

**CATALOGADO**



Buenos Aires, 29 de septiembre de 1976.

Al Señor  
Interventor del  
Consejo Federal de Inversiones  
Cnel. (Re.) Julio César Medeiros  
S. / D.

REF.: Expediente 6765

De mi mayor consideración:

Por la presente elevo, para su consideración,  
el Primer Informe Parcial sobre el expediente 6765 "Banco de Datos".

Sin otro particular, saludo al señor Interven  
tor con distinguida consideración.

A handwritten signature in dark ink, appearing to be "Lucio Librado Arce".

E.C. Lucio Librado Arce

Adj.: Original estudio y dos copias.

21094

INFORME N° 1.

BANCO DE DATOS DEL CONSEJO FEDERAL  
DE INVERSIONES

C.C. Lucio Librado Arce

B.34

0

V. 151

A 28

## INDICE

### DESCRIPCION DE PROCESOS

Carga de Series.....	1
Emisión de Catálogos.....	3
Almacenamiento de Arboles.....	3

### PROCEGRAMAS

Compilación de Descriptores.....	4
Compilación de Observaciones.....	5
Carga de Series.....	6
Emisión de Catálogos.....	7
Carga de un árbol geográfico.....	8
Carga de un árbol temático.....	8

### DESCRIPCION DE ARCHIVOS

Descriptores Compilados.....	9
Observaciones Compiladas.....	10
Palabras Claves.....	11
Geográfico .....	12
Temático.....	14
Indice.....	15
Series.....	17
Carga de Series.....	18
Carga de índices.....	19
Bajas de Series Compiladas.....	20

## CARGA DE SERIES

El sistema de carga de series se compondrá de tres procesos: el proceso de compilación de descriptores, el proceso de compilación de observaciones y el proceso de carga de series compiladas.

Proceso de compilación de descriptores: Está diseñado con el fin de pre-procesar los descriptores de las series, esto es detectar los errores de los mismos y formar una cola de espera de descriptores depurados y ya preparados para una carga instantánea, mientras se pre-procesan sus observaciones. Este proceso comprende cuatro etapas:

### 1-SORT

Una vez que se ha grabado el formulario N° 1 del banco de datos, diseñado por el Sector de Información, se realiza en el mismo un ordenamiento creciente de los números de serie que existen en el archivo.

### 2-COMPILACION

Comprende las verificaciones sobre código geográfico, códigos temáticos y palabras claves. El objetivo de esta etapa es emitir un listado de los descriptores incluyendo los errores encontrados en el mismo e interpretación de códigos y palabras claves a efectos de verificación. Los descriptores que pasen el control de esta etapa serán insertados en una cola de espera para ser cargados con sus observaciones en el banco. Ellos se grabarán en un archivo en una forma intermedia entre su formulario original y la representación de los descriptores de las series almacenadas.

/C1

### 3-MERGE

Intercalación de los descriptores compilados en la etapa anterior y el archivo de cola de espera con los descriptores ya compilados existentes.

### 4-ELIMINACION

Procede de la detección de errores que hayan pasado la etapa 2. Con la consecuente eliminación de los descriptores del archivo de descriptores compilados almacenados en cola de espera.

Proceso de compilación de observaciones: Tiene la misma estructura que el proceso anterior. Su diferencia es el preproceso de lote de observaciones de un descriptor dado, formando un lote de espera de onjuntos de observaciones o lote de observaciones compiladas.

Proceso de carga de series compiladas: Comprende la carga de las series propiamente dichas en 5 etapas.

### 1-MERGE

Se realiza la unión de los descriptores y lotes de observaciones almacenados en sus respectivos lotes de espera.

### 2-ELIMINACION

Se eliminan de las colas de espera las series que se aparearon en la etapa anterior.

### 3-CARGA

Se incluyen en el banco de datos las series proveniente de la etapa 1. Asimismo, se cargan nuevas palabras claves que se incluyan en estas series. Por último se genera un archivo intermedio el número de serie, el código geográfico, el código temático y el código palabra clave que pertenecía a esa serie (archivo de carga de índice).

### 4-SORT

El archivo de carga de índice se sortea por código geográfico,

/c2

código temático, código de palabra clave y número de serie.

#### 5-CARGA DE INDICE

El archivo preparado en la etapa anterior es procesado de la siguiente manera: si en el archivo de índices ya existía la misma tri-upla (de código geográfico, código temático, código de palabra clave) se agregan al registro del archivo catálogo todas las series que pertenezca a esta tri-upla. En el caso que no aparezca en el archivo de índices se almacena la misma y luego se almacenan en un nuevo registro de catálogo los números de series que corresponden a la misma.

