

Chapter 1

VSAM: INTRODUCTION AND OVERVIEW

*Get on the
Fast Track!*



TM

**SYS-ED/
Computer
Education
Techniques, Inc.**

Objectives

You will learn:

- VSAM components.
- Organization of VSAM - catalogs and datasets.
- VSAM dataset types: ESDS: Entry-sequence Datasets, RRDS: Relative Record Datasets, KSDS: Key-sequenced Datasets, VRRDS: Variable Relative Record Dataset, and LDS: Linear Dataset.
- Alternative indexes and paths.
- Control Intervals and Control Areas.
- VSAM dataset definition.
- Access Method Services.
- AMS commands, functional commands, and modal commands.

1 VSAM: Origin and Dataset Organization

In the late 1970's, IBM designed and established a collection of three dataset organizations.

Sequential	Indexed	Direct-access
------------	---------	---------------

VSAM: Virtual Storage Access Method is the collective name of the dataset organizations.

Each dataset has its own access methods and utilities.

VSAM:

- Replaced the ISAM: Indexed Sequential Access Method.
- Is used in mainframe installations to organize and access data.
- Datasets are cataloged in an ICF: Integrated Catalog Facility structure.
- Datasets must be defined in an ICF catalog.
- Records are arranged by an index key or by relative byte addressing.
- Uses direct or sequential processing of fixed and variable length records on DASD: Direct Access Service Devices.

2 VSAM Components

There are two major components to VSAM:

Record management	Catalog management
-------------------	--------------------

2.1 Record Management

The record management function in VSAM is what the application programmer uses most often. It maintains records in a VSAM dataset for an application.

Modern VSAM has five dataset organizations:

KSDS	Key-sequenced dataset
ESDS	Entry-sequenced dataset
RRDS	Fixed-length relative record dataset
VRRDS	Variable-length relative record dataset
LDS	Linear dataset

Each dataset organization stores and accesses its records differently.

3 VSAM Concepts

Logical record	Unit of information used to store data in a VSAM dataset. The terms logical record and record are the functionally equivalent.
Physical record	Is device dependent and is determined by the system.
Control interval	Logical records are contained in a CI: Control Interval. VSAM uses a control interval as its smallest unit of information to transfer. It is the building block of every component of a VSAM dataset.

A CI contains:

- 1- Logical records
- 2- Unused space, referred to as free space
- 3- Control information

The control information contains two types of fields:

- One CIDF: Control Interval Definition Field per CI.
- One or more RDF: Record Definition Fields per logical record.

3.1 Terminology and Definitions

Control area	A CA: Control Area is two or more CIs placed together into fixed-length contiguous areas of DASD.
Splits	CI and CA splits occur as a result of data record insertions. If a record is to be inserted and there is not enough free space in the CI, a reorganization of the data will occur.
RBA: Relative Byte Address	RBA is the physical location of a logical record contained in a dataset. It is represented by the offset of a logical record from the beginning of the dataset.
Data component	The data component contains the data records.
Index component	The index component contains the index. The key is a value taken from a fixed defined field in each logical record and determines the record's position in the dataset.
Cluster	For a KSDS, a cluster is the combination of a data component and an index component. RRDS, ESDS, and LDS formats are considered to be clusters without index components.
Sphere	A sphere is a VSAM cluster and its associated AIXs: Alternate Index datasets.
Alternate indexes	AIXs enable the logical records of an ESDS or of a KSDS to be accessed sequentially and directly by more than one key field.

4 Organization of VSAM

The VSAM objects are:

Catalogs	Datasets
----------	----------

Clusters are one or more datasets defined in the catalog and they consist of:

Primary Index	Alternate Indexes	Paths
---------------	-------------------	-------

4.1 VSAM Catalogs

All VSAM objects must be cataloged to a VSAM catalog. Uncataloged VSAM datasets do not exist.

There can be an unlimited number of user catalogs, but need not be any.

A user catalog must be connected to the master catalog, to be used as a catalog. Each user catalog points back to the master catalog to which it is connected.

A VSAM object must be in the master catalog, or in one user catalog, but not both.

User catalogs:

- Reduce the size of, and minimize contention for, the master catalog.
- Minimize the effect of a damaged master catalog.
- Make VSAM volumes portable from one system to another.

5 VSAM Datasets

Records are identified by a key or address.

