

Presentación del espacio curricular Base de Datos Distribuidas

En esta asignatura se exponen conceptos básicos en el campo de las bases de datos, desde el repaso de la definición y descripción del concepto base de datos hasta su implementación en un sistema comercial de gestión de bases de datos pasando por la presentación de las herramientas conceptuales necesarias para construir la base de datos, llegando a la definición de bases de datos distribuidas y su procesamiento y gestión.

Objetivos: los objetivos del espacio son:

Brindar conceptos avanzados de los elementos claves para el diseño, implementación y manipulación de Bases de Datos.

Específicamente

- adquirir capacidades para modelar y representar información persistente,*
- adquirir capacidades para discernir cuando es conveniente usar un sistema de manejo de bases de datos,*
- ejercitar el uso de modelo relacional, sus capacidades y limitaciones,*
- ejercitar el uso de lenguaje de consultas estándar como SQL,*
- adquirir conocimientos inherentes al manejo de transacciones,*
- adquirir conocimientos básicos del uso de bases de datos distribuidas,*
- conocer los distintos mecanismos de recuperación ante fallos,*
- adquirir conocimientos de las nuevas aplicaciones de bases de datos,*
- desarrollar capacidades para la presentación individual y colectiva de trabajos.*

Los temas que componen el contenido son:

I: CONCEPTOS BASICOS DE BASES DE DATOS.

II: DISEÑO DE BASES DE DATOS RELACIONALES.

III: MANEJO DE TRANSACCIONES.

IV: CONCURRENCIA DE BASES DE DATOS.

V: BASE DE DATOS DISTRIBUIDAS. Y RECUPERACIÓN ANTE FALLOS

VI: NUEVAS APLICACIONES DE BASE DE DATOS

Organización del espacio curricular:

Quisiéramos comentarles algunas cuestiones sobre el uso de plataforma de enseñanza virtual para poder dar comienzo a la cursada.

- *En cuanto a la casilla de correo, es conveniente que la revisen con regularidad, ya que toda la información que reciban sobre la asignatura se concentrará allí. También es recomendable que cuando necesiten comunicarse con un docente, lo hagan desde el correo.*

- *En el espacio de la asignatura encontrarán todo lo referente a la programación del curso: bibliografía, programa, avisos del profesor, listado de alumnos que toman el curso con sus respectivos correos, foros, etc.*

Metodología de cursada

Como es habitual encontrarán en la plataforma, los recursos correspondientes a cada nueva clase (páginas web, archivos, actividades, foros, etc.). Cuando “se suba” una nueva clase o recurso importante, se les estará avisando a través de un mensaje del/la docente a sus casillas de mail.

- *Los cursos tienen una extensión de un cuatrimestre, cuentan con una clase semanal que se sube en avisos profesor, y generalmente requieren al menos una intervención semanal por parte del alumno (prácticas, participación en debates, etc.). El foro abierto es el espacio recomendado para la comunicación entre alumnos y para dudas generales. Es un espacio más libre en cuanto a lo que pueden colgar, solicito que allí vuelquen todas las dudas así logramos compartir también las respuestas que pueden ser de utilidad para muchos/as. En cambio el espacio de debates es donde se concentran las participaciones específicas del curso y las obligatorias. Es conveniente que estén pendientes y revisen los foros, ya que el medio de comunicación, en la modalidad virtual, entre el docente y el alumno.*

Ante cualquier inquietud no duden en escribirnos. Si el asunto a atender requiere demora en averiguaciones, de todos modos les contestaré en lo inmediato ofreciéndoles las explicaciones pertinentes.

Les dejamos un saludo cordial, y les deseamos un buen comienzo!!.

Equipo de catedra

- *Ing. Altamirano Walter*
- *Lic. Bianchiotti Franco*

Algunos conceptos previos a modo de introducción.

Una Base de Datos Distribuidas (BDD) puede ser definida como una colección integrada de datos compartidos que están físicamente repartidos a lo largo de los nodos de una red de computadoras. Un DDBMS es el software necesario para manejar una BDD de manera que sea transparente para el usuario. Un DBMS centralizado es un sistema que maneja una BD simple, mientras que un DDBMS es un DBMS simple que maneja múltiples BD. El término global y local se utiliza, cuando se discute sobre DDBMS, para distinguir entre aspectos que se refieren al sitio simple (local) y aquello que se refiere al sistema como un todo (global). El modelo distribuido de datos hace posible la integración de BD heterogéneas proveyendo una independencia global del DBMS respecto de esquema conceptual. Además, es posible implementar una integración tal, que reúna varios modelos de datos, representado cada uno de ellos características propias de empresas diferentes, asociadas para un trabajo conjunto. Este modelo de distribución, genera las denominadas Bases de Datos Federativas.

El procesamiento de bases de datos distribuidas (BDD) consiste en trabajar con información en la cual la ejecución de transacciones y la recuperación y actualización de los datos se produce en dos o más computadoras independientes, por lo general separadas geográficamente. La utilización de bases de datos distribuidas (BDD) representa una solución viable para los usuarios cuando deben optar para la generación de sus sistemas de información. La utilización de estas bases de datos para

el mantenimiento de la información requiere el estudio de una gran cantidad de casos particulares, a fin de determinar las mejores condiciones de trabajo para cada problema real. La replicación de datos es uno de los casos particulares que se presentan cuando se estudia la utilización de BDD. Esta replicación es mucho más que simplemente copiar el dato a lo largo de varios nodos de la red. Está acompañado de análisis, diseño, implementación, administración y monitoreo de un servicio que garantiza la consistencia de la información a través de múltiples localidades en el ambiente distribuido. El propósito de distribuir datos sobre múltiples sitios es aumentar la confiabilidad, disponibilidad y eficiencia, entre otros. En estas aplicaciones, a menudo, el acceso concurrente de una gran cantidad de usuarios al sistema de base de datos, demanda alta disponibilidad y rápidos tiempos de respuesta. La replicación es el factor clave para una mayor disponibilidad (confiabilidad) y eficiencia. Los datos replicados son almacenados en múltiples sitios para que estos puedan ser accedidos por los usuarios cuando alguna de esas copias no está disponible debido a fallas en los sitios, proporcionando mayor disponibilidad. También la performance de los sistemas de base de datos distribuidas es mejorada ubicando los datos en los lugares próximos a donde se usan y poder de esta manera acceder a copias locales, reduciendo la necesidad de costosos accesos remotos

El procesamiento sobre bases de datos distribuidas continúa en evolución, por lo tanto no es posible hablar de una disciplina en plena madurez, donde para todos los problemas que se presentan existe una solución óptima y efectiva; y donde existe una caracterización única que represente una solución eficiente para aplicar en todo caso. El control de concurrencia distribuido, en comparación con uno centralizado está en pleno desarrollo. Se han propuesto una serie de algoritmos para su tratamiento y algunos de ellos cuenta con implementaciones que, si bien se adaptan a problemas específicos del mundo real, distan de ser el óptimo aplicable en todos los casos. Además, existen problemas identificados con BDD, pero con soluciones limitadas y con falta de eficiencia. Algunos resultados teóricos son difíciles de utilizar en aplicaciones reales de BDD.