#### EMISION DE CATALOGOS

Comprende la impresión de las series almacenadas en el banco de datos desagregada por 1 a 3 de sus atributos, geográficos y/o temático y/o palabra clave.

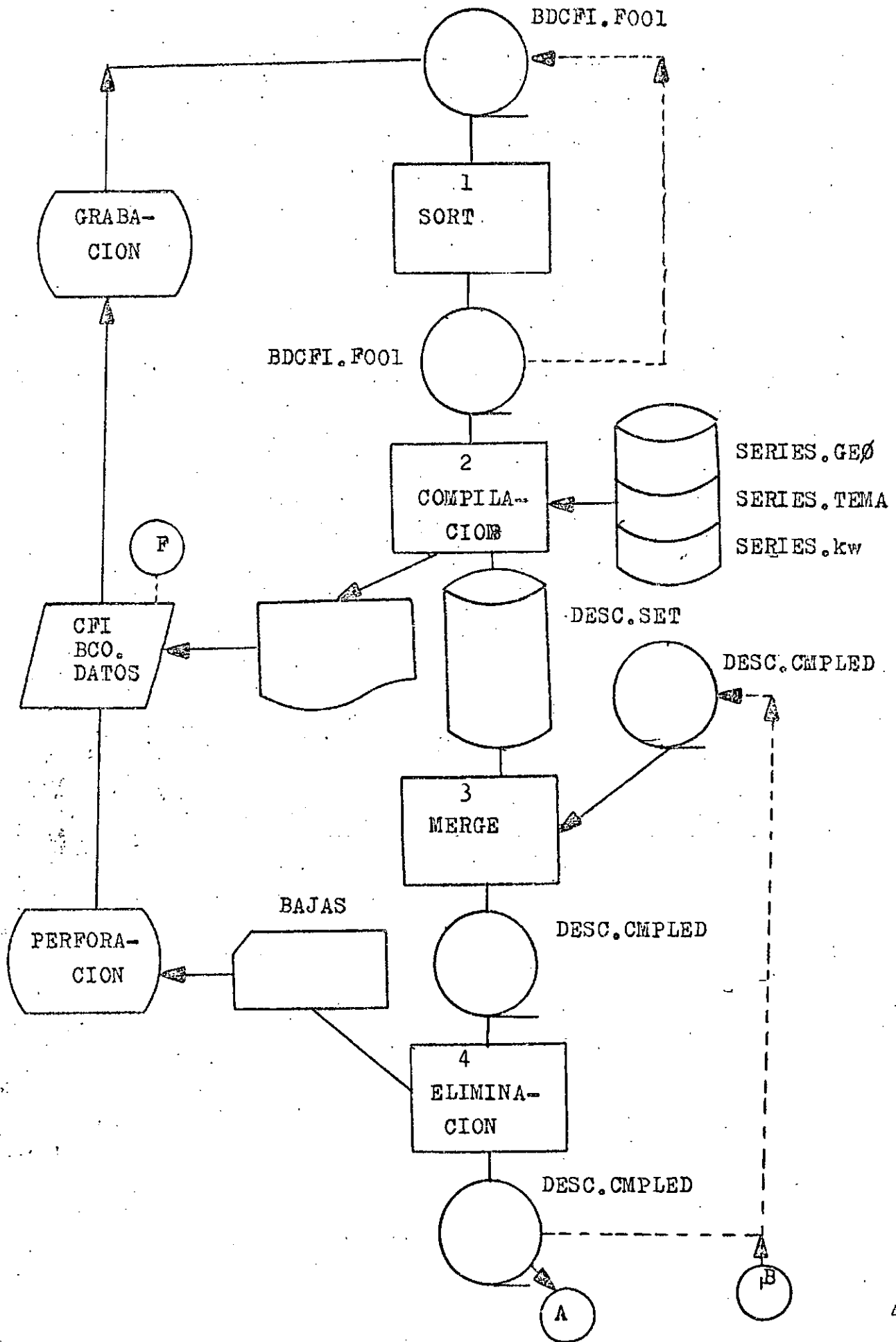
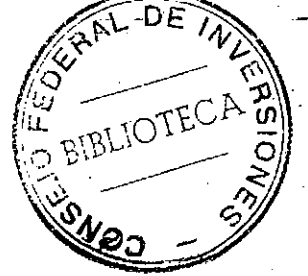
Dada la indicación mediante una tarjeta de control del escalar, bi-upla o tri-upla seleccionado, el sistema sorteará y/o unirá o no las tri-uplas almacenadas en el archivo índice y luego con los números de series almacenados en el registro catálogo imprimirán las series que correspondan a la condición establecida.