Key

Logical key	Contents of a field in the same place, for the same length, in each record; independent of the physical location of the record in the dataset.
Relative record number	Displacement in records from the beginning of the dataset. The first record has a RRN: Relative Record Number of 1.

Address

Relative byte address	Displacement in bytes of the first byte of a record from the beginning of the dataset. The RBA of the first record is zero; the RBA of the second record is equal to the length of the first record.
-----------------------	--

5.1 ESDS: Entry-sequence Datasets

Records are stored in the sequence in which they are added or loaded.

These characteristics apply to the records:

- Records are identified by RBA.
- Records can be variable-length or fixed-length.
- Sequential and direct retrieval of the records can be used.
- Adds are done to the end of the dataset.
- Updates cannot change record length.
- Deletes are not supported.
- Spanned records can be utilized.
- Alternate Indexes can be utilized.

5.2 RRDS: Relative Record Datasets

A RRDS: Relative Record Dataset consists of strings of fixed-length slots, which may or may not contain records.

RRDS are characterized by:

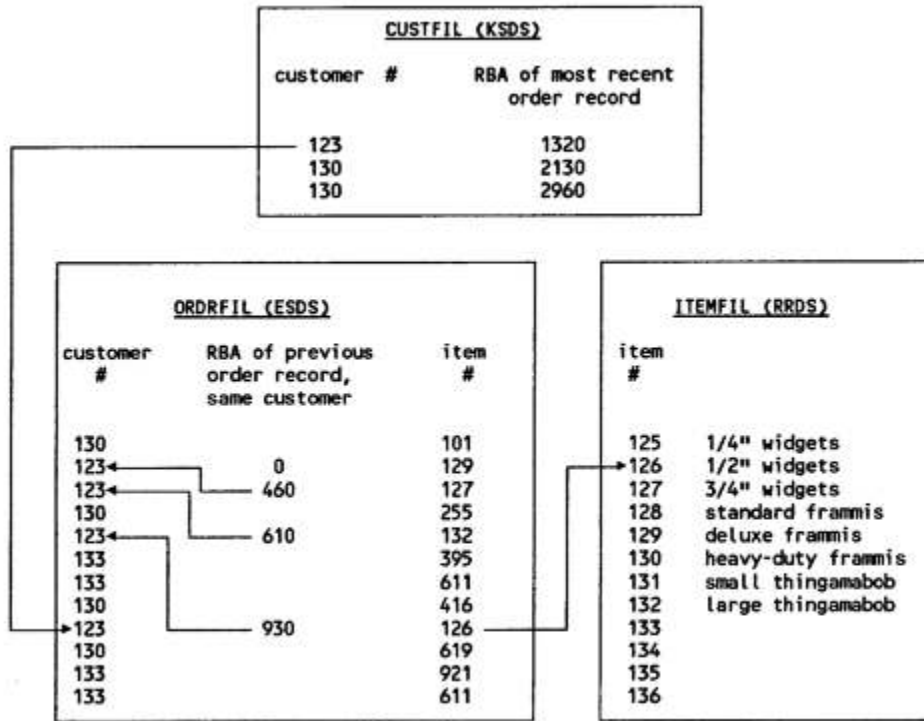
- Records identified by RRN.
- Fixed-length records only.
- Sequential and direct retrieval.
- Records added to empty slots.
- Updates which cannot change record length.
- Deletes marking an slot empty.
- Not allowing spanned records.
- Not allowing alternate indexes.

5.3 KSDS: Key-sequenced Datasets

Key-sequenced cluster consists of data component and index component.

KSDS are characterized by:

- An index component which contains records which relate logical keys to physical locations (RBA) of data records.
- Records which are identified by logical data (which must be unique) or RBA (but RBA may change if record is moved because of update or add).
- Variable-length or fixed-length records.
- Sequential and direct retrieval.
- Records added in place using distributed free space.
- Updates which may change record length.
- Deletes which physically erases a record, making space available.
- Spanned records.
- Alternate Indexes.



5.4 VRRDS: Variable Relative Record Dataset

A VRRDS is similar to a fixed-length RRDS, except that it contains variable-length records.

VRRDS are characterized by:

- Each record having a unique relative record number.
- Each record being placed in ascending relative record number order.
- Each record being stored and retrieved using its relative record number.
- Having no slots.
- The relative record number of a record not be changeable.
- Erased records can be reused for a new record.
- VRRDS is a KSDS processed as an RRDS; therefore an index will be created.

