

Replicación de Bases de datos PostgreSQL utilizando Slony-I

Víctor Ostorga
vostorga@gentoo.org

Julio de 2010

II Encuentro Centroamericano de Software Libre

Tabla de contenidos

1 Introducción a Slony-I

Tabla de contenidos

- 1 Introducción a Slony-I
- 2 ¿Cómo funciona Slony?

Tabla de contenidos

- 1 Introducción a Slony-I
- 2 ¿Cómo funciona Slony?
- 3 Modelos de replicación

Tabla de contenidos

- 1 Introducción a Slony-I
- 2 ¿Cómo funciona Slony?
- 3 Modelos de replicación
- 4 Conceptos Importantes

Tabla de contenidos

- 1 Introducción a Slony-I
- 2 ¿Cómo funciona Slony?
- 3 Modelos de replicación
- 4 Conceptos Importantes
- 5 Requisitos previos

Intención de esta presentación

Intención

Mostrar las capacidades iniciales de replicación del Sistema Gestor de Bases de Datos PostgreSQL auxiliándose de herramientas extras, tales como lo es Slony.

¿Qué es Slony-I?

Definición

Slony-I es un sistema de replicación asíncrono para PostgreSQL de una base de datos maestra hacia múltiples bases de datos hijas.

¿Cómo funciona Slony-I?

Definición

Realiza las actualizaciones utilizando triggers, lo que significa que no puede propagar cambios de esquemas y operaciones con objetos. Actualmente Slony solamente replica cambios a Tablas y Secuencias.

Para quiénes no es Slony-I

- Sitios con redes altamente inestables

Para quiénes no es Slony-I

- Sitios con redes altamente inestables
- Bases de datos donde regularmente se hagan cambios a esquemas y definiciones de tablas (DDL)

Para quiénes no es Slony-I

- Sitios con redes altamente inestables
- Bases de datos donde regularmente se hagan cambios a esquemas y definiciones de tablas (DDL)
- Para quiénes quieran un sistema de replicación de múltiples maestros a hijos

Para quiénes no es Slony-I

- Sitios con redes altamente inestables
- Bases de datos donde regularmente se hagan cambios a esquemas y definiciones de tablas (DDL)
- Para quiénes quieran un sistema de replicación de múltiples maestros a hijos
- Para quiénes quieran un sistema de replicación que haga algo automáticamente al detectar el fallo de un nodo

Modelos de replicación

Definición

Actualmente existen muchos modelos de replicación, dentro de los cuales los siguientes son populares en las bases de datos:

Modelos de replicación

Definición

Actualmente existen muchos modelos de replicación, dentro de los cuales los siguientes son populares en las bases de datos:

- Replicación síncrona de un único origen a múltiples bases hijas

Modelos de replicación

Definición

Actualmente existen muchos modelos de replicación, dentro de los cuales los siguientes son populares en las bases de datos:

- Replicación síncrona de un único origen a múltiples bases hijas
- Replicación síncrona de múltiples orígenes a múltiples bases hijas

Modelos de replicación

Definición

Actualmente existen muchos modelos de replicación, dentro de los cuales los siguientes son populares en las bases de datos:

- Replicación síncrona de un único origen a múltiples bases hijas
- Replicación síncrona de múltiples orígenes a múltiples bases hijas
- Replicación asíncrona de múltiples orígenes

Conceptos Importantes antes de poner en marcha la replicación

Definición

Antes de poner en marcha un cluster de replicación, es importante tener claros los siguientes conceptos:

Conceptos Importantes antes de poner en marcha la replicación

Definición

Antes de poner en marcha un cluster de replicación, es importante tener claros los siguientes conceptos:

- Cluster

Conceptos Importantes antes de poner en marcha la replicación

Definición

Antes de poner en marcha un cluster de replicación, es importante tener claros los siguientes conceptos:

- Cluster
- Nodo

Conceptos Importantes antes de poner en marcha la replicación

Definición

Antes de poner en marcha un cluster de replicación, es importante tener claros los siguientes conceptos:

- Cluster
- Nodo
- Set de replicación

Conceptos Importantes antes de poner en marcha la replicación

Definición

Antes de poner en marcha un cluster de replicación, es importante tener claros los siguientes conceptos:

- Cluster
- Nodo
- Set de replicación
- Origen, proveedores y suscriptores

Conceptos Importantes antes de poner en marcha la replicación

Definición

Antes de poner en marcha un cluster de replicación, es importante tener claros los siguientes conceptos:

- Cluster
- Nodo
- Set de replicación
- Origen, proveedores y suscriptores
- Demonios slon

Conceptos Importantes antes de poner en marcha la replicación

Definición

Antes de poner en marcha un cluster de replicación, es importante tener claros los siguientes conceptos:

- Cluster
- Nodo
- Set de replicación
- Origen, proveedores y suscriptores
- Demonios slon
- Comandos slonik

Conceptos

- Cluster
Un cluster es un conjunto de bases de datos PostgreSQL dentro de las cuales sucede la replicación

Conceptos

- Cluster
Un cluster es un conjunto de bases de datos PostgreSQL dentro de las cuales sucede la replicación
- Nodo
Un nodo es una base de datos que participará en el proceso de replicación

Conceptos

- Cluster
Un cluster es un conjunto de bases de datos PostgreSQL dentro de las cuales sucede la replicación
- Nodo
Un nodo es una base de datos que participará en el proceso de replicación
- Set de replicación
Es un conjunto de tablas y secuencias que se replicarán.
Nótese que no se replican todas las tablas, sino que se tienen que especificar explícitamente.

