

Ficheros en MVS (III) : VSAM

Para terminar el tema ficheros en MVS, este mes se dedica esta sección a estudiar la organización VSAM, la cual da servicio a un tipo de ficheros muy importante, nacido en los años 60, antes de la aparición de los sistemas gestores de Bases de datos, pero que a pesar de su edad no solo mantiene totalmente su vigencia, sino que sigue jugando un papel muy importante en cualquier GRAN SISTEMA.

1.- Introducción:

En las entregas anteriores se han enumerado los distintos métodos de almacenamiento y acceso a la información bajo el S.O. MVS de IBM, y también se hizo una introducción a la memoria virtual, por lo que el presente artículo se limita simplemente a la organización VSAM (Virtual Storage Access Method).

En esta organización dejan de tener importancia los términos de **cilindro**, **cabeza** o **pista** vistos en los artículos anteriores, pero en cambio aparecen otros conceptos que sirven para referenciar un registro, estos son los de **RBA** (Relative Byte Address) e **IC** (intervalo de control).

VSAM es una organización de datos que admite registros de longitud fija y variable. Realmente este método está pensado para optimizar el almacenamiento liberando al programador de la gestión que requiere el uso de registros de longitud variable al tiempo que gestiona el bloqueo y desbloqueo de registros calculando el factor de bloqueo a fin de aprovechar al máximo el espacio del dispositivo.

2.- Organización de los ficheros VSAM:

Existen tres tipos de ficheros:

- **ESDS** (Entry Sequence Data Set): Equivalentes a ficheros secuenciales
- **KSDS** (Key Sequence Data Set): en secuencia de clave. Este tipo de fichero es muy parecido al de organización secuencial indexada, teniendo la información dividida en dos partes: Área de índices y Área de datos.
- **RRDS** (Relative Records Data Set): equivalentes a ficheros con organización directa.

3.- Estructura física de ficheros VSAM:

Como se acaba de decir, VSAM libera al informático de hacer consideraciones sobre cilindros y cabezas (direcciones normales en disco) para dejar que el programador piense que el espacio de disco donde se graban los registros es continuo, comenzando en la dirección cero. Este espacio en disco, que puede abarcar uno o varios volúmenes, se denomina **DATA SPACE** y se define mediante la etiqueta correspondiente en la VTOC ya que al fin y al cabo es un fichero más del sistema.

En el caso de que por su tamaño abarcara varios volúmenes estos deberán ser del mismo tipo de dispositivo ya que considera todo su espacio como homogéneo en cuanto a características. Y, como ya se dijo en el primer artículo, cada tipo de disco tiene sus propias características, tales como longitud de pista, número de pistas por cilindros, etc. Por otra parte, y como es lógico, en un mismo dispositivo pueden existir uno o varios DATA SPACES junto con cualquier otro tipo de organización de datos.

4.- Conceptos específicos para VSAM

- **Registro Lógico:** Es la unidad básica de datos. Es decir, es la información que será tratada por el informático dentro del programa. Admite registros de longitud fija y variable.

- **Registro Almacenado:** Es un registro lógico más un campo de control que se denomina **RDF** (Record Definition Field, o campo de definición de registro). Este campo de control, RDF, consta de tres octetos (Figura 1):

VSAM: Registro Almacenado

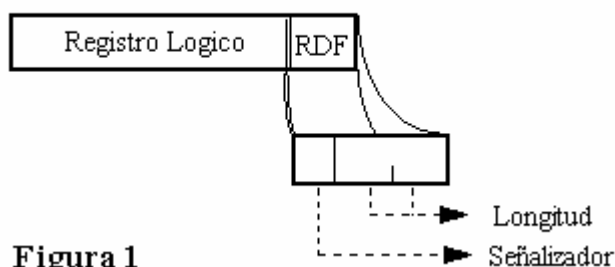


Figura 1

- byte 1: contiene indicadores de control, teniendo significado los bits 1 y 4.

Valores y significado de:

- bit_1 = 0 indica que no sigue más información de control relacionada con el RDF
- bit_1 = 1 indica que sigue más información de control relacionada con el RDF.
- bit_4 = 0 indica que el descriptor del registro es único (solo hay un registro con esa longitud)
- bit_4 = 1 indica que es un descriptor de cuenta, es decir indica el número de registros que tienen la misma longitud.