Si seleccionamos (geográfico, temático) se imprimirán para cada código geográfico, dentro de éste para cada código temático las series que cumplen con esta desagregación.

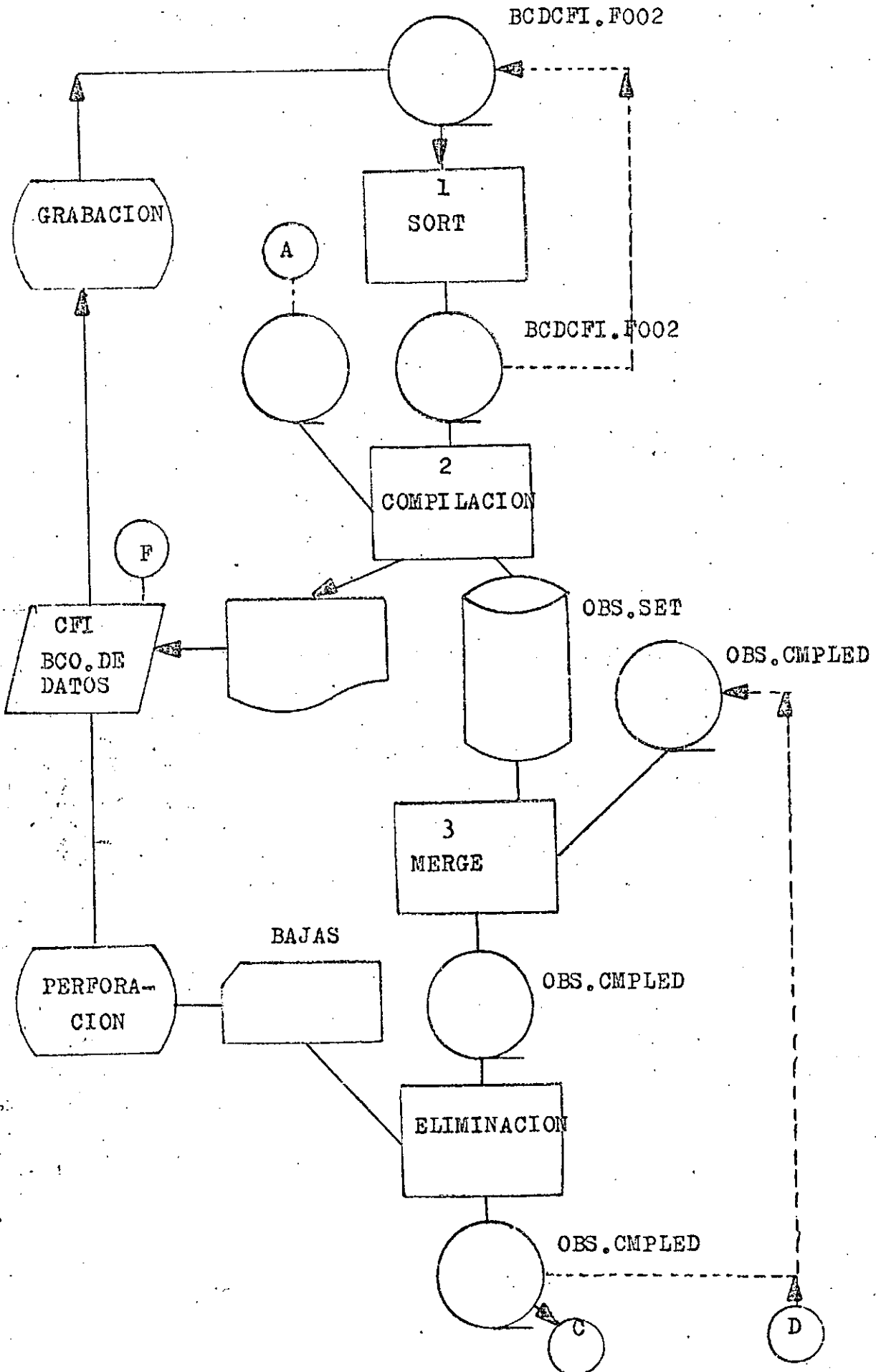
#### ALMACENAMIENTO DE UN ARBOL

Dado un lote de datos en tarjetas especificando el código geográfico o temático y el nombre del código el sistema almacenará el árbol geográfico o temático en su correspondiente archivo.

