

11/15/2014

Por Luis Daniel Aguiar y Cesar Celis

ITESM, cAMPUS gUADALAJARA

Proyecto Final

Base de Datos Distribuidas

Índice

[Introducción 2](#_Toc403842753)

[¿Qué vamos a hacer? 2](#_Toc403842754)

[¿Qué contiene el documento? 2](#_Toc403842755)

[¿Qué no pretende abordar este documento? 2](#_Toc403842756)

[Desarrollo 3](#_Toc403842757)

[Análisis y descripción del problema 3](#_Toc403842758)

[Diseño funcional 3](#_Toc403842759)

[Fragmentación utilizada 3](#_Toc403842760)

[Estadísticas de uso de la información 4](#_Toc403842761)

[Solución técnica 5](#_Toc403842762)

[Herramientas utilizadas para la solución 5](#_Toc403842763)

[Manual de uso de la aplicación Web 7](#_Toc403842764)

[Conclusión 8](#_Toc403842765)

[Logros: 8](#_Toc403842766)

[Áreas de Oportunidad: 8](#_Toc403842767)

[Bibliografía 10](#_Toc403842768)

# Introducción

## ¿Qué vamos a hacer?

En este proyecto final, haremos una implementación de una base de datos distribuida que contendrá la información necesaria para la validación de hardware que se llevara a cabo en una empresa de semiconductores. Con el objetivo de hacer más eficiente el uso y acceso de la información que se localiza en las diferentes sedes de dicha empresa. Además, se busca proveer de una interfaz gráfica amigable que le permitirá a nuestro usuario final consultar y actualizar la información de las diferentes sedes o nodos. Dicha interfaz gráfica se desarrollara sobre un entorno Web para el fácil acceso de los usuarios a la base de datos.

## ¿Qué contiene el documento?

Este documento contiene información detallada sobre diferentes aspectos que debemos cubrir para el desarrollo e implementación de la base de datos distribuida. Dichos aspectos los dividimos en tres ramas principales. Por un lado, tenemos el análisis y descripción del problema, que nos permitirá comprender las necesidades que tiene el usuario. Por otro lado, tenemos el diseño funcional, que nos permitirá ver las fragmentaciones de las tablas y las estadísticas de uso que tiene la información. Y finalmente esta la solución técnica, en donde se abordan temas sobre las herramientas utilizadas para resolver las necesidades y el modo de uso de la aplicación Web que permitirá al usuario interactuar con la base de datos.

## ¿Qué no pretende abordar este documento?

Este documento no pretende abordar los detalles de instalación de los servidores de SQL, dado que esta información se encuentra en la documentación de Microsoft en la Web. Tampoco se pretende documentar o explicar el código utilizado para la creación de la aplicación Web. Dado que resulta intrascendente en términos del sinfín de posibles alternativas que existen en la actualidad para mostrar la información al usuario final. Otros de los aspectos que no se abordaran en este documento son las especificaciones del sistema que se requieren para este sistema, ni las versiones de herramientas utilizadas, dado que el documento es una simple descripción de cómo se pudiera lograr la creación de una base de datos distribuidas. Además recordemos que en la nube podremos encontrar mucha de la información necesaria para la solución de problemas específicos a ciertas plataformas.

# Desarrollo

## Análisis y descripción del problema

Existe una empresa dedicada a la creación de hardware dentro del sector de semiconductores que requiere de la constante validación de sus productos. Para garantizar al OEM por sus siglas en inglés (Original Equipment Manufacturer) que el producto final cumple con las especificaciones de diseño para las cuales fue creado. Para llevar a cabo esta labor de validación, se desarrollan pruebas en diferentes lenguajes de programación que son utilizadas para detectar errores en las diferentes fases o etapas por las cuales atraviesa un proyecto. Para lo cual hemos desarrollado una base de datos que contendrá toda la información concerniente a dichas pruebas. Esto con el fin de garantizar el fácil manejo y administración de la ejecución que se lleva a cabo dentro de las diferentes etapas de validación del hardware.