Conceptos

- Origen, proveedores y suscriptores
Cada set de replicación tiene un nodo de origen, que es la base de donde se obtiene información. Los nodos que reciben la información se denominan nodos suscriptores. Dichos nodos suscriptores se pueden convertir en proveedores si a su vez tienen otros nodos suscriptores

Conceptos

- Origen, proveedores y suscriptores
Cada set de replicación tiene un nodo de origen, que es la base de donde se obtiene información. Los nodos que reciben la información se denominan nodos suscriptores. Dichos nodos suscriptores se pueden convertir en proveedores si a su vez tienen otros nodos suscriptores
- Demonios slon
Para cada nodo en el cluster, existe un proceso slon que administra la actividad de replicación para ese nodo.

Conceptos

- Origen, proveedores y suscriptores
Cada set de replicación tiene un nodo de origen, que es la base de donde se obtiene información. Los nodos que reciben la información se denominan nodos suscriptores. Dichos nodos suscriptores se pueden convertir en proveedores si a su vez tienen otros nodos suscriptores
- Demonios slon
Para cada nodo en el cluster, existe un proceso slon que administra la actividad de replicación para ese nodo.
- Comandos slonik
Es un pequeño intérprete de comandos que incluye comandos utilizados para la manipulación del cluster de replicación

Entonces, ¿qué es un cluster de slony?

Definición

Un cluster de Slony es una agrupación básica de instancias de bases de datos donde ocurre la replicación. Si el cluster se llama micluster, entonces slony creará al momento de inicializar cada nodo un esquema con el nombre `_micluster`, dentro del cual se crearán tablas que almacenan información de configuración y replicación.

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Cualquier plataforma que pueda ejecutar PostgreSQL puede correr en principio Slony-I

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Cualquier plataforma que pueda ejecutar PostgreSQL puede correr en principio Slony-I
- Utilizar **exactamente** la misma versión de Slony en todos los nodos

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Cualquier plataforma que pueda ejecutar PostgreSQL puede correr en principio Slony-I
- Utilizar **exactamente** la misma versión de Slony en todos los nodos
- En caso de compilar Slony
 - GNU make

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Cualquier plataforma que pueda ejecutar PostgreSQL puede correr en principio Slony-I
- Utilizar **exactamente** la misma versión de Slony en todos los nodos
- En caso de compilar Slony
 - GNU make
 - Compilador ISO/ANSI C

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Cualquier plataforma que pueda ejecutar PostgreSQL puede correr en principio Slony-I
- Utilizar **exactamente** la misma versión de Slony en todos los nodos
- En caso de compilar Slony
 - GNU make
 - Compilador ISO/ANSI C
- Versión 8.3 o superior de PostgreSQL para utilizar Slony-I 2.0.4

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Cualquier plataforma que pueda ejecutar PostgreSQL puede correr en principio Slony-I
- Utilizar **exactamente** la misma versión de Slony en todos los nodos
- En caso de compilar Slony
 - GNU make
 - Compilador ISO/ANSI C
- Versión 8.3 o superior de PostgreSQL para utilizar Slony-I 2.0.4
- Codificación similar de la base de datos (UNICODE, SQL_ASCII, LATIN1)

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Sincronización de hora de sistemas operativos donde estén las bases

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Sincronización de hora de sistemas operativos donde estén las bases
- Red Estable

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Sincronización de hora de sistemas operativos donde estén las bases
- Red Estable
- Tablas a replicar con llaves primarias (asegurar una forma unica de identificar filas)

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Sincronización de hora de sistemas operativos donde estén las bases
- Red Estable
- Tablas a replicar con llaves primarias (asegurar una forma unica de identificar filas)
- `tcpip_socket=true` en el archivo `postgresql.conf`

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Sincronización de hora de sistemas operativos donde estén las bases
- Red Estable
- Tablas a replicar con llaves primarias (asegurar una forma unica de identificar filas)
- `tcpip_socket=true` en el archivo `postgresql.conf`
- Accesos correctos en `pg_hba.conf`

Requisitos

Definición

Requisitos importantes antes de configurar la replicación

- Sincronización de hora de sistemas operativos donde estén las bases
- Red Estable
- Tablas a replicar con llaves primarias (asegurar una forma unica de identificar filas)
- `tcpip_socket=true` en el archivo `postgresql.conf`
- Accesos correctos en `pg_hba.conf`
- Utilizar un ambiente de pruebas antes de realizarlo en producción

Replicando la primera base de datos

Datos necesarios

En la demostración configuraremos la replicación entre dos bases de datos, para lo cual utilizaremos la siguiente información:

- NOMBRECLUSTER=ecsl2010
- BDMAESTRA=maestra
- BDHIJA=hija
- USUARIOREPLICACION=slony