PROCESO: COMPILACION DE DESCRIPTORES

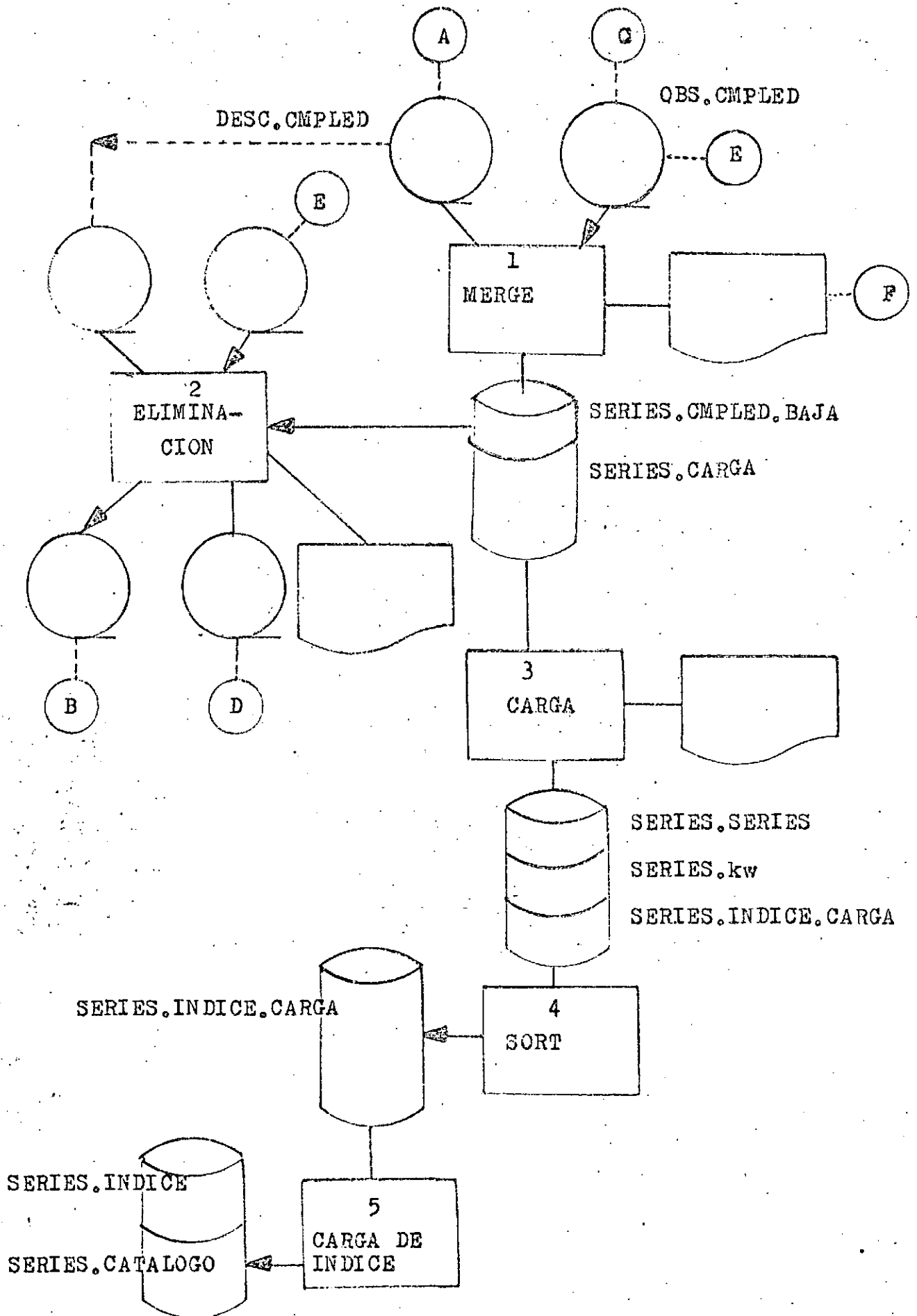


PROCESO: COMPILACION DE OBSERVACIONES

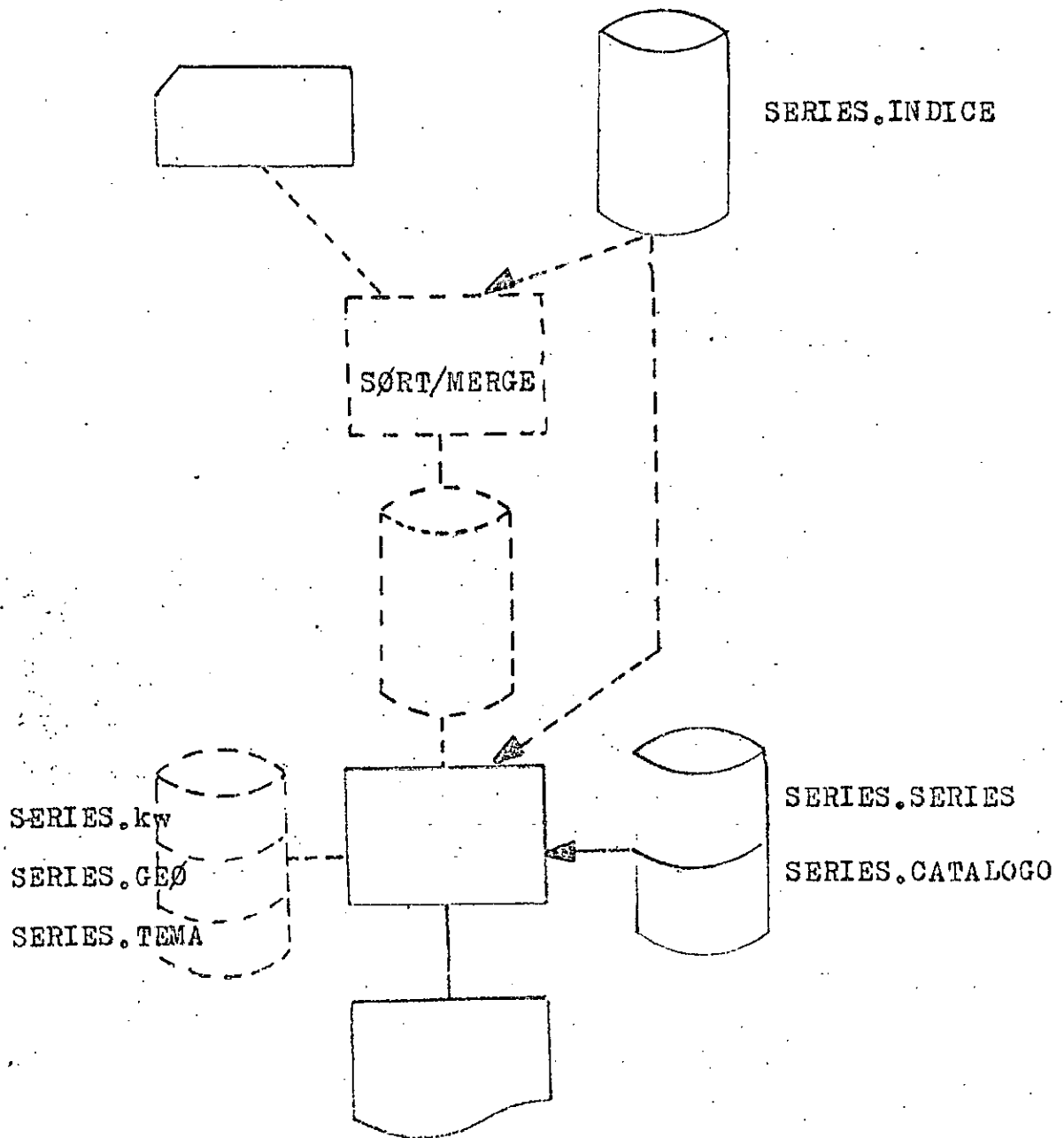




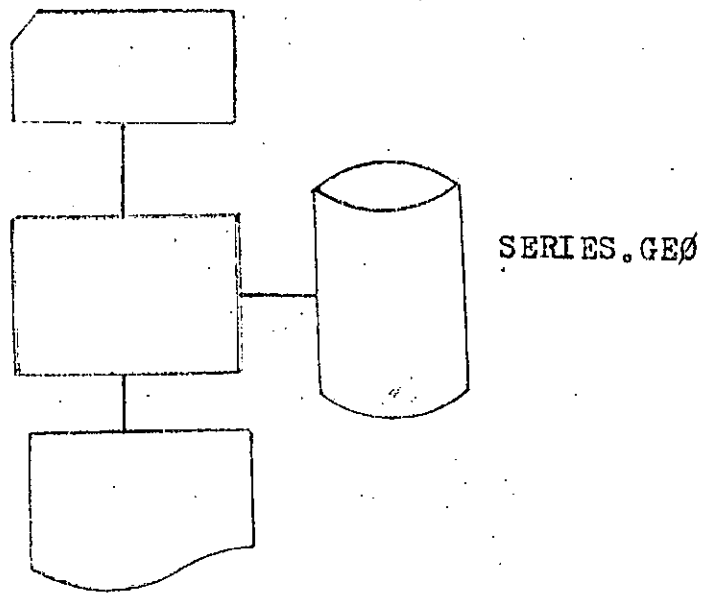
PROCESO: CARGA DE SERIES



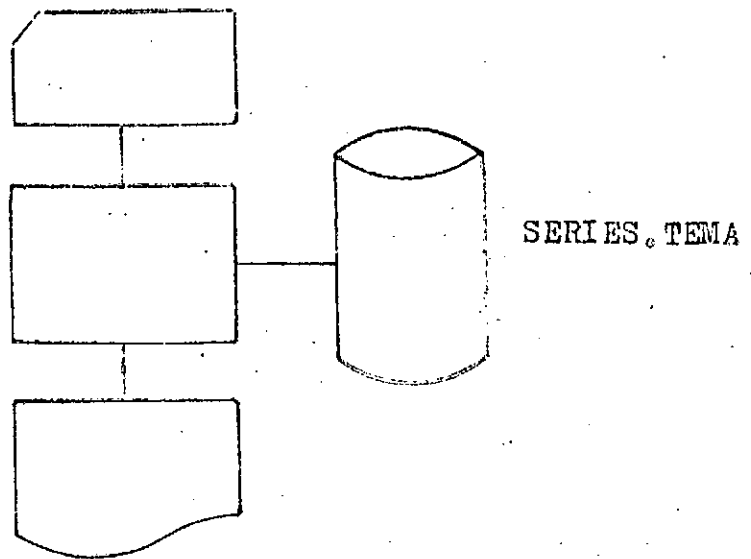
EMISION DE CATALOGOS



CARGA DE UN ARBOL GEOGRAFICO



CARGA DE UN ARBOL TEMATICO



## ARCHIVO DE DESCRIPTORES COMPILADOS

El contenido de cada registro es el siguiente:

- .número de serie (INTEGER)
- .título (LOGICAL \* 1, 80 bytes)
- .código geográfico (INTEGER)
- .exponente y unidad (INTEGER \* 2)
- .periodicidad de la serie (INTEGER \* 2)
- .fecha de referencia (INTEGER \* 2)
- .año/mes inicial (INTEGER \* 2)
- .año/mes final (INTEGER \* 2)
- .existencia de comentarios (INTEGER \* 2)
- .cantidad de decimales (INTEGER \* 2)
- .cantidad de temas (INTEGER \* 2)
- .3 temas ( 3 \* INTEGER)
- .cantidad de palabras claves (INTEGER \* 2)
- .cantidad de palabras claves ya existentes en el archivo de palabras claves (INTEGER \* 2)
