

INSTITUTO TECNOLÓGICO DE COSTA RICA

PROYECTO #2

Alejandro Rojas
Saúl Zamora

profesor
Kevin Moraga

June 19, 2017

1 Introducción

Se requiere un sistema para una empresa productora de dispositivos de IoT (Internet Of Things), los cuales se distribuyen a nivel nacional y se exporta a Nicaragua, Panamá y Guatemala.

Uno de los principales negocios de la empresa son las “Smart Houses”, lo cual implica brindarle al cliente una experiencia única al darle control total de su casa y los dispositivos en ella.

Las bases de datos deben almacenar información de forma local y otros datos de forma remota en las instalaciones de la empresa. Actualmente, las tecnologías usadas son MySQL, SQL Server y Oracle.

La base de datos de producción posee un inventario de todas las partes que la empresa adquiere como materia prima para la creación de los dispositivos. Existe otro sistema que usa otra base de datos donde se controlan las ventas de los productos y dispositivos, la cual es básicamente un catálogo de los productos; además de un historial con los cambios en los precios de dichos productos.

Cada casa de habitación cuenta con una base de datos local, cuya función principal es recolectar información post-venta de los dispositivos e información generada por su uso.

Es requerido mantener las bases de datos locales y remotas replicadas para brindarle al usuario la experiencia que desea.

Además, el sistema deseado debe ser capaz de producir estadísticas sobre inteligencia de negocios con el fin de ayudar a la gerencia a tomar decisiones más acertadas. Algunos de los puntos más relevantes son:

- Costos de producción
- Sugerencias de consumo de productos del cliente en los próximos 15 días.
- Dispositivos IoT más utilizados por un cliente
- Dispositivos IoT más vendidos
- Tiempo de entrega de materiales relacionados a los dispositivos IoT más vendidos
- Estadísticas sobre los proveedores (calidad de materiales, precios, etc)
- Determinar los materiales de mayor y menor circulación
- Estadísticas sobre las ventas anuales
- Estadísticas sobre los carriers de los productos (volumen enviado, más utilizados, destinos, etc)
- Índices de ganancias vs índices de gastos

2 Ambiente de desarrollo

- Sistema operativo host utilizado: Windows 10
- Python 3.6 (Simulación)
- Oracle VirtualBox 5.1
- Windows Server 2008 R2 (Dos máquinas virtuales)
- SQL Server 2008 Enterprise
- Pentaho 7.1

3 Estructuras de datos usadas y funciones

Para la simulación se usaron listas para almacenar datos aleatorios, los cuales luego se guardan en archivos CSV. El archivo de simulación genera todos los datos en una sola corrida, por lo que no utiliza otras funciones en particular. Sin embargo, se utilizaron las siguientes librerías para su correcta ejecución:

- csv
- random (randint)
- datetime (date, timedelta)
- operator
- numpy

4 Instrucciones para ejecutar el programa

Para generar los archivos CSV con los datos simulados, hay que colocar el archivo *dataSimulator.py* en la carpeta que se desea almacenarlos. Luego se debe ejecutar ya sea mediante el IDLE de Python, presionando F5, o mediante la línea de comandos, ejecutando directamente el programa.

Una vez generados los archivos, se puede utilizar el servicio de *import* de SQL Server para insertar los datos en las bases de datos respectivas.

5 Bitácora de trabajo

5.1 Alejandro Rojas

- 12-06-2017:
 - 1.5 horas - Descarga e instalación de Pentaho y Windows Server 2008 R2.

- 13-06-2017:
 - 4 horas - Diseño de base de datos y simulación. Diseño y simulación de productos y distribuidores.
- 14-06-2017:
 - 8 horas - Diseño y simulación de materiales, categorías, modelos, etc.
- 15-06-2017:
 - 5 horas - Simulación de ventas y despachos de productos.
- 17-06-2017:
 - 4 horas - Detalles de la simulación.
- 18-06-2017:
 - 3 horas - Conectar 2 máquinas virtuales con Windows Server 2008 R2.
 - 2 horas - Crear carpetas compartidas para archivos CSV, Excel y Replicaciones. Configuración correcta de la replicación de la base de datos de inventario en la segunda máquina virtual.
 - 1.5 horas - Configuración correcta de la replicación de la base de datos de ventas en la primera máquina virtual.
 - 1 hora - Terminar llenado de la base de datos de inventarios.
- 19-06-2016:
 - 8 horas - Configuración de Pentaho y proceso de ETL / reportes.
- Total horas: 38 horas.

5.2 Saúl Zamora

- 12-06-2017:
 - 2 horas - Investigación sobre replicación de base de datos en SQL Server 2008 Enterprise.
- 13-06-2017:
 - 5 horas - Descarga de Windows Server 2008 R2 y configuración de la primera máquina virtual.
- 14-06-2017:
 - 1 hora - Configuración de la segunda máquina virtual con Windows Server 2008 R2.
 - 3 horas - Descarga de SQL Server Enterprise.

- 15-06-2017:
 - 2 horas - Instalación de SQL Server en ambas máquinas virtuales.
 - 2 horas - Configuración de carpetas compartidas en las máquinas virtuales para archivos CSV.
- 17-06-2017:
 - 5 horas - Intento de configuración de replicación de bases de datos con las máquinas virtuales de Windows Server y SQL Server.
- 19-06-2017:
 - 4 horas - Documentación.
- Total horas: 25 horas.

6 Comentarios finales

7 Conclusiones

- Es importante tomar en cuenta la configuración en el firewall de Windows Server 2008 R2 a la hora de configurar máquinas virtuales con el propósito de que sean capaces de “verse entre sí”.

