphabetical, special, numeric (0-9)**

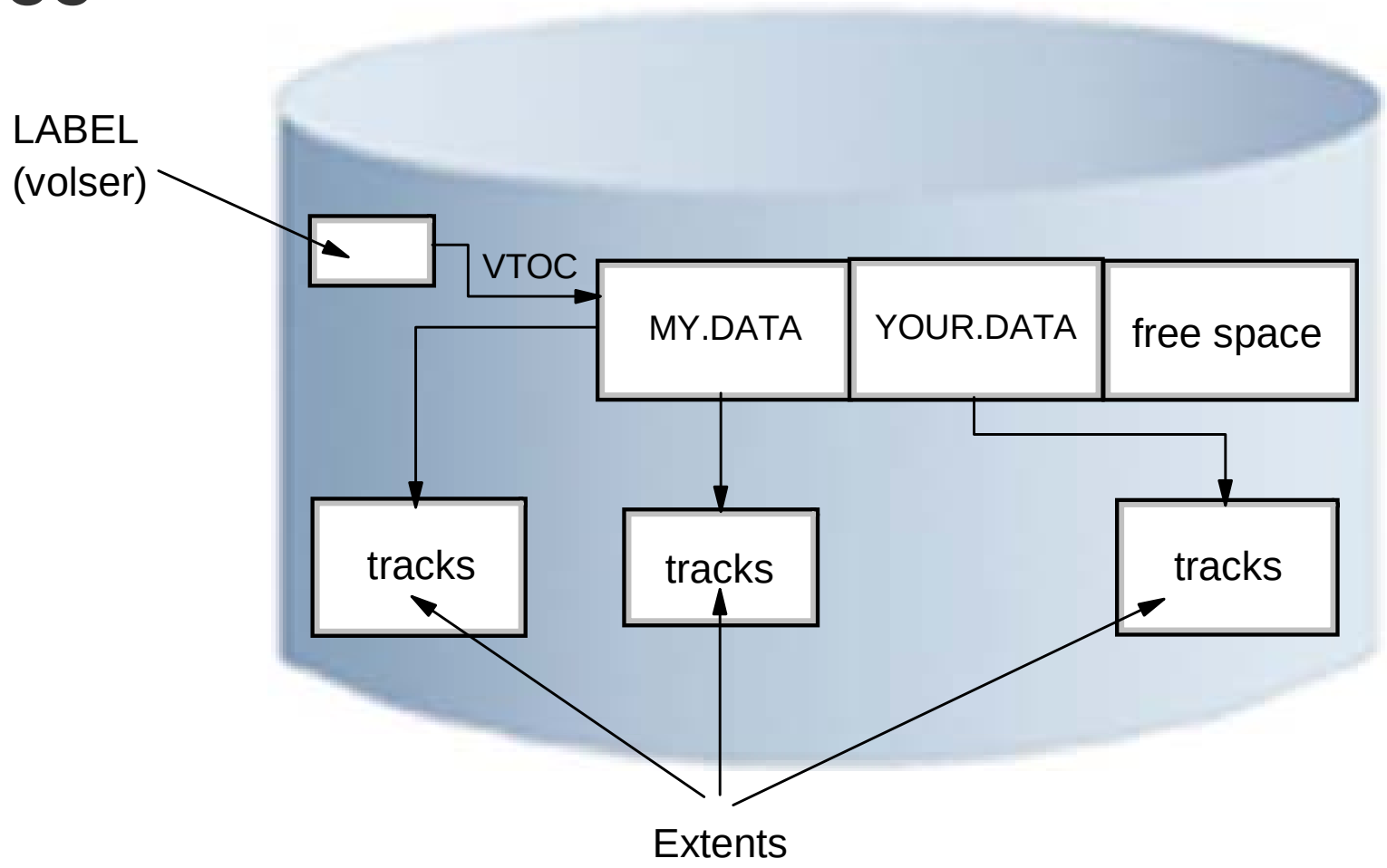
Catalogs and VTOCs

z/OS uses a catalog and a volume table of contents (VTOC) on each DASD volume to manage the storage and placement of data sets.

