

# LOTR: One Ring To Rule Them All

---

Estudiantes:

- José Gustavo González González.
- Cristopher Agüero
- Juan José Salas Arguedas

## Introducción

---

### Objetivo

“One Ring To Rule Them all”. Realizar una implementación de Inteligencia de Negocios para una empresa virtual costarricense. Incluyendo un sistema a prueba de fallos que permita aumentar el uptime del servicio.

### Requerimientos Funcionales

Los datos que se desean analizar de los sistemas son:

- Costos de producción
- Sugerencia de consumo de productos del cliente en los próximos 15 días.
- Dispositivos IoT más utilizados por el cliente
- Dispositivos IoT más vendidos
- Tiempo de entrega de materiales relacionados a los Dispositivos IoT más vendidos
- Malos proveedores ya sea por tiempo de respuesta o por precios muy elevados o por mala calidad de los artículos o materiales
- Determinar aquellos materiales de mucha circulación y de poca circulación
- Las ventas que realizan durante todo el año teniendo en cuenta los destinos de las mismas
- Les es importante saber cuáles son los carrier con los cuales están enviando más volumen y relativo a los destinos de los productos como es que se comportan los diferentes carriers
- Como todo gerente le interesa mucho saber cuánto gasta versus cuánto gana

### Simular el estado actual del sistema de LOTR

- Diseñe e implemente la base de datos de inventario de materiales en bodega. Rellene la base de datos con los 6000 materiales que permitan realizar unos 100 dispositivos IoT, clasificados en distintos números de serie, y modelos, entradas y salidas que completen unos 80000 movimientos realizados en los últimos 5 años no uniformemente distribuidos.
- Diseñe e implemente la base de datos de precios de productos, agregando un total de 100 productos reportando de 30 a 70 variaciones de precio por producto distribuidos en los mismos 5 años. Tenga presente que la empresa tiene que estar ganando dinero de lo contrario ya hubiera quebrado.
- Diseñe e implemente la hoja de Excel ingresando despachos para todos los productos existentes, concretando mínimo 9000 no uniformemente distribuidos en los 5 años
- Diseñe e implemente una recolección de datos de los dispositivos IoT que represente el uso de 1 año, en la casa de habitación. Por lo que al menos deberían de tener 52560 entradas tomando en cuenta que se realiza una toma cada 10 minutos. Esta base de datos debe de ser de al menos 25 clientes.
- Se revisará el diseño de las bases de datos

### Infraestructura

- Deberá usar mínimo dos servidores virtuales Windows 2008 Server y SQL Server 2008 R2 o bien MySQL Server sobre GNU/Linux. Cada servidor tendrá una base de datos y la hoja de Excel puede estar en un folder compartido de la red
- Utilizando replication services proceda a implementar el esquema de replicación de forma que la base de datos de

inventario se replique en la de productos y la de productos se replique en la de inventario. Se harán pruebas de replicación en la revisión

- Utilizando clustering services y otros dos servidores virtuales (un pasivo y un domain controller) haga que uno de los servidores de base de datos cuente con un esquema fail over en modo cluster activo/pasivo. Se hará la prueba de que el fail over funciona.
- Opcional 10pts: Utilizando los servicios (si la implementación utilizó SQL Server) de mirroring y otros dos servidores virtuales (un pasivo y un witness que también funciona de domain controller) implemente una arquitectura failover con testigo. Se hará la prueba de que el failover funciona

## Business Intelligence

- Según las necesidades de los gerentes diseñe e implemente el datawarehousing necesario utilizando modelo en estrella.
- Haga uso de integration services para hacer el ETL con el que se va a popular el datawarehousing, use las 3 fuentes de datos: Excel, base de inventario, base de productos y base de IoT.
- Por medio de análisis services implemente el o los cubos necesarios para ser usados en la solución de BI y haga deployment de los mismos.
- Demuestre que es posible ahora hacer consultas y reportes a los data marts creados utilizando reporting services y Excel

### Otras consideraciones

Además de las definiciones anteriores tome en cuenta:

- Se evaluará aspectos de diseño, calidad de código, optimización de código
- Deberá seguir las pautas de calidad de código y el estándar de código adjunto a este enunciado
- Las revisiones preliminares pueden ser por correo o presenciales
- Cualquier otra propuesta hecha por el estudiante debe de consultarla con el profesor.

## Ambiente de desarrollo

---

Herramientas utilizadas por los estudiantes:

- Python
- MySQL-Cluster
- LXD
  - De la mano con el sistema operativo Ubuntu 16:04
- Pentaho
- Spoon
- Workbench

## Estructuras de datos usadas y funciones

---

Para esta tarea usamos 3 bases de datos, las cuales son **bodega**, **producción e iot**.

### Bodega

En esta base de datos encontramos todo lo relacionado a la bodega e inventario de materias primas para la elaboración de dispositivos. sus principales tablas son:

#### Proveedores

En esta tabla tenemos todos los proveedores de los cuales compramos materiales.

#### Empleados

En esta tabla tenemos la lista de empleados de nuestra empresa.

#### Departamento

En esta tabla tenemos la lista de departamentos a lo interno de la empresa.

### **Materiales**

En esta tabla tenemos la lista de materiales con los cuales fabricaremos dispositivos IoT.

### **Ordenes\_de\_compra**

En esta tabla simulamos una factura de compra de materiales a los proveedores.

### **Detalle\_orden**

En esta tabla tenemos el detalle de cada linea de factura de ordenes\_de\_compra.

### **Salida\_de\_bodega**

En esta tabla simulamos una salida de la bodega de materiales para ser convertidos en dispositivos IoT.