- .4 códigos de palabras claves (4 \* INTEGER \* 2)
- .cantidad de palabras claves a incluir (INTEGER \* 2)
- .4 longitudes (4 \* INTEGER \* 2)
- .palabras claves (LOGICAL \* 1, 70)

Soporte unidad: 2400, nombre: DESC.CMPLED, longitud del registro lógico 206 bytes, el factor de bloqueo se determinará de acuerdo a la capacidad de memoria de la mayor extensión del programa que utilice este archivo y la longitud del buffer.

Con la misma especificación se construye el archivo intermedio DESC.SET en un soporte 2314.

## ARCHIVO DE OBSERVACIONES COMPILADAS

El contenido de cada registro es el siguiente:

- . número de serie (INTEGER)
- . cantidad de observaciones (INTEGER)
- . 260 observaciones (REAL \* 8)

Soporte:Unidad:2400, nombre:OBS.CMPLED, longitud del registro lógico: 2092 bytes, el factor de bloqueo se determinará de acuerdo a la capacidad de memoria de la mayor extensión del programa que utilice este archivo y la longitud del buffer. Con la misma especificación se construye el archivo intermedio OBS.SET en un soporte: 2314.

## ARCHIVO DE PALABRAS CLAVES

Cada registro del archivo de palabras claves coincide con un nivel del hash-code generator, es decir, que cada registro contendrá los sinonimos de cada nivel. Este archivo se divide en dos partes: el índice y el archivo de palabras claves propriamente dicho.

### Registro Índice:

Por cada nivel del hash-code generator existirá un registro de índice. Este contendrá:

- primer byte disponible en el registro de palabras claves (INTEGER \* 2)
- total de palabras claves en este nivel (INTEGER \* 2)
- vector que contiene el primer byte del registro de palabras claves donde comienza cada una de ellas. Las palabras claves se encontrarán almacenadas en orden alfabético creciente (INTEGER \* 2, ...)
- vector cada posición representa a la i-ésima palabra clave almacenada. Se utiliza como pointer a una 'PILA' (INTEGER \* 2, ...)
- vector PILA cada posición representa el valor asignado a una palabra clave, su contenido indicará la posición real que ocupa a partir de la primera palabra clave (INTEGER \* 2, ...)