La figura 3, que se comenta posteriormente, muestra el uso de este campo de control.

Intervalo de control: El espacio total ocupado por un fichero VSAM está dividido en áreas que se denominan intervalos de control. VSAM almacena los registros de datos y la información de control que los describe en estas áreas. (figura 2)

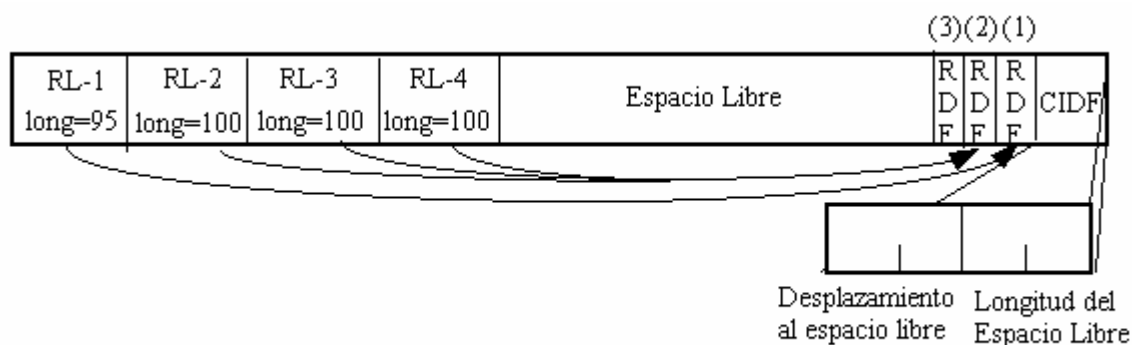


Figura 2 VSAM: Intervalo de control

Un registro de datos, junto con la información de control es lo que configura el **Registro Almacenado**. Todo intervalo de control lleva adicionado un campo de control que se denomina **CIDF** (Control Interval Description Field)

El intervalo de control puede variar de un fichero a otro, pero dentro de un fichero la longitud del intervalo de control permanece constante, es decir, a lo largo de un fichero la longitud de los intervalos de control no varia. VSAM elige esta longitud teniendo en cuenta lo siguiente:

- debe ser múltiplo de 512
- longitud de los registros de datos
- tipo de DAAD (Dispositivo de Almacenamiento de Acceso Directo)
- Memoria disponible
- Tamaño de las IOAREAS facilitado en la creación del fichero.

CIDF: es el campo de definición del intervalo de control.

Tiene 4 octetos de longitud, de los cuales, los 2 primeros indican el desplazamiento en octetos desde el principio del intervalo de control a las áreas de espacio libre; y los dos siguientes indican el tamaño del área de espacio libre contenida en el intervalo de control, pues VSAM intenta dejar espacio libre en los intervalos de control para las adiciones.

Ejemplo:

La figura 3 muestra el contenido de los RDF que acompañan a los registros lógicos grabados en el intervalo de control de la figura 2.

Byte	-----1-----	-----2-----	-----3-----
bit	0123-4567	0123-4567	0123-4567
RDF	(1)	0000-0000	0000-0000
	(2)	0100-1000	0000-0000
	(3)	0000-0000	0000-0011
		0000-0000	0110-0100

Figura 3 : RDFs de la figura 2

Como el registro lógico 1 va seguido por un registro de distinta longitud, utiliza solo un RDF. Como los registros lógicos 2, 3 y 4 tienen la misma longitud, se utilizan 2 RDF para describir a los 3 registros lógicos, con lo que se ahorra espacio. El primer RDF tiene un 1 en el bit 4 para indicar que la información contenida en los dos bytes siguientes especifican, en binario, el numero de registros que tienen esa longitud (11 (binario)=3(decimal)). El segundo RDF contiene 0 en los bits 1 y 4 pues no sigue mas información de control; y en los dos bytes finales contiene la longitud del registro (1100100=100).

Área de control: Recibe esta denominación cada uno de los conjuntos formados por un numero entero de intervalos de control. Este numero lo fija el propio VSAM, siendo como mínimo de tres.

