MAINTEC

VSAM

MAINTEC

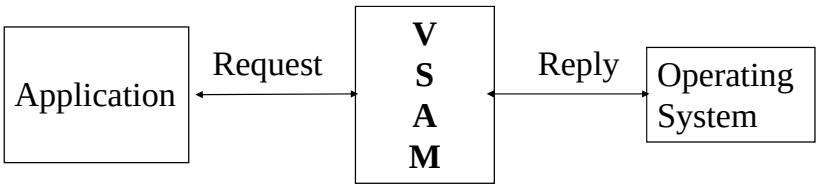
VSAM Fundamentals

- Course Objective
  - Introduction to VSAM
  - VSAM Data Set organization
  - IDCAMS

MAINTEC

What is VSAM?

- VSAM is Virtual Storage Access Method
- It is a method used to move data between Disk and Main Storage
- VSAM operates in Virtual Environment
- VSAM acts as interface between Operating System and Application Program



MAINTEC

Types of VSAM data sets

- ESDS Entry Sequenced Data Set
- KSDS Key Sequenced Data Set
- RRDS Relative Record Data Set
- LDS Linear Data Set

MAINTEC

Traditional access methods

- QSAM (Queried Sequential Access Method)
- BSAM (Basic Sequential Access Method)
  - for 'flat' files
- ISAM (Index Sequential Access Method)
  - for Index files
- BDAM (Basic Direct Access Method)
  - for direct access files

MAINTEC

VSAM History

- 1973 - VSAM was introduced with ESDS & KSDS
- 1975 - Alternate Index, RRDS, Catalog Recovery features introduced
- 1979 - VSAM re-introduced with ICF (Integrated Catalog Facility)
- 1983 - DFP / VSAM ran under MVS / XA & (Data Facility Product)
- 1987 - DFP / VSAM 2.3 introduced with LDS
- 1988 - VSAM 3.1 under MVS / ESA with DFSMS (Data Facility Storage Management System)
- 1991 - VSAM 3.3 introduced with variable length records for RRDS.

MAINTEC

Advantages

- One method to support all types of data retrieval
- Can be a stand alone file or within a DBMS
- Supports fixed and Variable Records
- One utility, IDCAMS, to manage everything (IEBGENER for seq files, IEBCOPY for PDS, IEBISAM for ISAM files)
- Supports alternate index
- Device Independence (EXPORT and IMPORT)
- Portable across systems
- Modular, catalog contains all the information
- Catalogs are protected by internal security

MAINTEC

Disadvantages

- Require lot of DASD space
- For KSDS, primary key cannot be changed
- Performance can be slow because of complexity of Index
- Only resides on DISK

MAINTEC

IDCAMS

- Handles VSAM data sets (exclusively)
- Some of the functions
  - Creates (DEFINE)
  - Copies (REPRO)
  - Prints (PRINT)
  - Delete (DELETE)
  - Lists Characteristics (LISTCAT)

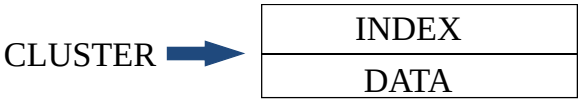
MAINTEC

VSAM data set organization

MAINTEC

VSAM data set organization

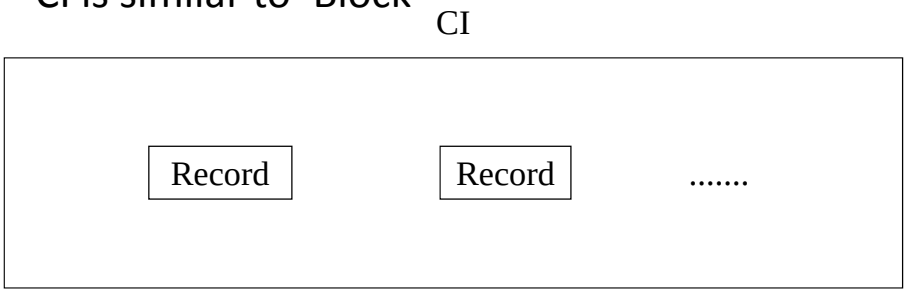
- VSAM Data Set can contain three major components
  - CLUSTER (Catalog entry)
  - INDEX
  - DATA (Actual data)
- Data Set is referred by cluster name in JCL



MAINTEC

VSAM internals

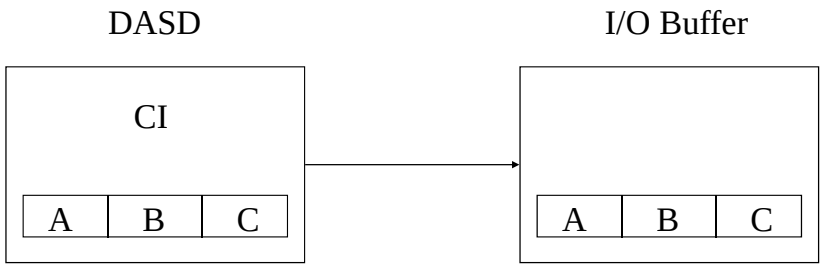
- CONTROL INTERVAL (CI)
- VSAM stores Data and Index in Control Intervals (CI)
- CI is similar to 'Block'



