

Trabajo Práctico 1

1.23: Índice

El índice es una estructura de datos que sirve como identificador único para las filas de una tabla en una base de datos. En palabras simples, es un puntero a los datos en una tabla.

1.24: Repositorio de datos

Es una colección centralizada de datos que se almacena en una base de datos o en un conjunto de bases de datos interconectadas. Este repositorio puede contener datos estructurados (como tablas relacionales), semiestructurados (como JSON o XML) o incluso no estructurados (como documentos o archivos multimedia). Su objetivo principal es servir como una fuente única de verdad (Single Source of Truth) para toda la organización.

2.23: Apache Thrift

Es un framework desarrollado por Apache que permite la creación de servicios escalables y eficientes para la comunicación entre diferentes lenguajes de programación. Thrift simplifica la creación de servicios distribuidos al proporcionar una forma de definir interfaces y tipos de datos que pueden ser utilizados por múltiples lenguajes, lo que lo hace ideal para sistemas heterogéneos donde diferentes componentes están escritos en diferentes lenguajes.

2.24: RDBMS dump files

Son archivos que contienen una copia de una base de datos en un momento específico. Estos archivos se utilizan comúnmente para realizar copias de seguridad, migrar datos o compartir información entre sistemas. Incluyen tanto el esquema de la base de datos (tablas, índices, restricciones, etc.) como los datos almacenados en las tablas.

3.23: RTMP

Es un protocolo de comunicación desarrollado por Adobe Systems para la transmisión de audio, video y datos a través de Internet. Se utiliza principalmente para la transmisión en vivo y la entrega de contenido multimedia bajo demanda. Aunque fue creado originalmente para trabajar con Adobe Flash, sigue siendo relevante en la actualidad debido a su baja latencia y eficiencia en la transmisión de datos.

3.24: RTSP

Es un protocolo de red diseñado para controlar la transmisión de medios en tiempo real, como audio y video. Se utiliza comúnmente en sistemas de videovigilancia, transmisiones en vivo y aplicaciones de streaming. A diferencia de los protocolos de entrega de contenido (como HTTP),

RTSP no transmite los datos directamente, sino que actúa como un controlador para iniciar, pausar, detener y gestionar la transmisión de medios.

4.2.5: **Amazon DynamoDB**

Es un servicio de base de datos NoSQL completamente gestionado y altamente escalable, ofrecido por AWS. Está diseñado para proporcionar un rendimiento rápido y predecible con una escalabilidad sin problemas, lo que lo hace ideal para aplicaciones que requieren baja latencia y alta disponibilidad, como aplicaciones web, móviles, juegos, IoT y más.

4.6.4: **Pentaho Data Integration**

También conocido como Kettle, es una herramienta de integración de datos open source desarrollada por Hitachi Vantara. Es ampliamente utilizada para la extracción, transformación y carga de datos (ETL). PDI permite a las organizaciones integrar datos de múltiples fuentes, transformarlos según sus necesidades y cargarlos en un destino final, como un data warehouse, una base de datos o un sistema de análisis.

4.10.3: **Nagios**

Es una herramienta de monitoreo de infraestructura de TI open source ampliamente utilizada para supervisar sistemas, redes y aplicaciones. Su objetivo principal es detectar y alertar sobre problemas en la infraestructura antes de que afecten a los usuarios finales.

5.9: **Amazon S3**

Es un servicio de almacenamiento de objetos altamente escalable y duradero ofrecido por AWS. Está diseñado para almacenar y recuperar cualquier cantidad de datos desde cualquier lugar en la web, lo que lo hace ideal para una amplia variedad de casos de uso, como almacenamiento de datos, respaldo, análisis, aplicaciones web, y más. S3 = Simple Storage Service.

5.36: **Asana**

Es una plataforma de gestión de proyectos y tareas diseñada para ayudar a equipos y organizaciones a planificar, organizar y realizar un seguimiento de su trabajo. Es una herramienta popular para la colaboración en equipo, ya que permite a los usuarios crear tareas, asignarlas a miembros del equipo, establecer plazos, y visualizar el progreso del trabajo en diferentes vistas, como listas, tableros, calendarios y diagramas de Gantt.

8.5: **Apache Spark**

Es un framework open source diseñado para procesar grandes volúmenes de datos de manera rápida y eficiente. Es una de las herramientas más populares en el ecosistema de Big Data debido a su velocidad, facilidad de uso y soporte para una amplia variedad de cargas de trabajo, como procesamiento por lotes, streaming, ML y análisis interactivo.

11.2: **Informatica Axon**

Es una plataforma de gestión de datos empresariales y gobernanza de datos diseñada para ayudar a las organizaciones a mejorar la calidad, confiabilidad y usabilidad de sus datos. Es parte del

ecosistema de herramientas de Informática, una de las empresas líderes en soluciones de integración y gestión de datos. Axon se centra en proporcionar una visión unificada y colaborativa de los datos en toda la organización, permitiendo a los usuarios empresariales y técnicos trabajar juntos para maximizar el valor de los datos.

14.2: **SQL**

Es un lenguaje de programación estándar utilizado para gestionar y manipular bases de datos relacionales. Es ampliamente utilizado para realizar operaciones como consultas, inserciones, actualizaciones y eliminaciones de datos, así como para definir y modificar la estructura de las bases de datos.

15.12: **Minería de Texto**

Es un proceso de análisis de grandes volúmenes de texto no estructurado para extraer información útil, patrones y conocimientos. Combina técnicas de Procesamiento del Lenguaje Natural (NLP), ML y análisis estadístico para transformar texto en datos estructurados que pueden ser analizados y utilizados en la toma de decisiones.