VTOC:

- **Lists the data sets on a volume**
- **Lists the free space on the volume.**

VTOC



How a catalog is used

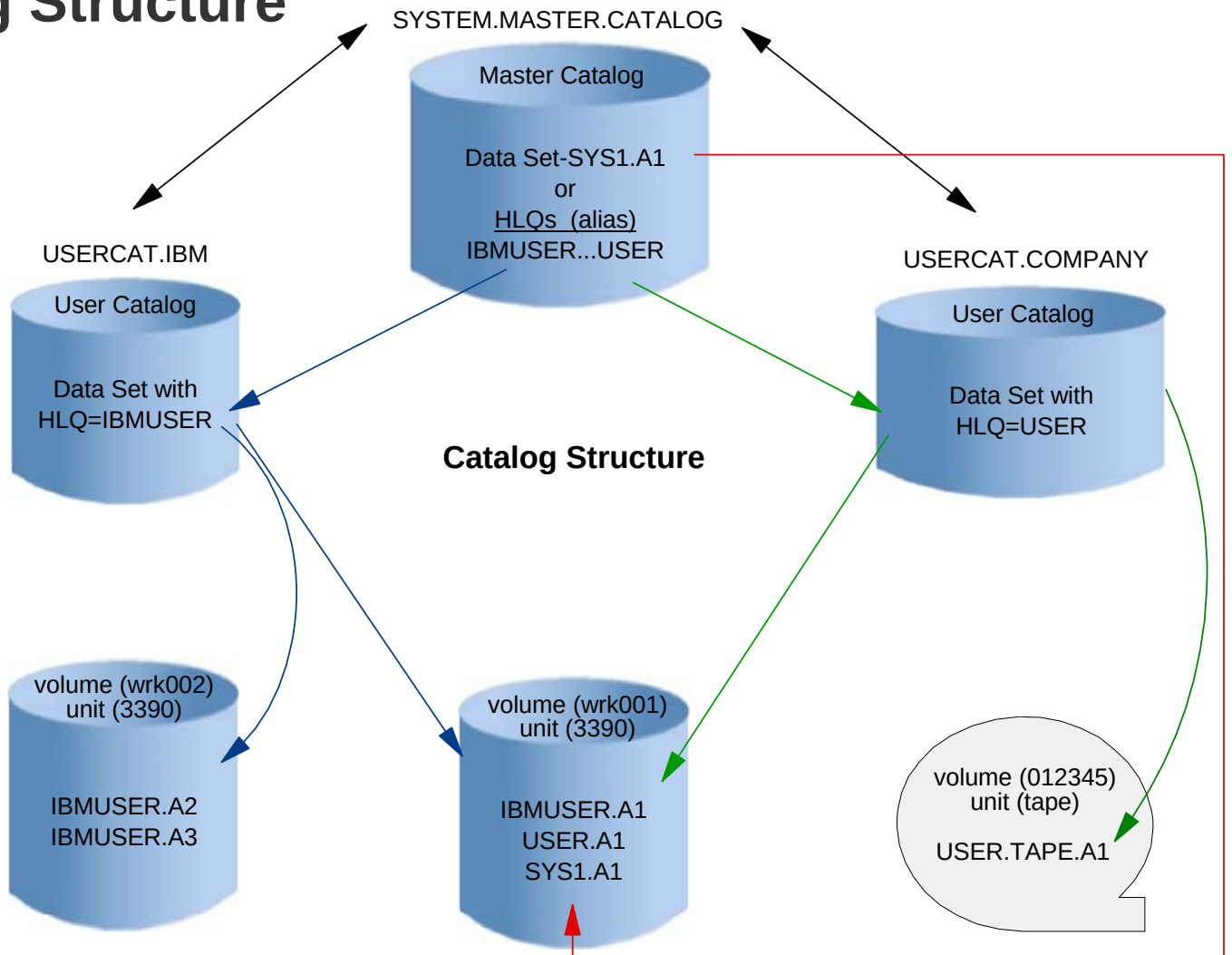
A catalog associates a data set with the volume on which the data set is located.

Locating a data set requires:

- **Data set name**
- **Volume name**
- **Unit (volume device type)**

Typical z/OS system includes a master catalog and numerous user catalogs.

