

Taller de Administración de Servidores Linux CC5308

Clase 14: 12 de Abril de 2011

Base de Datos: SQL y Encoding

José Urzúa
jose@nic.cl

Copyright © 2011
Creative Commons 3.0-cl by-nc-sa

Temario

- Bases de Datos
 - Políticas de implementación
 - ANSI SQL, SQL Propietario
 - Representaciones de datos
- Encoding
 - Definiciones
 - Decisiones en bases de datos

Políticas Implementación

- Qué base de datos?
 - Nuevos sistemas desarrollados internamente
 - Sistemas que se adquieren
- Qué versión?
 - Actualización de sistemas
 - Mejoras de seguridad, performance, etc
- Actualizaciones
 - Changelog

Políticas Implementación

- Accesos
 - Usuarios y roles
 - Otros sistemas y hosts
- Respaldos
 - Periodicidad (bajar servicios?)
 - Ubicación servidores
 - Recuperación ante fallos
- Monitoreo
 - Performance, disponibilidad, logs

Políticas Implementación

- Seguridad
 - Políticas de acceso
 - Seguridad sobre datos: claves, certificados digitales, n° tarjetas, etc.
- Trabajo con desarrolladores
 - ¿Compresión de datos?
 - Disco
 - Red
 - Bases de datos no solo para datos: procedimientos, funciones, cálculos..

Políticas Implementación

- Trabajo con desarrolladores
 - Archivos binarios
 - Almacenar en BD?
 - Almacenar PATH al archivo?
 - Scripts creación, deploy, testing sobre BD
 - Separación ambientes: desarrollo, prototipo, rc, producción

SQL

- ANSI SQL

- Lenguaje de acceso a base de datos relacionales
- Permite hacer consultas y cambios sobre los datos

- Motores de datos y ANSI SQL

- MySQL:

<http://dev.mysql.com/doc/refman/5.0/en/differences-from>

- PostgreSQL

<http://www.postgresql.org/docs/9.0/interactive/unsupported>

SQL

- Representaciones de datos

- Tipos de datos de otras bases de datos en MySQL

<http://dev.mysql.com/doc/refman/5.0/en/other-vendor-data-types.html>

- Ejemplo: Datos Binarios

- MySQL: Binary, VarBinary, TinyBlob, BLOB, MediumBlob, LongBlob

<http://dev.mysql.com/doc/refman/5.0/en/storage-requirements.html>

- PostgreSQL: bytea

<http://www.postgresql.org/docs/9.0/interactive/datatype-bytea.html>

Encoding

Encoding

- No existe el texto plano
 - Siempre va acompañado de una codificación: ASCII, UTF-8, ISO-8859-1, etc
- Conjunto de caracteres usados en Internet:
<http://www.iana.org/assignments/character-sets>
- Unicode
 - Provee un número único a cada caracter, sin importar la plataforma, programa e idioma
<http://www.unicode.org/>
 - UTF-8, UTF-16, UTF-32

Encoding

- Ejemplo:
 - Servidor HTTP enviando encoding de sitios web a browser
 - Revisar cabeceras
 - Comando iconv: convertir de un encoding a otro
 - Revisar archivos iso8859-1.txt y utf-8.txt

Encoding

- ¿Qué encoding usar en un servidor?
 - Sistemas operativos, aplicaciones
- ¿Qué encoding usar para el servidor de base de datos?
 - Múltiples encoding?
- ¿Qué encoding usar para cada base de datos?
 - MySQL: `mysql> show collation;`
 - Postgres:
`createdb -E ISO-8859-1 -T template0 --lc-collate=es_CL.iso88591 -lc-ctype=es_CL.iso88591..`

Encoding

- ¿Qué encoding usar para cada tabla?
 - MySQL:

```
CREATE TABLE t1 ( ... )  
CHARACTER SET latin1 COLLATE latin1_danish_ci;
```
- ¿Qué encoding usar para cada columna?
 - Ojo con llaves foráneas!!
- Encoding en terminales clientes
 - Siempre revisar en que encoding está trabajando el shell del operario!

Encoding

- Consideraciones para una aplicación web
 - Encoding sistema operativo
 - Encoding servidor
 - Encoding base de datos
 - Encoding aplicación (Java, PHP, Python, etc)
 - Encoding servidor aplicaciones (Tomcat, Jboss, Glassfish, etc)
 - Encoding servidor web (Apache)
 - Encoding navegador (cliente)
 - Encoding sistema operativo (cliente)

Actividad Sugerida

- Cree una base de datos con al menos una tabla que permita almacenar datos binarios. Haga **varias** pruebas de almacenar los datos comprimidos y no comprimidos, compare el espacio utilizado y el tiempo extra que toma el software al comprimir-descomprimir ¿Qué conviene más?
- Compare el caso de almacenar una foto y un archivo XML