### **Detalle\_salida**

En esta tabla tenemos el detalle de cada linea de factura de salida\_de\_bodega.

### **Devoluciones**

En esta tabla tenemos materiales defectuosos los cuales son devueltos a los proveedores y cancelar la factura de dichas partes.

## **Producción**

En esta base de datos tenemos los detalles de los dispositivos que creamos para su posterior venta.

### **Productos**

En esta tabla tenemos los dispositivos en stock para su posterior venta.

### **Precios**

En esta tabla tenemos el historial de precios de los dispositivos IoT.

## **IOT**

En esta base de datos tenemos los datos recolectados de nuestros clientes, así como sus datos.

### **Clientes**

Simplemente es una tabla con los datos de nuestros clientes.

### **Casa**

Esta es una tabla que asocia un cliente con una casa.

### **Dispositivos**

Esta tabla contiene los dispositivos que se encuentran recolectando datos para luego ser analizados.

### **Datos**

En esta tabla tenemos los datos recolectados a travez del tiempo por los dispositivos en distintas casas.

## **Archivo excel**

En este archivo tenemos los datos de entrega de dispositivos IoT a nuestros clientes al rededor del mundo. En esta hoja de cálculo tenemos datos del cliente, el empleado encargado, el carrier y demás.

## **Instrucciones para ejecutar el programa**

---

Para ejecutar la tarea se debe seguir una serie de pasos:

### **Inicializar el Cluster**

Para hacer esto, brindamos toda una documentación acerca del cluster y cómo realizarlo.

## Montar las bases de datos SQL.

En este punto se deben montar las bases de datos en el cluster, para esto, se ejecuta en algún cliente del cluster el script provisto en la carpeta de la tarea.

### Business Intelligence

Realizar la conexión con ETL y el diagrama estrella, para de ahí realizar las relaciones importantes.

## Actividades realizadas por estudiante

---

Se desglosan en el formato:

Fecha - Cantidad Horas Invertidas - Tarea - Estudiante

- 13 de Junio al 15 de Junio - 10 horas - Diseño de las bases de datos bodega y producción. - Gustavo
- 15 de Junio - 5 horas - Diseño de la base de datos hogar IoT. - Christopher y Juan Jose
- 16 de Junio al 19 de Junio - 20 horas - Llenado de la base de datos mediante scrips de Python - Gustavo
- 16 de Junio al 19 de Junio - 15 horas - Investigación de ETL - Juan Jose .
- 16 de Junio al 19 de Junio - 15 horas - Investigación Cluster .
- 17 Junio - 4 horas - Documentación del Cluster - Juan Jose .
- 18 de Junio al 19 de Junio - 15 horas - Realización del Cluster - Christopher y Juan Jose .
- 18 de Junio al 19 de Junio - 15 horas - Realización del ETL - Juan Jose .
- 19 Junio - 4 horas - Documentación - Todos .

Total de horas trabajadas por todos los estudiantes: 100 horas

## Comentarios finales (estado del programa)

---

El proceso de replicación de una base de datos fue el que más se dificultó, costó varios días de trabajo y mucha investigación, aparte de poner en práctica gran número de tutoriales. Finalmente se dejó al lado de las máquinas virtuales del cluster.

Se logró establecer las conexiones de ETL, sin embargo, finalmente daba problemas y no se podía obtener los datos de manera correcta.

La simulación e inserción de datos se logró de manera completa y satisfactoria. Todos los datos fueron insertados en las bases de datos correspondientes.

El trabajo es realmente largo y complejo, por lo que se recomienda empezar a investigar con bastante tiempo, ya que el trabajo contempla una curva de aprendizaje bastante alta.

## Conclusiones

---

Se estudió el problema de BI, replicación y clusterización de bases de datos. BI debía integrarse con distintas fuentes de datos, entre ellas, otras tablas de bases de datos y páginas / tablas de excel o csv.

Clusterización de bases de datos pretende manejar las bases de datos y tener los datos disponibles en distintos nodos del master. Clusterización se logró implementar de manera exitosa e incluso se realizó un tutorial/documento completo y extenso de cómo realizar el cluster en futuros proyectos usando LXD.

## Bibliografía

---

stackoverflow. (2015). *MySQL: Grant all privileges on database*. 17 junio 2017, de stackoverflow Sitio web: <https://stackoverflow.com/questions/5016505/mysql-grant-all-privileges-on-database>

Canonical Ltd.. (2017). *Installing LXD and the command line tool*. 17 junio 2017, de Canonical Ltd. Sitio web: <https://linuxcontainers.org/lxd/getting-started-cli/>

Randall Araya. (2016). *cluster-de-mysql-en-lxd*. 15 Junio 2017, de tec.siuva.ac.cr Sitio web: <https://tec.siuva.ac.cr/documentacion/cluster-de-mysql-en-lxd>

Aciddrop com. (2017). *step-by-step-how-to-setup-mysql-database-replication*. 15 Junio 2017, de [Aciddrop.com](http://aciddrop.com) Sitio web: <http://aciddrop.com/2008/01/10/step-by-step-how-to-setup-mysql-database-replication/>

Sawiyati. (2015). *Master - Slave MySQL Replication Tutorial for Newbie*. 15 Junio 2017 , de servermom Sitio web: <http://www.servermom.org/master-slave-mysql-replication-tutorial/>

Muhammad Arul. (2016). *How to Install a MySQL Cluster on Ubuntu 16.04*. 15 Junio 2017 , de howtoforge Sitio web: <https://www.howtoforge.com/tutorial/how-to-install-a-mysql-cluster-on-ubuntu-16-04/>