



Voto Electrónico

Etapa 1

Organización de Datos – Cátedra Servetto 2do cuatrimestre 2011

21/10/2011

Alfredo Scoppa	89149
Cecilia khalil	
Fernando Romera Ferrio	88406
Juan Manuel Romera Ferrio	88405



Tabla de contenidos

Manual de Usuario	3
Organización de registros.....	4
Organización Balanceada	5
Organización Directa	6
Funciones	6
Cubetas.....	6
Primitivas implementadas.....	7
Archivos.....	7
Diagrama de clases.....	7



Universidad de Buenos Aires
Facultad de Ingeniería
75.06 Organización de Datos
Segundo Cuatrimestre 2011

Manual de Usuario



Universidad de Buenos Aires
Facultad de Ingeniería
75.06 Organización de Datos
Segundo Cuatrimestre 2011

Organización de registros



Universidad de Buenos Aires
Facultad de Ingeniería
75.06 Organización de Datos
Segundo Cuatrimestre 2011

Organización Balanceada



Organización Directa

Se implementó dispersión estática. Los desbordes se resuelven usando direccionamiento abierto y doble dispersión. Se usa una función para obtener la dirección base y otra para dar los saltos en caso de desborde.

Funciones

Función principal: Resto de División o Módulo. La clave se divide entre el número de direcciones. Para asegurarnos una dispersión eficiente nos aseguramos que el tamaño de la tabla sea un número primo.

Función secundaria: También es Resto de División o Módulo solo que esta función es alimentada con el resultado de una función de hashing previa.

Se pensó el hash para trabajar con registros de tamaño variable que se almacenaran en cubetas con capacidad para varios registros.

Cubetas

Estructura de una cubeta:

Overflow (1 byte)	FreeSpace (2 bytes)	Count (2 bytes)	bytes
-------------------	---------------------	-----------------	-------

El tamaño de las cubetas es parametrizable como también el tipo de dato que se almacena en la tabla. De acuerdo al tipo de registro a insertar se tomaron distintos múltiplos de 512 para determinar el tamaño de la cubeta.

Para determinar la cantidad óptima de cubetas en diferentes escenarios se consideraron diferentes parámetros:

C: número de registros por cubeta

R: número total de registros almacenados

DE: Densidad de empaquetamiento la cual se fijó en 0.7

C y R son parámetros estimativos que deben determinarse conociendo los registros a insertar.



Así se despeja el número óptimo de cubetas D , como $D = R/(C \cdot DE)$.

Finalmente se verifica que D sea primo y en caso contrario se toma el primo más cercano.

Primitivas implementadas

Se implementaron las siguientes primitivas:

- Creación: Se crea un nuevo archivo y se calcula la cantidad de cubetas necesarias.
- Carga: Se abre un archivo y se determina la cantidad de cubetas existentes.
- Inserción: Se intenta insertar un registro y lanza una excepción en caso de que ya exista.
- Recuperación: Se intenta recuperar un registro y lanza una excepción en caso de que no exista.
- Actualización: Se intenta actualizar un registro y lanza una excepción en caso de que no exista.
- Borrado: Se intenta borrar un registro y lanza una excepción en caso de que no exista.

Archivos

Se utiliza un solo archivo que es el que contiene la totalidad de las cubetas.

Diagrama de clases

A continuación se presenta un diagrama de clases resumido de las clases involucradas en esta organización.