MAINTEC

Control Interval

- CI is the unit of retrieval between DASD and I/O Buffer (Virtual Storage)



MAINTEC

Control Interval

- CI contains
  - Records (or DATA)
  - Free space (Optional)
  - Control Interval Definition Field (CIDF)
  - Record Definition Field (RDF)

MAINTEC

RDF and CIDF

- RDF
  - 3 bytes long
  - Indicates length of records
  - And also, how many adjacent records are of the same length
- CIDF
  - 4 bytes long
  - One per CI
  - Indicates Free space

MAINTEC

CONTROL AREA (CA)

- CIs are grouped into CA
- Can have more than one CA in a VSAM data set
- CA is VSAMs internal unit for allocating space
- Smallest is a TRACK, and the largest is a CYLINDER
- VSAM determines CA size & number of CIs in CA based on
  - CI size specified
  - Record size

MAINTEC

Index

- Separate entity
- Organized similar to Data component (CI & CA)
- Has different CI size
- VSAM builds Index when data is loaded
- Index is organised as Inverted Binary Tree
- VSAM compresses keys to conserve space
- Can have several Levels of Indexes
  - Lowest level of index is called ‘Sequence set’
  - Next level is called index set

MAINTEC

CI & CA Split

- CI split happens while
  - Adding new records or extending an old record
  - And not enough room in the CI to complete the operation
- CI split may trigger a CA split
- Splits generally degrades performance
- Specify freespace to reduce CI & CA splits

MAINTEC

ESDS

- Similar to Sequential File
- Sequenced by the order in which data is entered/loaded
- New Records are added at the end only (chronological order)
- Supports both Fixed and Variable formats
- Contains only CLUSTER & DATA components
- Only sequential access in Batch Cobol Programs
- Random access is supported in on-line applications (CICS) using Relative Byte Address (RBA)
- Alternate Index is supported in on-line applications (CICS)
- NO primary index

MAINTEC

RRDS

- Has only CLUSTER and DATA components
- Records are stored in numbered, fixed length slots
- Each slot is given a number 'Relative Record Number (RRN)'
- VSAM determines the number of slots by
  - Size of CI
  - Length of Record
- Records can be deleted physically
- Empty slots are filled up with new records without shifting existing records
- No primary Index or Alternate Index
- Supports Fixed and Variable formats
- RRN cannot be changed

MAINTEC

KSDS

- Has all three components of VSAM (CLUSTER, INDEX and DATA)
- Key sequenced
- Primary key should be
  - Unique
  - Same position in every record
  - Is not split (has to be contiguous)
- Records can be deleted physically
- Primary key cannot be changed
- Allows Alternate Index
- Has all the access methods
  - Sequential
  - Random
  - Dynamic (SKIP sequencial)

MAINTEC

VSAM dataset choice

	KSDS	ESDS	RRDS
Less DASD		✓	
On-line Direct Access	✓		✓
Batch Sequential		✓	
Alternate Index	✓		
Static data			✓
Ease of programming and maintenance	✓		

MAINTEC

Working with VSAM Datasets

Access Method Services (AMS)
IDCAMS

MAINTEC

IDCAMS

- Normally executed in Batch
- Always has the following JCL structure

```
//STEP010 EXEC PGM=IDCAMS
//SYSPRINT DD SYSOUT=*
//SYSIN DD*
<IDCAMS command from col.2  to  col.72>
/*
//*
```

MAINTEC

Basic IDCAMS Commands

- DEFINE (Cluster, Alternate Index etc.)
- BUILDINDEX (Alternate Index)
- REPRO (Cluster)
- LISTCAT (Catalog Entries)
- IMPORT / EXPORT (Cluster)
- VERIFY (Cluster)

MAINTEC

Basic define command syntax for cluster

```
DEFINE CLUSTER                -
(NAME (XIND.NLT.CLUSTER)      -
CYLINDERS (5 1)               -
    VOLUMES (WOR01)           -
    <Data set type>           -
)
```

{`-' is the continuation character and are normally aligned}

MAINTEC

Define Command (Contd..)