## Diseño funcional

A continuación se muestra un diagrama (Diagrama [UML](#_Bibliografía) de Base de Datos) de la base de datos que se implementó para el proyecto. Que consta de varias entidades. Es un diagrama básico que deberá ir creciendo conforme a las necesidades del proyecto, pero que cuya función se centra en las pruebas que se utilizan para la validación, siendo el campo de los resultados de las pruebas uno de los más consultados para el reporteo del avance de la validación. Permitiendo además al cliente mostrar el progreso de la ejecución a los ingenieros de validación y otras áreas funcionales de la empresa.



Diagrama [UML](#_Bibliografía) de Base de Datos

### Fragmentación utilizada

En el proyecto hemos decido hacer uso de la fragmentación horizontal para dividir las tablas que se utilizan en las diferentes sedes basados en la propiedad de las unidades funcionales del sistema a validar. Es decir, el cliente cuenta con 2 sedes, una localizada en Estados Unidos y la otra localizada en Mexico. Las personas que trabajan en Estados Unidos son dueños de las siguientes áreas: eDP, DP y HDMI. Mientras que en Mexico son dueños de las siguientes áreas: WD, FBC y Pipes. Por lo tanto la tabla llamada Pruebas estará fragmentada de forma horizontal en 2 áreas. La primera área contendrá eDP, DP y HDMI y la segunda área contendrá WD, FBC y Pipes. Como muestra a dicha fragmentación véase la Imagen 1. En donde las pruebas con ID: [1, 2, 3] pertenecen a un servidor y las pruebas con ID: [4, 5, 6] perteneces a otro servidor y la unión de ambas tablas forman la vista que se presentara al usuario en la aplicación Web.

**test\_table**

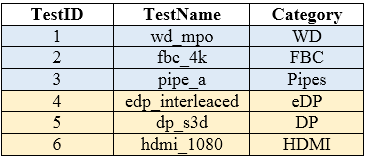
 

Imagen 1

### Estadísticas de uso de la información

Cada categoría contiene un determinado número de pruebas que son consultadas por el dueño de esa categoría en la mayoría de los casos y con contadas excepciones. Por lo que debido a esta situación el balanceo de la carga se determinó con base a esta restricción. Y con lo cual podemos generar una estadística del porcentaje de pruebas que se localizan en cada servidor. Véase la Imagen 2 en donde se ilustra el porcentaje de registros que hay en cada servidor para el caso de la tabla de pruebas.

Imagen 2

## Solución técnica

### Herramientas utilizadas para la solución

#### Servidores

Se optó por utilizar [My SQL Server 2014](#_Bibliografía_1) debido a la magnífica integración que ofrece Microsoft® con otras herramientas de desarrollo como la [IDE](#_Bibliografía) por sus siglas en inglés (Integrated Development Environment) llamada Visual Studio que permitió al proyecto la integración y creación de la aplicación Web con conexión a los diferentes servidores que contienen la información fragmentada de las pruebas.

Mediante Microsoft SQL Server Magnamente Studio se realizó la creación y configuración de los servidores. Ambos servidores pueden ser observados en la imagen 3 en donde se muestra como pueden coexistir en la misma instancia de manejador de servidores. Sin necesidad de usar instancias por separado siempre que ambos servidores estén dentro del mismo ordenador.

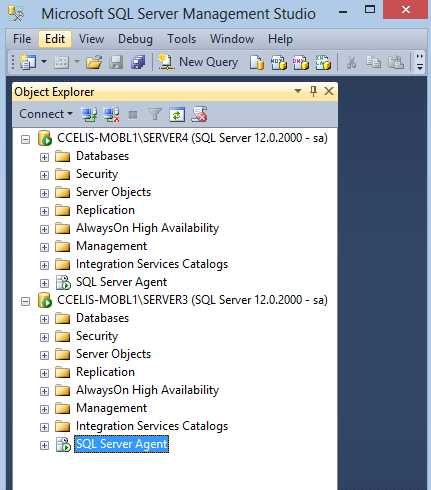


Imagen 3

Así mismo, dentro de este mismo manejador se crearon las bases de datos, las tablas y los scripts de SQL o Queries para configurar las tablas y restricciones de inserción en los registros entre muchas otras cosas. Veamos un ejemplo básico del tipo de cosas que se pueden hacer con esta herramienta:

1. Query básico para la creación de una tabla:

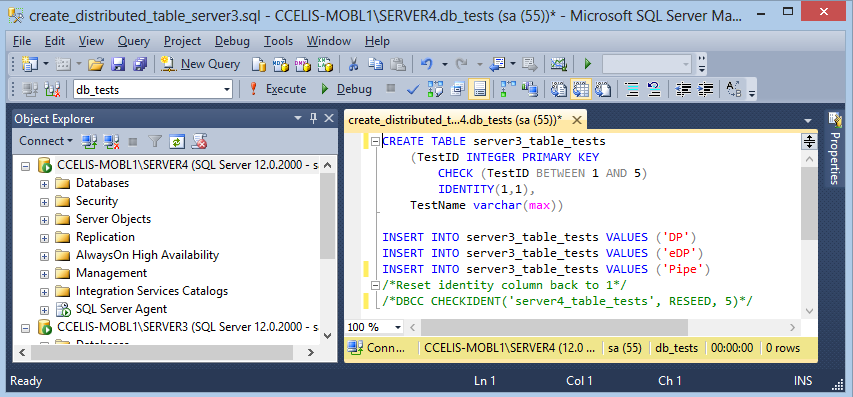


Imagen de Query básico en SQL Server

#### Desarrollo de la aplicación Web

Una vez creadas las tablas con sus restricciones y reglas de inserción se procede a crear la aplicación Web haciendo uso de la [IDE](#_Bibliografía) Visual Studio. Que mediante el paquete [Entity Framework](#_Bibliografía) y el uso del lenguaje C# se hace la conexión a los servidores para intercambiar información entre el usuario final y la base de datos. Véase la imagen flujo de información que se presenta a continuación. Para observar de forma gráfica el flujo que se da entre el usuario y la base de datos distribuida.



Imagen flujo de información

### Manual de uso de la aplicación Web

La aplicación web se representa por una interfaz gráfica muy sencilla pero poderosa, que le permite al usuario las siguientes funcionalidades:

* Insertar registros
* Actualizar registros
* Borrar registros
* Consultar vistas

Cada una de estas funciones está representada por un botón web y un campo de texto web para que el usuario final pueda seleccionar que registro desea insertar, actualizar, borrar y/o consultar. Véase la imagen Versión 1 de la aplicación Web. Sin embargo, la interfaz puede cambiar durante el desarrollo del proyecto para incorporar nuevas funcionalidades o formas de operación. Pero en esencia se espera el mismo comportamiento o funcionalidad de la misma a lo largo del proyecto. También cabe destacar que la interfaz gráfica debe ser intuitiva y fácil de comprender para el usuario. Por ejemplo en la Imagen Versión 1 de la aplicación Web podemos ver que existen colores que nos permiten diferenciar entre el servidor 3 y el servidor 4. También los botones tienen nombres específicos a su función, por ejemplo:

1. Insertar Server3, que inserta un registro a la tabla del servidor 3
2. Delete Server3, que borra un registro de la tabla del servidor 3
3. Un campo Result, para desplegar los resultados

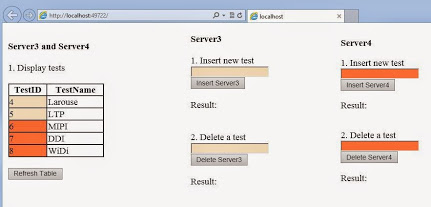


Imagen Versión 1 de la aplicación Web.

# Conclusión

Este proyecto final permitió el aprendizaje de la creación y uso de herramientas de una base de datos distribuida. Sin bien, el alcance va mucho más allá de lo realmente logrado durante este proyecto semestral, se cumple con el objetivo principal de dejar la semilla del conocimiento en cada alumno. Para que con el tiempo se puedan realizar obras de un grado mayor de complejidad que contribuyan al desarrollo de una mejor organización e impacten positivamente la productividad de los usuarios y creadores de las bases de datos. Bases de datos que hoy en día son tan necesarias e indispensables para almacenar la gran cantidad de información que se encuentra disponible en cada empresa y en cada proceso.

