

Instituto Tecnológico de Costa Rica
Administración de Tecnología de Información
TI4601 Bases de datos avanzados
Semestre II, 2018
Profesora: Ing. María José Artavia Jiménez, MAE.

Proyecto Programado 1

Diseño e implementación de una base de datos distribuida

1. Introducción

La vida ajetreada actual, ha llevado al desarrollo de nuevos servicios y soluciones innovadoras que faciliten tareas cotidianas a las personas. Uno de estos nuevos servicios en auge es el almacenaje y entrega de cualquier producto que el cliente desee. En Costa Rica, se encuentran actualmente disponibles los servicios suministrados por empresas como Correos de Costa Rica o Aeropost y una nueva compañía llamada courierTEC desea incursionar en este negocio. Usted ha sido seleccionado por esta compañía para desarrollar un servicio que permita registrar la información necesaria para las operaciones de la misma, en las secciones siguientes encontrará información que le permitirá comprender lo que debe desarrollar.

2. ¿Qué se busca con este proyecto?

- Mejorar las habilidades de resolución de problemas.
- Aumentar el conocimiento del estudiante sobre bases de datos distribuidas.
- Practicar la experimentación y la resolución de problemas (divide y vencerás).
- Ejercitar la toma de decisiones.
- Fomentar el trabajo **grupal**.
- Fomentar la **investigación** por parte del estudiante sobre como:
 - Implementar una base de datos distribuida en DBMS en particular.
 - Formato de paper propuesto por IEEE

3. Contexto y las funcionalidades esperadas

El courierTEC es una empresa distribuidora de paquetes que ha decidido implementar un sistema distribuido que le permita registrar información de los servicios que utilizan sus clientes en cada sucursal, así como monitorear y generar diversos indicadores en tiempo real. Los paquetes que despacha la empresa se clasifican en dos grandes categorías: regular o especial. En el caso de la primera, incluye electrónicos, ropa, juguetes, elementos para el hogar, entre otros. La segunda categoría incluye paquetes como comida, baterías, químicos, herramientas.

La empresa cuenta con varias sucursales, la sucursal principal se encuentra en Heredia y allí también están las oficinas administrativas, por lo que se considera a esta ubicación como las oficinas centrales de la empresa. Además, existe una sucursal en San José centro y otra en Cartago. Cada sucursal tiene un administrador que reporta al gerente en la central.

Se requiere una alta disponibilidad del sistema. De forma tal que si uno de los nodos (una sucursal) sale de la red, aún así se puedan consultar sus datos desde las oficinas centrales.

Algunos de los datos que se necesitan gestionar incluyen:

- Detalle de la sucursal, nombre, información del administrador, teléfono, correo electrónico, dirección y horario
- Para el cliente se debe almacenar la cédula, nombre, apellidos, número de cuenta, teléfono, tipo de cliente (bronce, oro o platino), provincia de residencia del cliente.
- Cada vez que se recibe un paquete en una sucursal para ese paquete se debe registrar la fecha de recepción del paquete, el cliente al que pertenece el paquete, categoría de paquete (regular o especial), tipo de paquete (electrónicos, ropa, juguetes, elementos para el hogar, comida, baterías, químicos, herramientas), descripción del contenido, el peso, el valor del paquete y el monto que debe pagar el cliente por el servicio que realiza la empresa.
- Costos básicos de paquete según el tipo y el peso.

Usted debe definir campos o relaciones adicionales que requiera registrar en el sistema.

Por lo general, los administradores de San José y Cartago utilizan el sistema para realizar las siguientes consultas:

- Cantidad de dinero recaudado (se asume que se recauda cuando el paquete es retirado) en la sucursal
- Cantidad de paquetes según cliente para un rango de fechas específico. El sistema debe presentar un listado de los clientes y los paquetes registrados durante esas fechas inclusive.
- Monto promedio pagado por paquete por cliente para un periodo de fechas específico. El sistema debe mostrar un listado de los clientes y los montos promedios de los paquetes para el periodo indicado.
- Monto de paquete para un tipo de paquete específico para un mes en particular. El sistema debe permitir escoger el tipo de paquete y el mes, presentando luego el monto de paquete que corresponda (total de montos de paquetes de ese tipo en ese mes).

En las oficinas centrales de la empresa, el gerente realiza, entre otras las siguientes consultas:

- Monto recaudado por sucursal para un periodo específico.
- Monto recaudado por sucursal y por tipo de paquete para un periodo específico.
- Listado de los tres mejores clientes (los que tenga un monto mayor en el total de paquetes que hayan traído) en un periodo específico.