5.5 LDS: Linear Dataset

A linear dataset contains data that can be accessed as byte-addressable strings in virtual storage. It is a VSAM dataset with a control interval size of 4096 bytes.

A LDS has no imbedded control information in its CI, there are no RDFs and CIDFs.

Logical records must be blocked and deblocked by the application program; but records do not exist from the point of view of VSAM.

As with the ESDS and RRDS, an LDS contains a data component only.

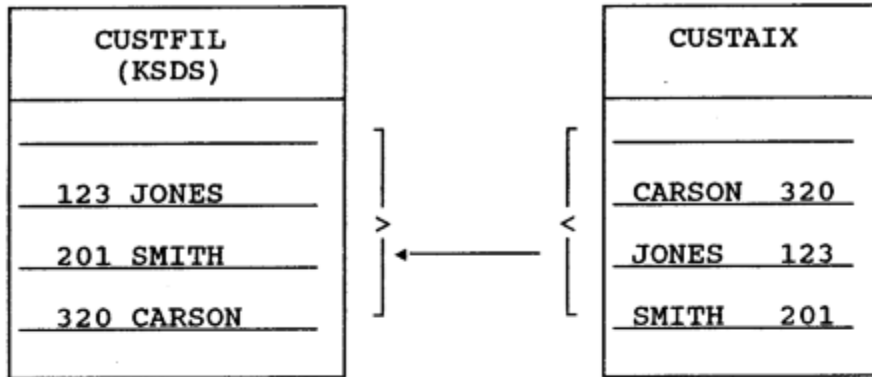
6 Alternative Indexes and Paths

A LDS: Linear Dataset contains data that can be accessed as byte-addressable strings in virtual storage. It is a VSAM dataset with a control interval size multiple of 4096 bytes.

Alternate indexes may be constructed over entry-sequenced and key-sequenced datasets.

VRRDS are characterized by:

- An alternate index which is an index to the data in the base cluster, based on an alternate key in each data record.
- Alternate keys not required to be unique.
- A base cluster having any number of alternate indexes.
- A path which is defined from a given alternate index to its base cluster.
- A base cluster or any of its paths which can be opened to provide access to a given dataset by different logical keys, as well as by RBA.
- An alternate index which is a KSDS and can be processed as a dataset in its own right, as well as a component of a path.
- The upgrade set of a base cluster consisting of those alternate indexes which are to be updated when the base cluster is updated.

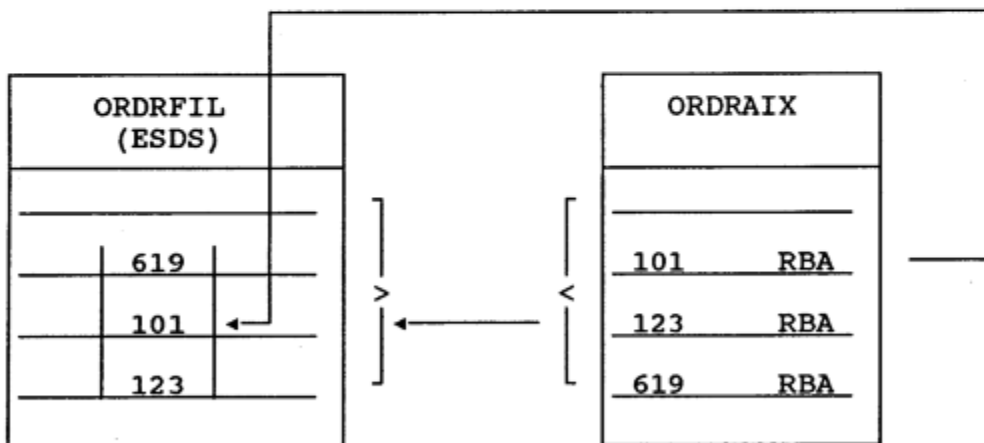


Base Cluster

Alternate Index

Base Key = customer #

Alternate key = customer name



Base Cluster

Alternate Index

No base key
Access is by RBA

Alternate key = customer #

7 Control Intervals and Control Areas

The control interval is the basic unit of information transferred by VSAM between main storage and peripheral storage.