Un área de control ocupa un numero entero de pistas, generalmente llegando a un cilindro, pero nunca mayor de un cilindro. Así, en el caso de elegir 100 intervalos de control por área, los primeros 100 intervalos constituyen la primera área de control, los 100 siguientes la segunda, etc.

Por áreas de control, VSAM reserva intervalos de control libres para adiciones y reorganizaciones. (**Figura 4**)

Cluster: En VSAM recibe el nombre de cluster el **data space** , el cual contiene:
 en KSDS : las áreas de datos y las áreas de índices
 en ESDS : el área de datos
 en RRDS : el área de datos.

RBA : dirección de byte relativa a origen de fichero

Para direccionar registros, VSAM no tiene en cuenta la ubicación física del registro en el dispositivo (concepto cabeza cilindro), sino que VSAM direcciona mediante el desplazamiento en octetos del registro respecto al comienzo del espacio reservado al fichero. Este desplazamiento se denomina **RBA**. (Relative Byte Address)

Para el calculo de los RBA, VSAM considera tanto registros de datos, como espacios libres y campos de control (**figura. 4**).

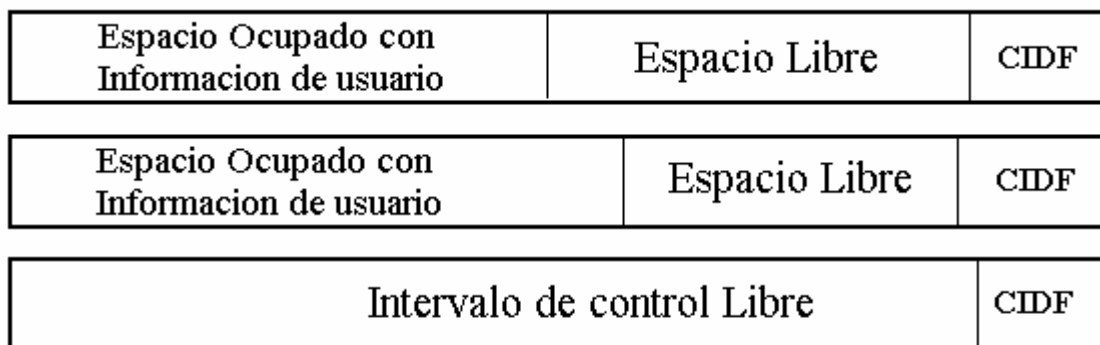


Figura 4

5.- Tipos de ficheros VSAM:

- ESDS - Fichero en secuencia de entrada.

En este tipo de ficheros, los registros se graban en el mismo orden en el que van llegando, es decir, uno detrás de otro, sin dejar huecos libres. Los registros nunca se intercalan, y tampoco se podrán ni acortar ni alargar. Las adiciones solo podrán hacerse al final del fichero por ser una organización compacta.

Cuando se quiere borrar un registro, se marcará físicamente pero se eliminará físicamente cuando se realice la posterior reorganización.

Los RBA no varían. Los registros siempre tienen el mismo RBA.

El acceso a los registros será secuencial, aunque también podrá hacerse al azar conociendo su RBA, para lo cual basta con leer previamente el registro y recoger así el RBA, pues VSAM no mantiene ningún tipo de índice con este tipo de fichero.

- KSDS : Fichero en secuencia de clave.

Se corresponde con la organización secuencial indexada.

Este tipo de fichero tiene almacenado sus registros en secuencia de clave llevando un orden lógico. Para poder realizar su cometido, un fichero de este tipo va siempre acompañado de un índice que se usa para localizar los registros de datos. Los ficheros KSDS están formados por dos partes: Índice y Espacio de datos, y no tendrá área de overflow porque VSAM reserva áreas libres en los intervalos de control.

a.- Índices

Los índices tienen por finalidad el relacionar las claves con la posición de los registros de datos en el fichero. Será VSAM el que genere y mantenga este conjunto de índices. Los índices se agrupan a su vez manteniendo la estructura de un fichero, con intervalos y áreas de control. Cada intervalo de control recibe el nombre de **registro**, y contiene varias **entradas**. (Figura 5)