4. Por hacer

1. Diseñar el sistema (la aplicación y la base de datos) de manera distribuida:
 - a. Diseñar la fragmentación y la asignación. Debe justificar claramente el porqué de la fragmentación que realiza y el porqué de la asignación.
 - b. La aplicación debe tener un módulo para que los empleados de las sucursales puedan registrar la entrada y salida de paquetes; un módulo para los clientes para consultar sus paquetes y otro para los administradores o gerentes (los administradores están asociados a las sucursales y los gerentes a las oficinas centrales). El sistema debe hacer validación de usuarios para permitir que solo usuarios registrados utilicen el sistema y que las funcionalidades utilizadas por los mismos, sean las que corresponden según su rol.
2. Se debe implementar el sistema de manera distribuida según el diseño propuesto en el punto 1.
3. Generar un paper donde se describa la metodología utilizada para generar la base de datos distribuida en el DBMS seleccionado, el paper debe contener las siguientes secciones: introducción, descripción del problema, un resumen de la investigación realizada para implementar una base de datos distribuida en el DBMS seleccionado, un ejemplo de la implementación de la base de datos distribuida para el contexto de este proyecto y referencias bibliográficas; conclusión. Se debe utilizar el formato de IEEE para papers.

5. Puntos a ser evaluados

1. **Correctitud** de la solución computacional
2. **Robustez y resiliencia** de la solución computacional
3. **Diseño y asignación de bases de datos distribuida.**
4. **Implementación de una base de datos distribuida**
5. **Implementación de aplicación** con sus respectivos módulos.
6. **Conocimiento** de la solución generada.
7. **Redacción de paper con el formato propuesto por IEEE.**
8. **Desarrollo** de documentación efectiva.
9. **Uso de mejores prácticas** de programación en su código. Puede revisar los principios del programador y olores de software (Clean Code).

6. Aspectos técnicos

Para la implementación del sistema puede utilizar el DBMS de su preferencia siempre y cuando se cumpla con la implementación de la base de datos distribuida.

Puede utilizar el lenguaje de programación de su preferencia para implementar la aplicación correspondiente.

7. Evaluación

Rubro	Puntaje
Base de datos distribuida	50
Implementación correcta de la asignación	15
Implementación correcta de la fragmentación	15
A través de la aplicación permite consultar información de cualquier nodo desde la sede central aún si el nodo no está disponible	10
Paquetes registrados en los nodos se reflejan inmediatamente en el nodo central	10
Requerimientos de la aplicación	30
Registros de paquetes de clientes	6
Consultas de administradores de sucursales	12 (3 pts c/u)
Consultas para gerente en oficinas centrales	12 (4 pts c/u)
Documentación	20
Aspectos generales	1
Diseño de la base de datos	2
Correctitud del diseño y justificación de la fragmentación	4
Correctitud del diseño y justificación de la asignación	3
Paper con formato IEEE	10
TOTAL	100

8. Aspectos administrativos

1. Entregar **un documento** con al menos los siguientes apartados. Los apartados DEBEN entregarse en ese orden:
 - a. Portada.
 - b. Aspectos generales: indicar en este apartado cuáles herramientas se utilizaron para implementar el sistema y el DBMS con el cual se implementó la base de datos distribuida.
Debe incluir también el estado actual del proyecto: si es 100% funcional o bien indicar qué no se pudo implementar y por qué.
 - c. Diseño de la base de datos (describir tablas y sus campos). Incluya el modelo (diagrama de la base de datos, antes de la fragmentación)
 - d. Diseño de la fragmentación: Indique claramente sobre cuáles tablas hace fragmentación y cuáles predicados utiliza. Debe explicar el diseño.
 - e. Diseño de la asignación, debe indicar claramente qué tablas y fragmentos quedaron en cada nodo y justificar por qué.
 - f. Para cada consulta se debe mostrar un *screenshot* donde se observe claramente la consulta realizada y su resultado.
 - g. Debe incluir *screenshots* para el ingreso de la información de los paquetes de los clientes.

Si alguna funcionalidad queda sin documentar, no será tomada en cuenta durante la evaluación y el grupo de trabajo perderá el puntaje respectivo.