Control interval size is set by the programmer or VSAM.

- All control intervals in a dataset are the same size.
- CI size must be between 512 and 32,768 bytes, and must be a multiple of 512, or of 2048 if CI size is greater than 8192 bytes.
- CI size can be greater than track size.

A control interval contains one or more logical records, unused space, and control information.

- Unused space may or may not be usable, depending on record size and dataset type.
- For KSDS and ESDS, a record may span multiple control intervals.
- A control interval is stored on DASD as one or more fixed-length physical records.
- When a CI is read or written all the physical records which make up the CI are read or written.
- Physical record size is chosen by VSAM to optimize DASD space utilization, based on CI size and track size. When possible it will be equal to CI size.

Control intervals are grouped by VSAM into control areas.

- VSAM determines for a given dataset how many control intervals per control area.
- A control area may consist of control intervals containing data or empty control intervals.
- A control area is a contiguous area of DASD storage.

minimum = 1 track

maximum = 1 cylinder

- A control area is the unit of space that VSAM preformats when a dataset is loaded or extended.

8 VSAM Dataset Definition

A VSAM dataset can be defined using these methods:

- IDCAMS DEFINE or ALLOCATE commands.
- ALLOCATE command from a TSO terminal.
- JCL DD statements; all datasets can be defined directly through JCL.
- Dynamic allocation.

9 Access Method Services

IDCAMS is the all-purpose VSAM utility program used to create and maintain VSAM catalogs, data spaces, datasets, etc.

AMS: Access Method Services functions are specified using AMS commands and their associated parameters.

AMS is invoked by:

- Executing IDCAMS as a job step.
- Calling IDCAMS from another program.
- Executing AMS commands under TSO.

10 AMS Commands

Format

COMMAND Parameters terminator

- Command must begin between left and right margins.
 - The defaults are columns 2 and 72.
- Command is separated from its parameters by one or more separators.
 - Separators are blanks, commas, and comments.
- A comment is any character string surrounded by /* and */.
 - Character string must not include /*.
- Most commands may be abbreviated.
- A command may be continued on several lines.
 - Each line except the last must have a hyphen or plus sign as the last non-blank character before or at the right margin.
- A hyphen indicates continuation of the command.
- A plus sign indicates continuation of the command and continuation of a value within the command.

Parameters

- Positional parameters are required and must be coded in a pre-defined sequence.
- All positional parameters must appear before any keyword parameters.
- A keyword parameter is preceded by a keyword which defines the meaning of the parameter.
- Keyword parameters may be coded in any sequence, following all positional parameters.
- A keyword parameter may be in the form of a list of subparameters.
 - If the subparameter list is positional, the first subparameter is required; the others may be optional.
 - If more than one positional subparameter is coded, the subparameter list must be enclosed in parentheses.
- Parameters and subparameters must be separated by separators.

Terminators

- End-of-command is indicated by a semi-colon or by the absence of a continuation mark.
- All characters to the right of a semi-colon are ignored.

10.1 Functional Commands

Functional commands are used to request AMS services.

Command	Description
ALTER	Alters previously defined catalog entries.
BLDINDEX	Constructs alternate indexes to existing datasets.
DEFINE	Creates catalog entries for catalogs, data spaces, and datasets.
DELETE	Deletes catalog entries.
EXPORT	Copies a VSAM dataset for backup or for transportation to another system.
IMPORT	Re-establishes a dataset from a backup copy or establishes a dataset in another system.
LISTCAT	Lists catalog entries.
PRINT	Prints VSAM, ISAM, and SAM datasets.
REPRO	Copies datasets, converts ISAM and SAM datasets to VSAM, backups VSAM catalogs, reads backup copies of catalogs, and copies catalogs.
VERIFY	Corrects a catalog entry for a dataset which was not closed properly.

These commands can be used on non-VSAM datasets:

DEFINE, ALTER, DELETE, LISTCAT, REPRO, PRINT

10.2 Modal Commands

Modal commands allow conditional execution of functional commands.

Modal commands are not available under TSO.

Command	Description
IF....THEN....ELSE	Allows the conditional execution of a command or sequence of commands based on the value of a condition code.
DO....END	Defines the beginning and end of a command sequence for reference in an IF command.
SET	Changes the value of the condition code.
PARM	Specifies diagnostic aids and printed-output options.