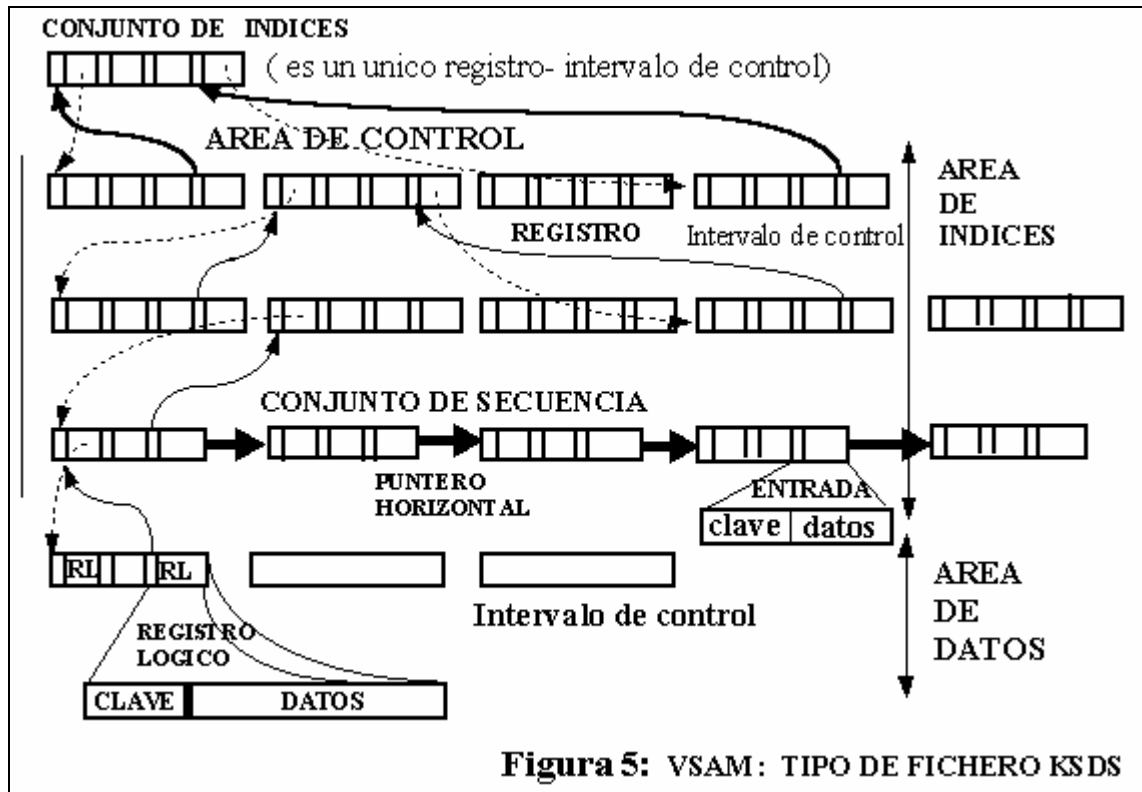


Figura 5: VSAM: TIPO DE FICHERO KSDS

El índice puede tener uno o mas niveles. Cada uno de los niveles está compuesto por un conjunto de registros que contendrán una serie de entradas con la dirección de los registros de índices de jerarquía inmediatamente inferior. El índice de mas bajo nivel recibe el nombre de Conjunto de secuencia (sequenced-set). Cada entrada de este conjunto de secuencia apunta a los intervalos de control que contienen los registros de datos en el fichero.

Cada **entrada** de un registro índice contiene:

- La clave mas alta de cada uno de los registros del nivel inferior de índices; y
- Un puntero que se podría llamar vertical, para indicar la dirección de ese registro, es decir, es parecido al campo de enlace de la organización secuencial indexada.

Esta organización se aprecia en todos los niveles, hasta llegar al **conjunto de secuencia** que es el nivel mas bajo de la jerarquía de índices.

Cada uno de los **registros** del *conjunto de secuencia* remite a un **área de control**; y Cada una de las **entradas** de un registro del *conjunto de secuencia*, remite a un **intervalo de control**.