### Registro de Palabras Claves:

- cada registro pertenecerá a uno y sólo un nivel, su longitud es de 7294 bytes. Se prevee para cada nivel 5 registros.
- longitudes y cantidad de registros índices provendrá del estudio estadístico de las palabras claves y la selección del hash-code generator.
- soporte unidad 2314. Estado inicial y '00'.

## ARCHIVO GEOGRAFICO

La información geográfica será desagregada en país, provincia, departamento y localidad. Estructuralmente representa un árbol de 4 niveles. Según el formulario diseñado por el Sector de Información del Consejo Federal de Inversiones, las claves para cada nivel serán numéricas y la longitud en dígitos decimales es la siguiente:

<u>NIVEL</u>	<u>DIGITOS</u>
1	2
2	2
3	3
4	3

De esta manera el árbol será recorrido en la forma pre-orden (nodo padre, sub-árbol izquierdo, sub-árbol derecho), en otras palabras, de nodos superiores a nodos inferiores y en el mismo nivel de izquierda a derecha. Por lo tanto es conveniente almacenarlo de la misma manera, logrando con ello en un barrido secuencial recorrerlo en forma preorden.

Existirá una relación directa entre código y dirección de almacenamiento.

El archivo estará dividido en tres partes:

- .a) descripción de árbol;
- .b) nodos; y
- .c) nombre del nodo.

.a) DESCRIPCION DEL ARBOL (TRACK-1):

- . dirección donde se almacena el nodo raíz (vector INTEGER, 99 elementos)
- . cantidad de elementos reservados para el nivel 2
- . cantidad de elementos reservados para el nivel 3
- . cantidad de elementos reservados para el nivel 3  
(vector INTEGER \* 2, 99 \* 3 elementos)

.b) NODOS (TRACK 2-40):

/

Cada nodo contendrá la longitud y la dirección donde se ha alm  
cenado el nombre del nodo. De la siguiente manera:

. número de registro y longitud del nombre (INTEGER \* 2)

. número del byte donde comienza el nombre (INTEGER \* 2)

Cada registro contendrá 1823 nodos y su longitud será 7292 by-  
tes.

.c) NOMBRE (TRACK 41-180):

Cada registro tendrá 7294 bytes y dentro del mismo se almacena  
rán los nombres en forma secuencial.

Soporte: unidad 2314, estado inicial X'00', recuperación: ac-  
ceso directo. Nombre: Series. GEØ



### ARCHIVO TEMATICO

La información temática tiene estructura arborea de 6 niveles. El Sector de Información del Consejo Federal de Inversiones ha previsto una clave numérica cuya descripción es la siguiente (Ver Formulario BDCFI001).

<u>NIVEL</u>	<u>DIGITOS</u>
1	1
2	1
3	2
4	2
5	2
6	2

El tratamiento de cada nodo y su almacenamiento se realizará de la misma forma que el caso geográfico, su diferencia es la cantidad de niveles.

#### a) DESCRIPCIÓN DEL ARBOL (TRACK-1):

.dirección donde se almacena el nodo raíz (vector INTEGER, 9 elementos)

.cantidad de elementos reservado para el nivel 2

.cantidad de elementos reservado para el nivel 3

.cantidad de elementos reservado para el nivel 4

.cantidad de elementos reservado para el nivel 5

.cantidad de elementos reservado para el nivel 6

(vector INTEGER \* 2, 9 \* 6 elementos)

#### b) NODOS (TRACK 2 -40)

#### c) NOMBRE (TRACK 41-180)

Soporte: unidad 2314, estado inicial X'00', recuperación: acceso directo. Nombre: Series. Temas.

## ARCHIVO INDICE

Cada elemento de este archivo contendrá:

- .código geográfico (INTEGER);
- .código temático (INTEGER);
- .código de palabra clave (valor)(INTEGER \* 2);
- .número del registro del archivo catalogo que contiene los números de serie que cumplen la condición de contener ese código geográfico, ese tema y esa palabra clave (INTEGER \* 2).

Los registros tendrán una longitud de 7294 bytes. El tratamiento del archivo será pseudo-indexado, el primer registro contendrá el índice, existirán registros (pistas) que contengan los elementos descriptos y registros de overflow para los registros.

Registro Indíce (Registro Número 1):

- .cantidad de registros (INTEGER \* 2);
- .cantidad de registros alocados (INTEGER \* 2);
- .cantidad de registros de overflow (INTEGER \* 2);
- .cantidad de registros de overflow alocados (INTEGER \* 2);
- .vector: contiene la clave más alta que contiene cada registro. Cada elemento posee 12 bytes de longitud;
- .vector: contiene, si existe overflow de registro, el número de registro de overflow donde se ha comenzado a almacenar las claves que han provocado esa condición;
- .vector: contiene, si existe overflow de registro, la clave más alta insertada en él.