2. Debe crear un archivo .zip ("PP1.zip") que contenga únicamente un archivo **info.txt** y 3 carpetas llamadas **documentación**, **aplicacion** y **baseDatos**, en la primera deberá incluir el documento de *word* o *open office* (no pdf) solicitado y en la segunda el código fuente de la aplicación y en la tercera los scripts de la base de datos. El archivo **info.txt** debe contener la siguiente información (cualidades):
 - a. Nombre del curso
 - b. Número de semestre y año lectivo
 - c. Nombre del Estudiante x4
 - d. Número de carnet x4
 - e. Número de proyecto programado
 - f. Fecha de entrega
 - g. Estatus de la entrega (definido por el responsable de la implementación de la tarea, debe ser **CONGRUENTE** con la solución entregada):
[Deplorable|Regular|Buena|MuyBuena|Excelente|Superior]
3. Deberá subir el archivo antes mencionado al TEC Digital en el curso de BASES DE DATOS AVANZADOS GR 0X, en la asignación llamada "P01" debajo del rubro de "Proyectos programados". El proyecto debe entregarse a más tardar el **domingo 16 de setiembre a las 11:55pm para el grupo 1, y el lunes 17 de setiembre a las 11:55 pm para el grupo 2**
4. Dentro de la carpeta de **aplicacion**, deberá incluir un archivo .txt (**PrimerNombreMiembro1.PrimerNombreMiembro2.PrimerNombreMiembro3.PrimerNombreMiembro4.txt**) que contenga todo el texto de la solución del o de los archivos presentados (las implementaciones). Este archivo puede ser revisado en el sistema de Control de Plagio del TEC Digital. **Todo el código de cada proyecto debe ser 100% original, y no se va a tolerar el plagio.** Este archivo también deberá ser subido al TEC Digital en el curso de BASES DE DATOS AVANZADOS GR 0X, en la asignación llamada "P01 (archivo TXT)" debajo del rubro de "Proyectos". Se deberá subir el archivo siguiendo los mismos rangos de tiempo que los descritos en el punto anterior.
5. Los proyectos deberán ser revisadas con el profesor o el asistente. Todos los miembros del grupo deberán participar de la revisión, de lo contrario a la nota obtenida grupalmente se le descontará 15%. Además, si durante la revisión el estudiante no puede demostrar conocimiento de la solución presentada, se le descontará hasta un 15% de la nota grupal.
6. La tarea se hará en grupos de 4 personas, solamente.
7. La tarea vale un 15% de la nota del curso.
1. Antes de la entrega del proyecto se pasará en clase una lista donde deben anotar los miembros de su equipo de trabajo. Opcionalmente, pueden mandar un correo electrónico

con el asunto “[Avanzados 218] PP1:GRUPO” e indicar los nombres de las personas que conforman el grupo de trabajo para este proyecto programado. Esta información es necesaria para crear los grupos en el TEC Digital. Si no se anotan en la lista, o no envían el correo, se calificará la tarea de forma individual al miembro del equipo que la suba.

9. Condiciones

Este proyecto programado se rige por las siguientes condiciones:

1. El desarrollo del proyecto es estrictamente en grupos de 4 integrantes.
2. Debe cumplir con todo lo indicado en esta especificación.
3. Deberá entregarse en tiempo y forma según el plazo establecido por la profesora al momento la lectura de este documento.
4. Por simplicidad se supondrá que además de las oficinas centrales solo existen dos sitios más desde donde los usuarios pueden utilizar el sistema. Por lo que la red contará solo con 3 nodos.
5. La aplicación debe poder ejecutarse desde cualquiera de los tres nodos simultáneamente inclusive.
6. No se recibirán proyectos fuera de la fecha/hora indicadas, tampoco se recibirán por correo electrónico ni otro medio que no sea el TECDigital. **Todo trabajo entregado después de la fecha y hora indicada tendrá una nota de cero.**
7. Si se detecta copia o plagio se procederá según lo indicado en el Reglamento del Régimen de Enseñanza Aprendizaje del ITCR.
8. Las citas de revisión del proyecto se asignarán después de la entrega del mismo.
9. Para la revisión, el grupo de trabajo deberá llevar los equipos necesarios para poder mostrar el sistema distribuido en funcionamiento.
10. No se harán revisiones previas de los diseños.

Nota: El incumplimiento de alguna condición implicará una calificación de cero.

IMPORTANTE: CONOCIMIENTO DE LA SOLUCIÓN PRESENTADA. En la revisión del trabajo, los estudiantes deben demostrar un completo dominio de la solución implementada, tanto desde el punto de vista técnico (uso de herramientas) como de la funcionalidad del proyecto. La revisión se puede hacer individualmente o en grupos, examinando la solución o temas específicos aplicados en el proyecto. Todos los integrantes del grupo deben tener el mismo conocimiento de la solución presentada. Recuerde que de no contestar las preguntas durante la revisión, se puede descontar hasta un 15% de la nota obtenida de forma grupal.