Es decir, si en un fichero, un *área de control* tiene 8 intervalos de control, un registro del *conjunto de secuencia* tendrá 8 entradas.

- Los *intervalos de control* libres también están direccionados en el *conjunto de secuencia*.

El *conjunto de secuencia* relaciona a todos los registros por punteros horizontales para poder realizar un proceso secuencial. tal y como indica la figura 5.

- VSAM actualiza y reorganiza los índices cuando lo cree necesario.

-b.- ÁREA de DATOS:

Los registros con datos de usuario, se graban en intervalos de control siguiendo una secuencia lógica. Los intervalos de control se agrupan en áreas de control, y estas estarán referenciadas de alguna manera en cada uno de los intervalos de control del conjunto de secuencia como ya se ha visto.

-c.- ÁREA DE ESPACIO LIBRE

En los ficheros VSAM-KSDS se puede dejar espacio libre para añadir, alargar o acortar registros de datos. Estas áreas de espacio libre pueden ser de dos tipos:

- Espacio libre al final de los intervalos de control utilizados para datos.
- Intervalos de control libres. Varios en cada área de control.

El tamaño y el modo de dejar el espacio libre es facilitado por el propio programador en el momento de la creación, siendo VSAM el que se encargue de utilizarlo adecuadamente.

d.- División de los intervalos de control :

Cuando al añadir o alargar un registro se exceda de la capacidad de un intervalo de control, VSAM utiliza un intervalo de control libre que se encuentre dentro de la correspondiente área de control y reparte con él el contenido del intervalo de control excedido, quedando ambos con cierta cantidad de espacio libre.

En el caso de que el área de control se quede sin intervalos de control libres, VSAM generara una nueva área de control al final del fichero y repartirá con ella el contenido de esta.

Cuando se reparte áreas de control, el área de datos no lleva ni secuencia lógica ni física, la secuencia lógica será llevada por el conjunto de secuencia

e.- Procesos en los que puede participar un fichero VSAM en secuencia de claves:

- Carga: La carga se suele realizar de forma secuencial. Se dejará espacio libre dentro de los intervalos de control y también intervalos de control libres dentro de las áreas de control.

- Adiciones: Con las adiciones se pueden dar los tres casos siguientes.

- Se podrán hacer adiciones de registro sin que se pierda la secuencia lógica y física, gracias a los espacios libres dejados en los intervalos de control en la creación.

- Cuando se produce la división de los intervalos de control dentro de la misma área de control, puede perderse la secuencia física, pero nunca se pedrera la lógica gracias a los punteros del conjunto de secuencia.

- Adiciones repartiendo áreas de control: entonces se dividen las áreas de control y se pierde la secuencia física pero nunca la lógica.

- Supresiones: Los registros se borran físicamente y se condensarán todas las informaciones que se encuentran en el intervalo de control, de forma que el espacio ocupado por ese registro que es baja pasa a ser adicionado al espacio libre actualizándose los índices correspondientes.

Como consecuencia de estas adiciones, divisiones y supresiones, los RBA de los registros se alteraran.

- Recuperación: Puede ser secuencial o al azar.

- Secuencial: VSAM solo utiliza el conjunto de secuencia, utilizando los punteros horizontales para poder mantener la secuencia.

- Azar: VSAM va siguiendo todos los niveles de índices mediante los punteros verticales para localizar el intervalo de control del registro pedido.

RRDS : *Fichero de registro relativo*

Esta organizacion podría asimilarse con la Organizacion Directa Relativa vista anteriormente.

Este fichero consta de una serie de huecos o registros, todos de igual longitud y que de alguna forma se encuentran numerados correlativamente desde el primero de ellos. Se podría comparar con una tabla o array, donde cada registro de datos ocupa un único hueco y se carga o recupera con respecto a su posición relativa a origen de fichero.

Un fichero RRDS podrá actualizar registros, borrarlos o adicionarlos en los huecos libres.

Hay veces que se puede obtener a partir de una clave la dirección relativa a origen aplicando una función de direccionamiento.

6.- Catalogo VSAM:

VSAM utiliza un catalogo para mantener la información de todos los volúmenes y todos los ficheros VSAM.