//

Registros: 607 elementos ocupa 7284 bytes.

Registros de overflow. Sus elementos contienen:

- .código geográfico (INTEGER);
- .código temático (INTEGER);
- .código de palabra clave (INTEGER \* 2);
- .número del registro del archivo catalogo (INTEGER \* 2);
- .registro y elemento inferior (INTEGER \* 2);
- .registro y elemento superior (INTEGER \* 2)

Cada registro de overflow contendrá 455 elementos que ocupan 7280 bytes.

<u>REGISTRO N°</u>	<u>USO</u>
1	INDICE
2-50	REGISTRO
50-60	REGISTRO OVERFLOW

Soporte unidad 2314, estado inicial x'00'

## ARCHIVO DE SERIES

El contenido de cada registro es el siguiente:

- .título (80 bytes);
- .longitud del título (INTEGER \* 2);
- .exponente y unidad (INTEGER \* 2);
- .fuente productora (INTEGER);
- .mes/año inicial (INTEGER \* 2);
- .mes/año final (INTEGER \* 2);
- .fecha de referencia (INTEGER \* 2);
- .código geográfico (INTEGER);
- .cantidad de temas (INTEGER \* 2);
- .3 temas (3 \* INTEGER);
- .cantidad de palabras claves (INTEGER \* 2);
- .4 palabras claves (4 \* INTEGER \* 2);
- .continuación/comentarios (INTEGER \* 2);
- .registro de continuación (INTEGER \* 2);
- .40 años de observaciones/192 meses.

La longitud del registro de series anuales será de 450 bytes. Cada pista tendrá 13 registro y se han alocado 5200 registros. La longitud de registros de series mensuales es de 1693 bytes. Cada pista contiene cuatro registros se han alocado 3200. Ocupan 20 y 40 cilindros respectivamente, los números de series anuales son 1 a 5199 y para las mensuales es 10001 a 13199. Su estado inicial es x'00'. Soporte unidad: 2314.

ARCHIVO DE CARGA DE SERIES



El contenido de cada registro es el siguiente:

- .número de serie(INTEGER)
- .título(LOGICAL \* 1, 80 bytes)
- .código geográfico (INTEGER)
- .exponente y unidad (INTEGER \* 2)
- .periodicidad de la serie (INTEGER \* 2)
- .fecha de referencia (INTEGER \* 2)
- .año/mes inicial (INTEGER \* 2)
- .año/mes final (INTEGER \* 2)
- .existencia de comentarios (INTEGER \* 2)
- .cantidad de temas (INTEGER \* 2)
- .3 temas (3 \* INTEGER)
- .cantidad de palabras claves (INTEGER \* 2)
- .cantidad de palabras claves ya existentes en el archivo de palabras claves (INTEGER \* 2)
- .4 códigos de palabras claves (4 \* INTEGER \* 2)
- .cantidad de palabras claves a incluir (INTEGER \* 2)
- .4 longitudes (4 \* INTEGER \* 2)
- .palabras claves (LOGICAL \* 1, 70)
- .260 observaciones (REAL \* 8) máxima

Soporte:Unidad:2314, nombre:SERIES.CARGA, longitud del registro lógico y físico:2298 bytes máxima, registros por pista:3, disposición:transitoria, espacio alocado:10 cilindros, recuperación:secuencial, formato:VS.

## ARCHIVO DE CARGA DE INDICES

Cada elemento del registro contiene:

- . código geográfico (INTEGER)
- . código temático (INTEGER)
- . código de palabra clave (INTEGER \* 2)
- . número de serie (INTEGER)

Cada registro contiene 455 elementos, su longitud es de 7280 bytes.

Soporte.Unidad:2314, nombre:SERIES.INDICE.CARGA, disposición:transitoria, recuperación:secuencial.

Formato:VS.Registro por pista:1, disposición:transitoria, espacio a alocar: 10 cilindros.

ARCHIVO DE BAJAS DE SERIES COMPILADAS

Contiene los números de serie compiladas que se cargarán y por lo tanto serán dadas de baja como series compiladas.

El número de serie se almacena como INTEGER, cada registro contiene 1822 elementos. Soporte: Unidad: 2314, longitud del registro: 7288 bytes, Formato: VS, registros por pista: 1, espacio asignado: 10 pistas, disposición: transitoria.