



**Universidad Nacional Abierta**  
**Vicerrectorado Académico**  
**Área de Ingeniería**

**Carrera Ingeniería de Sistemas**

### **Trabajo Sustitutivo de Pruebas**

Asignatura: Procesamiento de Datos

Código: 330

Nombre del Estudiante: Angel Leon

Cédula de Identidad: 29772294

Centro Local / Unidad de Apoyo: Maracay/Aragua

Correo electrónico: [angelleonarmas23@gmail.com](mailto:angelleonarmas23@gmail.com)

Teléfono celular: 0426-332 10 74

Carrera: 237

Lapso: 2024-2

### **Resultados de Corrección**

| <b>OBJ N°</b> | <b>1</b> | <b>2</b> | <b>3</b> | <b>4</b> |
|---------------|----------|----------|----------|----------|
|               |          |          |          |          |

**Mod. I Unid.1 Obj. 1**

1. El procesamiento de datos implica transformar datos sin procesar en información valiosa para las empresas, lo que incluye recopilarlos, organizarlos, limpiarlos, verificarlos, analizarlos y convertirlos en formatos legibles, como gráficos o documentos.

**Por ejemplo**, en el departamento de una empresa de producción, consideran sus datos de información, sumamente valiosos y por esta razón desean cumplir a cabalidad con las operaciones que involucran el procesamiento de datos. Basado en lo anterior, describa las operaciones que involucran el procesamiento de datos.

**1. Recopilación:**

Primero se identifican y seleccionan las fuentes de donde se obtendrán los datos o **fuentes de datos**, estas pueden ser bases de datos, archivos, sensores, encuestas, etc. y se utilizan diversas técnicas o **métodos de recolección** para obtener los datos, como la extracción de bases de datos, el rastreo de redes sociales, la realización de encuestas, el uso de sensores, etc.

**2. Organización:**

Se define cómo se organizarán los datos, es decir, en qué formato o **estructura** se almacenarán (tablas, archivos planos, bases de datos relacionales, etc.), luego se establecen los campos y atributos que contendrá cada registro de datos y se crea un **esquema**.

**3. Limpieza:**

Se **Identificarán** o detectarán errores en los datos, como valores faltantes, duplicados, inconsistencias, formatos incorrectos, para luego **corregir** aplicando técnicas, como la imputación de valores faltantes, la eliminación de duplicados, la normalización de datos, etc.

**4. Verificación:**

Se verifica y **validara** que los datos sean coherentes y cumplan con las reglas de negocio establecidas, con una **garantía** de calidad que los datos sean precisos, completos y confiables.

## 5. Análisis:

Se realiza un análisis inicial de los datos para identificar patrones, tendencias y relaciones, se construirán modelos estadísticos o de **machine learning** para **extraer** conocimiento de los datos, así se representan los datos de forma gráfica para facilitar su comprensión e interpretación.

## 6. Conversión a formatos legibles:

Se crean **informes** personalizados para presentar los resultados del análisis de datos y se generaran **gráficos, diagramas y otras visualizaciones** para comunicar los hallazgos de manera efectiva.

**Mod. I Unid. 2 Obj. 2**

2. La creación de un archivo secuencial es un proceso sencillo, que implica almacenar datos uno detrás de otro en un orden predefinido. Imagina que estamos digitalizando los datos de proveedores una tienda llamada “SuperOfertas”.

Supongamos que tenemos tres proveedores con la siguiente información:

**Frutas Gutiérrez** Nombre: Antonio Gutiérrez, Teléfono: 607454545

**Hortalizas del Sur**, Nombre: Guillermo Morales, Teléfono: 652854874

**Azucarera Ortiz**, Nombre: Rodrigo Méndez, Teléfono: 22525885

Basado en lo anterior, realiza un esquema que represente el archivo secuencial en orden ascendente.

**Esquema del Archivo Secuencial**

| Posición | Campo                | Tipo de Dato | Tamaño | Descripción  |
|----------|----------------------|--------------|--------|--|
| 1        | Nombre del Proveedor | Alfanumérico | 30     | Nombre completo de la Empresa proveedora           |
| 31       | Teléfono             | Numérico     | 9      | Número de teléfono de la Empresa proveedora        |
| 40       | Nombre del Dueño     | Alfanumérico | 30     | Nombre completo del Dueño de la Empresa proveedora |

**Ordenamiento Ascendente:**

Los registros se ordenarán alfabéticamente según el nombre del proveedor.