References

- [1] Singh, S. (2017). SQL Server Performance Setting up Transactional Replication in SQL Server 2008 R2. [online] Sql-server-performance.com. Available at: <http://www.sql-server-performance.com/2010/transactional-replication-2008-r2/>

8 Anexos

Figure 1: Diagrama de la base de datos de Inventario

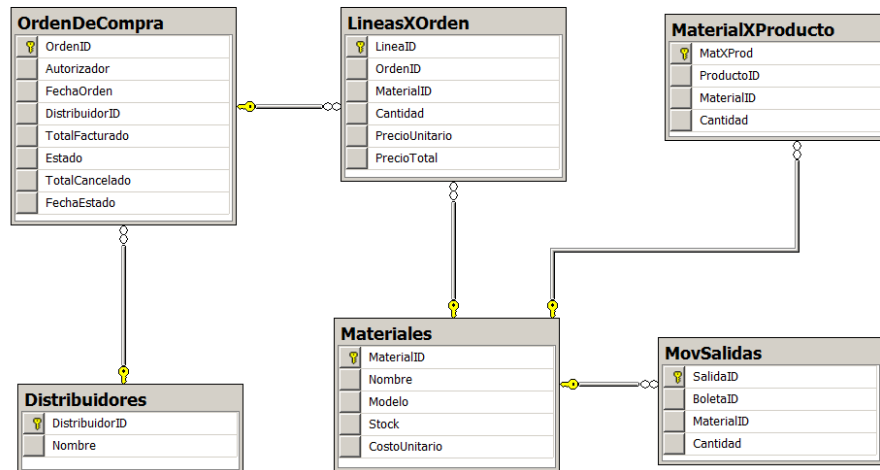


Figure 2: Diagrama de la base de datos de Ventas

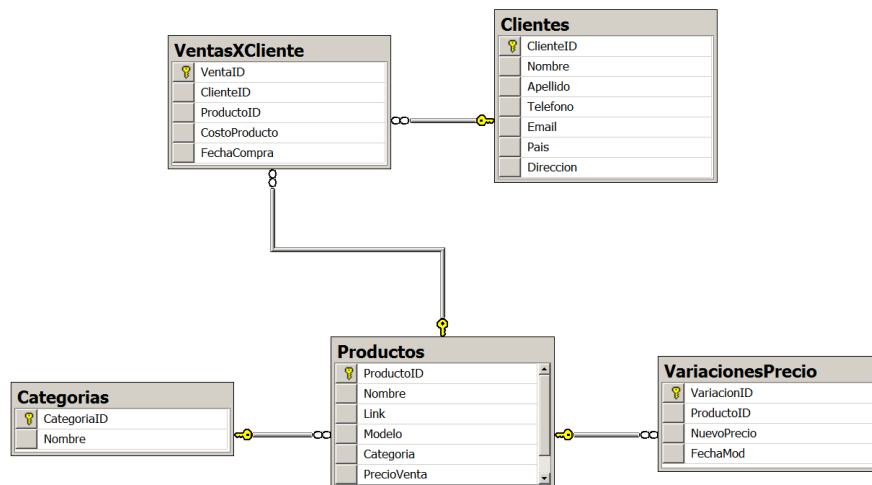


Figure 3: Configuración de conexión a Pentaho

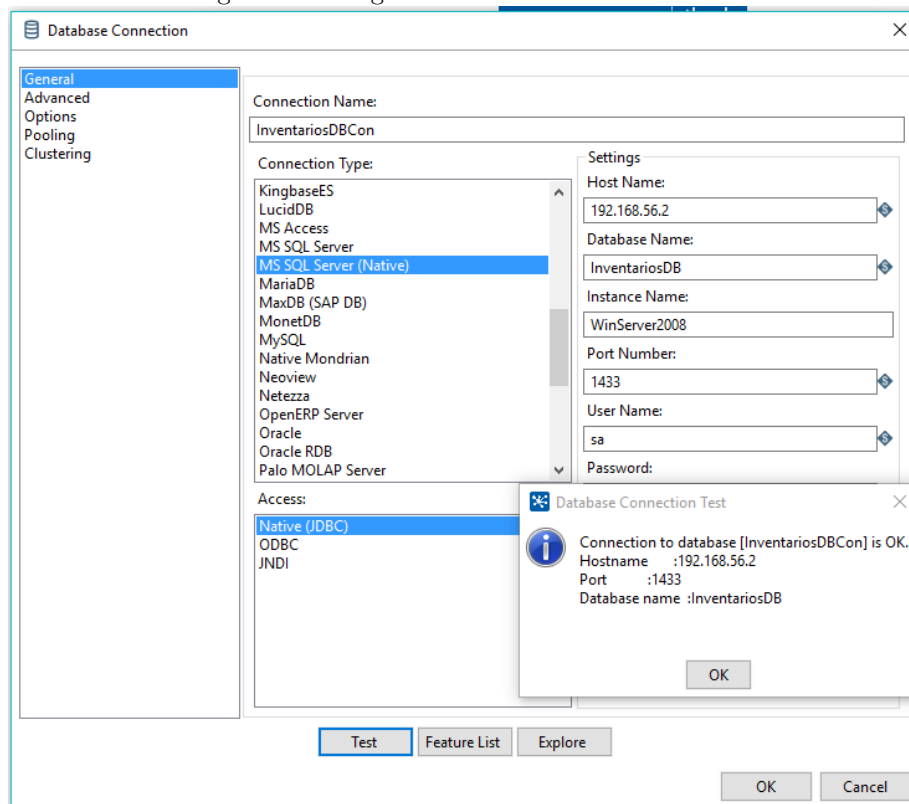


Figure 4: Estados de las bases de datos replicadas