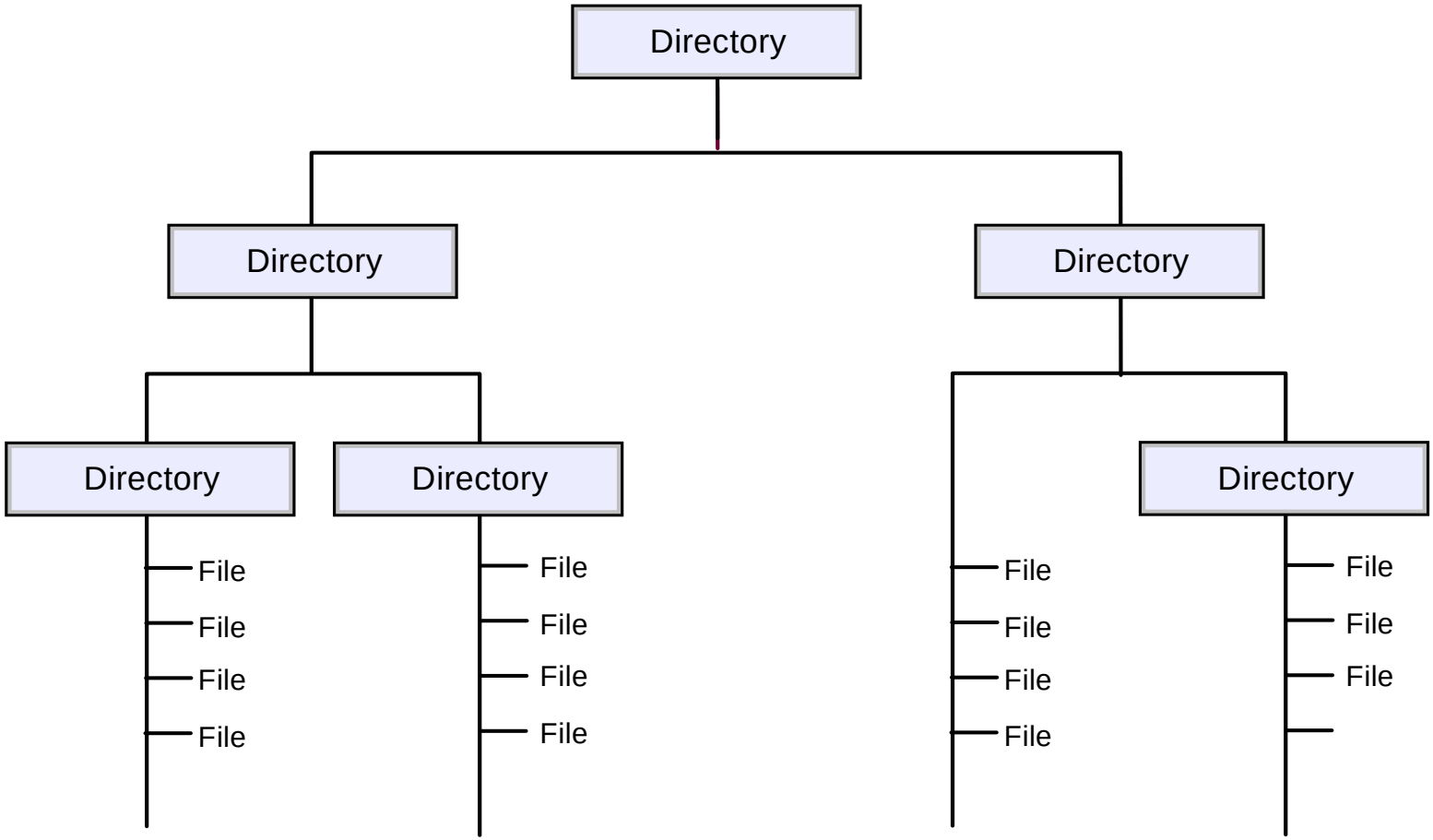
Catalog Structure



z/OS UNIX file systems

- **z/OS UNIX System Services (z/OS UNIX) allows z/OS to access UNIX files.**
- **A z/OS UNIX file system is hierarchical and byte-oriented.**
- **Files in the UNIX file system are sequential files and are accessed as byte streams.**
- **UNIX files and traditional z/OS data sets can reside on the same DASD volume.**

UNIX file system structure



Summary

- A data set is a collection of logically related data (programs or files)
- Data sets are stored on disk drives (DASD) and tape.
- Most z/OS data processing is record-oriented. Byte stream files are not present in traditional processing, except in z/OS UNIX.
- z/OS records follow well-defined formats, based on record format (RECFM), logical record length (LRECL), and the maximum block size (BLKSIZE).
- z/OS data set names have up to 44 characters, divided by periods into qualifiers.

Summary (continued)

- Catalogs are used to locate data sets.
- VSAM is an access method that provides more complex functions than other disk access methods.
- z/OS libraries are known as partitioned data sets (PDS or PDSE) and contain members.
- A file in the hierarchical file system can be either a text file or a binary file.
- z/OS treats an entire UNIX file system hierarchy as a collection of “data sets.” Each data set is a mountable file system.