

**Universidad de Buenos Aires**

**Facultad de Ingeniería**

**Departamento de Informática**

***Organización de Datos (75.06)***

**Voto Electrónico**

Cuatrimestre y año: 2do Cuatrimestre 2011

Docente a cargo del TP:Nicolás Pablo Fernández Theillet

Grupo: Lamas

Fecha de Entrega:

Integrantes: 2011-10-22

|  |  |  |
| --- | --- | --- |
| ***Padrón*** | ***Nombre*** | ***Email*** |
| 91187 | Gonzalez Durand, Juan Manuel | jmanuel.gonzalez.durand@gmail.com |
| 90762 | Ostrowsky, Gabriel | gaby.osto@gmail.com |
| 90728 | Schenkelman, Damián | damian.schenkelman@gmail.com |
| 91045 | Torrado, Alejandro | aletorrado@gmail.com |
| 90884 | Zamudio, Gonzalo | ahogadosderazon@gmail.com |

Índice

[Diseño 3](#_Toc304593810)

[Distrito 3](#_Toc304593811)

# Diseño

A continuación se detallan las decisiones de diseño tomadas para la organización de los archivos de las diferentes entidades del problema. Para cada una de ellas se detalla:

* La organización que se utilizara para guardar los datos de relacionados con la misma.
* Las razones por la que se decidió dicha organización.
* Características particulares dentro de la organización elegida (ej: función de dispersión para el Hash).
* Estructura de los registros utilizados (tanto en índices como para guardar datos).
* Estructura de los bloques utilizados.
* Una imagen o gráfico que sirve para ilustrar el diseño elegido.

## Distrito

**Organización**

Archivo de bloque con Registros de datos de longitud variable (RLV) guardados en un Hash de dispersión fija con zona de desborde. La zona de desborde será un archivo secuencial de RLV organizado por bloques para poder direccionar a otra cubeta.

**Razones**

* Ninguna funcionalidad de la aplicación requiere recorrer distritos secuencialmente una vez que encontramos a uno.
* La cantidad de distritos (al menos en un escenario real) es fija, no aumenta. Solo se harán operaciones de consulta una vez que la carga inicial fue realizada.
  + De esta forma es posible diseñar una función de dispersión así como el tamaño de los buckets para poder recuperar un distrito en la primera lectura a disco (2da en caso de que la función de hash no sea lo suficientemente efectiva y se produzca un desborde).

**Características Particulares**

* Función de Hash: A determinar.
* Tamaño Bucket/Bloque: A determinar.
* Factor empaquetamiento (registros ocupados / espacio disponible): A determinar.
* El bloque se considera desbordado cuando se algún distrito no pueda insertarse por tener tamaño mayor al espacio libre disponible. Sin embargo, si otro distrito es insertado y puede ser acomodado en el archivo de datos, el mismo puede ser dado de alta allí.

**Registros**

Se utilizara solo un registro de datos, de la forma:

* **R**(longitudRegistro,(distrito)i)

**Bloques**

Los bloques del archivo de datos y de desborde serán de la forma:

* **B**(bloqueDesborde, espacioOcupado, (Distrito)\*)

**Ilustración**

A continuación se muestra el posible contenido de un bloque:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Bloque Desborde | Espacio Ocupado | Distrito 1 | | Distrito 2 | | Distrito 3 | | Distrito 4 | |
| **4** | 42 | 5 | Chaco | 9 | Catamarca | 7 | Córdoba | 6 | Chubut |

Espacio ocupado: 4 (bloque desborde) + 4 (espacio ocupado) + 8 (longitud registro) + (5 + 9 +7 + 6) = 42