Con lo anterior en mente hacemos una breve síntesis de lo que se logró de forma exitosa en este proyecto, y a su vez de las áreas de oportunidad que quedan abiertas para el futuro.

## Logros:

1. Planteamiento de las necesidades del usuario
   1. Se logró plantear un escenario que requería del uso de una base de datos distribuida entre las sedes de una empresa de semiconductores para la validación de hardware.
2. Creación de una base de datos distribuida
   1. Fragmentación horizontal: Se logró establecer de forma exitosa la fragmentación horizontal de una tabla de pruebas basados en el hecho de que las categorías pertenecías a ingenieros localizados en sedes diferentes del mundo.
3. Creación de una aplicación Web para el manejo de la base de datos
   1. Conexión a los servidores de SQL: Gracias al manejo de herramientas como [Entity Framework](#_Bibliografía_1) se logró establecer una conexión de fácil acceso a los servidores SQL mediante código en C# y un buen ambiente de integración de desarrollo como el que encontramos con Visual Studio.
   2. Intercambio de información entre el usuario final y la base de datos
      1. Inserción: Se logró hacer la inserción de registros con el auto incremento de las llaves primarias, gracias al uso de comandos de SQL como Identity y además una validación de registros validos gracias a las restricciones que se utilizaron con el CHECK de SQL.
      2. Actualización: Se logró hacer la actualización de registros y/o tablas enteras
      3. Eliminación: Se logró hacer la eliminación de registros no deseados y además regresar la semilla de enumeración de la llave privada, gracias al uso de queries tales como:

‘DBCC CHECKIDENT('server4\_table\_tests', RESEED, 5)’

* + 1. Consulta: También se logró hacer la consulta tanto de tablas como de vistas

## Áreas de Oportunidad:

1. Aplicar el proyecto en un escenario real: Si bien se logró plantear un escenario ficticio para el desarrollo del proyecto. Un buen siguiente paso, sería el llevar a cabo la implementación de un proyecto sobre un escenario real que supondría mayores retos con un mayor número de variables que involucraría entre otras cosas, contratos, reuniones con el cliente, planes de mantenimiento, planes de actualización, acuerdos de confidencialidad, etc.
2. Creación de una base de datos distribuida más compleja
   1. Fragmentación vertical: Conforme el escenario se vuelve más complejo se haría necesario plantear la posibilidad de fragmentar las tablas de forma vertical para ocultar cierta información al usuario como podría ser el caso del número de tarjeta de crédito o resultados visibles solo para los administradores.
   2. Fragmentación mixta: También se podría recurrir a la combinación de ambas fragmentación para hacer más eficiente el uso de una tabla al dividir y/o ocultar cierta información que haría entre otras cosas más veloces a las búsquedas.
   3. Algoritmos de fragmentación: Utilizar algoritmos para automatizar la fragmentación y evitar la intervención manual del desarrollador para mantener un ambiente de crecimiento y portabilidad.
   4. Procedimientos de SQL avanzados para la inserción de nodos de forma automática: Buscar la manera de instalar de forma automática otros nodos o servidores mediante scripts que actualicen las vistas e incluyan de forma automática a los nuevos nodos agregados.

# Bibliografía

Referencia de la herramienta para crear un servidor SQL:

Microsoft Cloud Platform (2014) SQL Server, 15 de noviembre de 2014, <http://www.microsoft.com/en-us/server-cloud/products/sql-server/>

Referencia del término UML:

Unified Modeling Language (2014) Getting Started with UML, 15 de Noviembre de 2014, <http://www.uml.org/>

Referencia del término IDE:

Microsoft Documentation (2005) Walkthrough: Using the Visual Studio IDE, 15 de Noviembre de 2014, <http://msdn.microsoft.com/en-us/library/ms235632(v=vs.80).aspx>

Referencia del término OEM:

Barnor, Albert y Browne, Lynn Elaine (2003) Ken Olsen. PDP-1 and PDP-8, 15 de Noviembre de 2014, <http://www.economicadventure.org/pdfs/olsen.pdf>

Referencia de Entity Framwork:

Microsoft Data Developer Center (2014) Learn Section, 15 de Noviembre de 2014, <http://msdn.microsoft.com/en-us/data/ef.aspx>