- NAME is a positional keyword parameter and must be coded first.
- Other keywords can be placed anywhere
- `+' sign is used for continuation within a field

- e.g.

```
DEFINE CLUSTER        -
(NAME (XIND.   +
NLT.CLUSTER))
```

MAINTEC

Record size parameter

- Syntax : RECORDSIZE (Average Maximum)
- Optional :Default : 4086
- Average & Maximum are same for Fixed length records
  - DEFINE CLUSTER -
  - (NAME(XIND.NLT.CLUSTER) -
  - CYLINDERS (5 1) -
  - VOLUMES (WORK01) -
  - RECORDSIZE (120 124) -
  - INDEXED)

MAINTEC

KEYS Parameter

- SYNTAX : KEYS (length offset)
- Used for KSDS only
- Optional : Default Value : Length=64 & Offset=0
  - DEFINE CLUSTER -
  - (NAME (XIND.NLT.CLUSTER) -
  - CYLINDER (5 1) -
  - VOLUMES (WORK01) -
  - RECORDSIZE (120 124) -
  - KEYS(8 0) -
  - INDEXED)

MAINTEC

Dataset type parameters

- KSDS : INDEXED or IXD
- ESDS : NONINDEXED or NIXD
- RRDS : NUMBERED
- LDS : LINEAR

MAINTEC

Data & index components

Required when installation default names are to be overridden

```
DEFINE CLUSTER                                          -
   (NAME (XIND.NLT.CLUSTER)                             -
   CYLINDER (5  1) VOLUMES (WORK01) -
   RECORDSIZE (120   124) KEYS (8   0) INDEXED)         -
DATA                                                    -
   (NAME(XIND.NLT.CLUSTER.DATA))                        -
INDEX                                                   -
   (NAME(XIND.NLT.CLUSTER.INDEX))     -
```

MAINTEC

VOLUMES

- Can specify different volumes for
  - Data component
  - Index component

MAINTEC

MORE AMS COMMANDS

MAINTEC

REPRO

- All purpose load and backup utility command
- Can be used against empty / loaded VSAM file with another VSAM file or sequential file
- Much easier to use than IEBGENER
- Can be used against all four types of VSAM datasets

MAINTEC

REPRO (Contd..)

REPRO INDATASET (DSN) or INFILE (DD1)
OUTDATASET (DSN) or OUTFILE(DD2)

SKIP (count)
COUNT (count)

FROMKEY
FROMADDRESS
FROMNUMBER

TOKEY
TOADDRESS
TONUMBER
REUSE/REPLACE

MAINTEC

Repro (Contd..)

- INFILE or INDATASET parameter is mandatory, similarly OUTFILE or OUTDATASET is mandatory
- All other parameters are optional
- SKIP specifies number of input records to skip before beginning to copy
- COUNT specifies number of output records to copy

MAINTEC

Repro (Contd..)

- REUSE parameter
  - Can be used only if the VSAM dataset was originally defined with REUSE option
  - Has the effect of logically deleting records before loading
- REPLACE parameter
  - Replaces the records for which primary keys are matching between input and output records
  - If not specified, the matching key records are untouched
  - If the target is ESDS the records are appended and REPLACE is inappropriate.

MAINTEC

Repro - Example

```
//REPRO EXEC PGM=IDCAMS
//SYSPRINT DD SYSOUT=*
//DD2 DD DSN=XIND.NLT.CLUSTER.BACKUP,DISP=SHR
//DD1 DD DSN=XIND.NLT.CLUSTER, DISP=SHR
// SYSIN DD*
REPRO INFILE (DD1) OUTFILE (DD2) REUSE
/ *
```

MAINTEC

EXPORT/IMPORT

- Used for backup and recovery
- Catalog information also is exported along with the data, unlike REPRO
- DFSMS classes are preserved
- Cluster deletion and redefinition are not necessary during the import
- Can be easily ported to other systems

MAINTEC

EXPORT/IMPORT(Contd..)

- Disadvantages
  - The EXPORTED file not reusable until it is imported
  - Slower than REPRO

MAINTEC

VERIFY

- Syntax : VERIFY FILE (<ddname>)

  VERIFY DATASET(<datasetname>)
  - Can be issued from TSO or from a JCL
  - Verifies the catalog HURBA (High Used Relative Byte Address) field and stores the true values from the control block HURBA field.
  - Should be used against cluster name only and not against data or index components
  - Used to rectify some of the problems due to data corruption

MAINTEC

LISTCAT

Used to list the contents of a master or user catalog

LISTCAT CATALOG (name)

ENTRIES (name - of - entries)

LEVEL (generic-leve-names)

Other parms:

ALL	NAME
HISTORY	VOLUME
ALLOCATION	NOT USABLE
CREATION (days)	EXPIRATION

OUTFILE (dd name)

MAINTEC

DELETE

- DELETE <object name> (parameters)
- Example : DELETE XIND.NLT.CLUSTER
  - All the parameters are optional
  - Deletes all subordinate objects such as AIX, Path

MAINTEC

Some Common DELETE Parameters

- ERASE / NO ERASE : ERASE writes binary zeroes after deletion
- PURGE / NO PURGE : PURGE allows deletion even though expiration date is still due
- ALTERNATE INDEX Or AIX : Deletes only Alternate index of the cluster
- PATH : Requests only path name to be deleted.
- FORCE / NO FORCE : FORCE deletes the dataset even if it is not empty

MAINTEC

Thank you !