Este catalogo es un fichero especial , siendo su organizacion KSDS.

Gracias al catalogo VSAM, esta organizacion presenta una función nueva con las siguientes características:

- Gestión automática de espacio DAAD:

VSAM conoce gracias al catalogo el espacio disponible en los volúmenes, con esto se proporciona el espacio libre y el ocupado. Siempre que se vaya a crear un fichero, VSAM selecciona automáticamente un área libre para ese fichero, y siempre que queramos pedir información sobre un determinado fichero, habrá que proporcionar el nombre de ese fichero al catalogo.

- Protección de datos: VSAM es el que elige el área para un fichero nuevo. No será por tanto posible su destrucción por error ya que es el propio VSAM el que protege esos ficheros.

- Datos Privados(confidencialidad) Se definen unas palabras claves para poder localizar los datos de un fichero, por lo que cuando se desee recuperar información sobre un determinado fichero habrá que proporcionar la password o palabra de paso.

- Estadísticas: VSAM mantiene en el catalogo un conjunto de estadísticas de la utilización de las informaciones al tiempo que almacena en el catalogo toda la información asociada con las características físicas del dispositivo en el que reside la información.

7.- AMS (access Mode Service)

El AMS es un conjunto de programas de servicio, o utilidades que permiten:

- Definir, imprimir y copiar los ficheros VSAM
- Reorganizar ficheros VSAM
- Añadir, alterar, borrar e imprimir entradas en el catalogo.
- Convertir ficheros ESDS en KSDS.

8.- Consideraciones finales

Este artículo, como se dice en la presentación, cierra la miniserie dedicada a los ficheros. Todas las organizaciones y métodos vistos tienen su importancia pues son todos los que existen y además, todos están vivos. Unos, los más complejos, son los utilizados por los Gestores de Bases de Datos (SGDBs) ya que al tiempo que gestionan la complejidad establecen una barrera casi infranqueable ante los datos que mantienen (organización directa). Otros, más sencillos, secuencial, secuencial Indexados y particionados son los comúnmente usados por los programadores. Por último, la Organización VSAM es usada tanto por unos como por otros, citando a modo de ejemplo un Gestor basado en ficheros VSAM, el SGDB: DB1 de la casa SAPIENS.

9.- Utilidad:

La utilidad que acompaña este mes al artículo, consiste en una *clist* que lista del catálogo de usuario todas las entradas que cuelgan del nivel que se especifique, pasándolas a un fichero para su posterior tratamiento.

La utilidad está contenida en los siguientes ficheros:

- **UTILIDAD.JMP** : Hardcopys de tres paneles en tiempo de ejecución de la utilidad.
- El primero presenta un menú de utilidades, desde el que se puede acceder a la utilidad que acompaña al artículo.
- El segundo presenta el panel que aparece al elegir la opción 6 del menú anterior. En él se puede especificar el nivel cualificador de las entradas a listar del catálogo.
- El tercero presenta el resultado obtenido.
- **CATALOGO.CLS** : CLIST central de la utilidad. Básicamente el esquema de este programa de comandos consiste en pedir mediante un panel el nivel cualificador a listar. Ejecuta el comando **LISCAT** de TSO con el parámetro **LEVEL** para obtener todas las entradas del catálogo que cuelgan del prefijo especificado y pasar esa salida a un fichero, de forma parecida a como se comentó en los artículos de esta sección correspondientes a los números 11 y 12 de esta revista. Una vez hecho esto, el problema se reduce a tratar el fichero obtenido como una tabla presentándole mediante un panel.
- **CATALOGO.PAN** : Definición del panel invocado por la *clist* para pedir el nivel cualificador.
- **ENTORNO0.PAN** : Definición del panel utilizado para presentar los resultados.

- **LIBRO.DOC** : El autor de este artículo ha escrito un libro, inédito hasta la fecha, titulado “**TSO para desarrolladores**” en el cual describe los conceptos básicos de MVS y ya más exhaustivamente, con ejemplos concretos, los tres apartados que debería conocer todo desarrollador en una gran instalación JCL, CLIST e ISPF. El fichero citado contiene una presentación de dicho libro.

- José María Peco Palacios -