Ejemplo de Archivo Secuencial:

|                    |                             |
|--------------------|-----------------------------|
| Azucarera Ortiz    | 22525885 Rodrigo Méndez     |
| Frutas Gutiérrez   | 607454545 Antonio Gutiérrez |
| Hortalizas del Sur | 652854874 Guillermo Morales |

### Explicación del Esquema:

- **Posición:** Indica la posición inicial de cada campo dentro del registro.
- **Campo:** Describe el tipo de información que se almacenará en ese campo.
- **Tipo de Dato:** Especifica si el dato es alfanumérico o numérico.
- **Tamaño:** Define el número máximo de caracteres que puede contener el campo.
- **Descripción:** Ofrece una explicación breve del contenido del campo.

**Mod. I, Unid.3 Obj. 3**

3. Un archivo multillaves permite acceder a los registros utilizando múltiples claves, lo que facilita la búsqueda y recuperación de datos desde diferentes perspectivas.

Entre sus características están:

1. Llave Primaria.
2. Llaves Secundarias
3. Índices
4. Acceso Rápido.

S tiene, por ejemplo, la siguiente tabla acerca de un archivo principal con datos de empleados, y se quiere acceder a través del campo *N° de control* y *Nombre*.

| Indice | N° de Control | Nombre |
|--------|---------------|--------|
| 1      | 37            | Juan   |
| 2      | 29            | Ana    |
| 3      | 16            | Pedro  |
| 4      | 35            | Jose   |
| 5      | 15            | Carmen |
| 6      | 34            | Luis   |
| 7      | 28            | Hector |
| 8      | 33            | Carlos |
| 9      | 12            | Jesus  |
| 10     | 36            | Simon  |
| 11     | 27            | Mari   |
| 12     | 31            | Cris   |
| 13     | 18            | Ruben  |
| 14     | 26            | Omar   |

**Archivo Principal**

De acuerdo a la información dada, se le pide:

- a) Explicar cada una de las características de los archivos multillaves.

- b) Con los datos del archivo principal, construya e indique el contenido de las tablas para las llaves secundarias y acceder por *N° de control* y por Nombre.

### Archivos Multillaves:

Los archivos multillaves son una **estructura** de datos que permite acceder a los **registros** de un archivo a través de múltiples **campos o claves**. Esta característica los hace muy útiles cuando se necesita buscar información de diversas maneras.

### Características de los Archivos Multillaves:

- **Múltiples Claves:** A diferencia de los archivos secuenciales, que típicamente tienen una sola clave principal para ordenar los registros, los archivos multillaves permiten indexar los registros por varios campos. Esto significa que puedes buscar un registro no solo por su clave principal, sino también por otros atributos como nombre, fecha, categoría, etc.
- **Acceso Directo:** Gracias a los índices creados para cada clave, es posible acceder directamente a los registros que cumplen con un criterio de búsqueda específico, sin necesidad de recorrer todo el archivo secuencialmente. Esto mejora significativamente el rendimiento de las consultas.
- **Flexibilidad:** Los archivos multillaves son altamente flexibles y pueden adaptarse a diferentes tipos de consultas. Puedes crear índices para cualquier campo que sea relevante para tus aplicaciones.
- **Complejidad:** La implementación y mantenimiento de archivos multillaves es más compleja que la de los archivos secuenciales. Se requiere una estructura de datos más sofisticada para gestionar los índices y las relaciones entre ellos.
- **Espacio de Almacenamiento:** Los índices ocupan espacio adicional en el disco, lo que puede ser un factor a considerar en aplicaciones con grandes volúmenes de datos.

**Estructura Básica de un Archivo Multillave:**

- **Archivo Principal:** Contiene los registros de datos en sí mismos.
- **Índices:** Cada índice está asociado a una clave y contiene una lista de punteros a los registros que tienen ese valor de clave. Los índices pueden ser de diferentes tipos, como índices B+ o índices hash.

**Construyendo los Índices****Índice por Número de Control**

El índice por Número de Control será una tabla donde cada fila tendrá dos columnas:

- **Número de Control:** El valor del número de control.
- **Puntero al Registro:** Un valor que indica la posición del registro en el archivo principal.

Dado que no se especifica cómo se almacenarán los registros en el archivo principal, asumiremos que cada registro tiene un número de identificación único (por ejemplo, 1, 2, 3, ...) y usaremos este número como puntero.

| Número de Control | Puntero al Registro |
|-------------------|---------------------|
| 16                | 3                   |
| 15                | 5                   |
| 12                | 9                   |
| 18                | 13                  |
| 26                | 14                  |
| 27                | 11                  |
| 28                | 7                   |
| 29                | 2                   |
| 31                | 12                  |
| 33                | 8                   |
| 34                | 6                   |
| 35                | 4                   |
| 36                | 10                  |
| 37                | 1                   |



### Índice por Nombre

El índice por Nombre también tendrá dos columnas:

- **Nombre:** El nombre del empleado.
- **Puntero al Registro:** Un valor que indica la posición del registro en el archivo principal.

| Nombre | Puntero al Registro |
|--------|---------------------|
| Ana    | 2                   |
| Carlos | 8                   |
| Carmen | 5                   |
| Cris   | 12                  |
| Hector | 7                   |
| Jesus  | 9                   |
| Jose   | 4                   |
| Juan   | 1                   |
| Luis   | 6                   |
| Mari   | 11                  |
| Omar   | 14                  |
| Pedro  | 3                   |
| Ruben  | 13                  |
| Simon  | 10                  |

**Mod. I, Unid.4 Obj. 4**

4. El gerente de comunicaciones de una empresa, al analizar el tráfico, identifica cuellos de botella y bajo rendimiento. Ante esta situación desea realizar un estudio de optimización donde se presenten las soluciones, donde el objetivo sería minimizar el tiempo, esfuerzo y recursos necesarios para intercambiar datos entre dos puntos de la red.

Basado en lo anterior, aplique sus conocimientos sobre comunicación y redes, para sugerir una solución de optimización, presentándola en un esquema, considerando los siguientes aspectos:

- **Evaluación de la Red**
- **Gestión del Tráfico**
- **Actualización de la Infraestructura**
- **Supervisión Continua**
- **Compresión y Almacenamiento en Caché**
- **Equilibrio de Carga**

**Propuesta de Optimización de la Red para Minimizar Tiempo, Esfuerzo y Recursos****Diagnóstico:**

El gerente de comunicaciones ha identificado cuellos de botella y bajo rendimiento en el tráfico de la red. Esto sugiere que existen limitaciones en la capacidad de la red para manejar la carga de datos, lo que resulta en retrasos en la transmisión y una experiencia de usuario subóptima.

**Objetivo:**

Minimizar el tiempo, esfuerzo y recursos necesarios para intercambiar datos entre dos puntos de la red.

**Solución Propuesta:**

A continuación, se presenta en considerando los aspectos mencionados:

## 1. Evaluación de la Red

- **Análisis de tráfico:** Identificar los patrones de tráfico, los tipos de datos que se transmiten y las horas pico de uso.
- **Identificación de cuellos de botella:** Determinar los puntos de la red donde se producen congestiones y retrasos.
- **Evaluación del hardware:** Verificar si los equipos de red (routers, switches, servidores) tienen la capacidad suficiente para manejar el tráfico actual y futuro.
- **Análisis de la topología:** Evaluar la estructura de la red y determinar si existen redundancias o ineficiencias.

## 2. Gestión del Tráfico

- **Priorización del tráfico:** Asignar mayor prioridad a los datos críticos y menos a los datos menos importantes.
- **Control de congestión:** Implementar mecanismos para evitar la congestión en la red, como el control de flujo y la calidad de servicio (QoS).
- **Optimización de rutas:** Seleccionar las rutas más eficientes para el tráfico entre los dos puntos de la red.

## 3. Actualización de la Infraestructura

- **Aumento de ancho de banda:** Incrementar la capacidad de transmisión de los enlaces de red.
- **Actualización de hardware:** Reemplazar equipos obsoletos por equipos más modernos y potentes.
- **Implementación de nuevas tecnologías:** Considerar tecnologías como la fibra óptica, Wi-Fi 6 o 5G para mejorar el rendimiento de la red.

## 4. Supervisión Continua

- **Monitoreo del rendimiento:** Implementar herramientas para monitorear el tráfico de la red, la utilización de los recursos y la disponibilidad de los servicios.

- **Detección de anomalías:** Configurar alertas para detectar cualquier problema o degradación en el rendimiento de la red.
- **Análisis de logs:** Revisar los registros de la red para identificar las causas de los problemas y realizar un análisis de raíz.

## 5. Compresión y Almacenamiento en Caché

- **Compresión de datos:** Reducir el tamaño de los paquetes de datos para aumentar la eficiencia de la transmisión.
- **Almacenamiento en caché:** Almacenar los datos más solicitados en caché para reducir el tiempo de acceso y disminuir la carga en la red.

## 6. Equilibrio de Carga

- **Distribución del tráfico:** Distribuir la carga de trabajo entre múltiples servidores para evitar sobrecargar un solo equipo.
- **Algoritmos de balanceo de carga:** Implementar algoritmos efectivos para distribuir el tráfico de manera equitativa.

## Conclusión:

La optimización de una red es un proceso continuo que requiere una evaluación constante y ajustes. Al implementar las medidas propuestas, la empresa podrá mejorar significativamente el rendimiento de su red, reducir los costos y garantizar una mejor experiencia para los usuarios.