

# PostgreSQL: cuando "suficientemente bueno" no es "suficiente"

Jaime Casanova  
jaime@2ndQuadrant.com

8 de octubre de 2011

¿Quién soy yo?

## Permítanme presentarme

- Miembro de la comunidad de PostgreSQL desde el 2004
- Contacto Regional de PostgreSQL para el Ecuador
- Beta-tester
- Revisor de parches
- Administrador del Grupo de Usuarios de PostgreSQL del Ecuador: [ecpug@postgresql.org](mailto:ecpug@postgresql.org)
- Consultor y representante en Ecuador de la empresa 2ndQuadrant Ltd

# Tabla de Contenido

1 Introducción

2 Enova Financial

3 Mind Candy

# ¿Sobre qué es esta charla?

- Todo software tiene límites
- Los casos de estudio

# Los límites de PostgreSQL

Límites impuestos por:

- ... el sistema operativo
- ... el hardware
- ... el diseño de PostgreSQL

# Tabla de Contenido

1 Introducción

2 Enova Financial

3 Mind Candy

# ¿Qué hace Enova?



- es una financiera
- ... con clientes en Australia, Canadá, Estados Unidos y Reino Unido
- cada año provee mas de \$1 billón en préstamos

# PostgreSQL en Enova

- Fundada en 2004
- Empezo a brindar servicios financieros en 2005
- Migro de MySQL a PostgreSQL en 2006
- ... sin MySQL desde el 35 de Agosto!

# PostgreSQL en Enova

- Sus sistemas principales están en PostgreSQL
- ... y muchos de los internos también

Perder el servicio durante el día, costaría 100mil por hora

# La elección de la mejor base de datos

Jim Nasby - 1999

“All databases suck, they just suck in different ways.”

# La elección: MySQL

- Rails trabajaba con MySQL de forma predeterminada (en ese tiempo)
- "Todo el mundo usa MySQL"
- Es fácil de usar y rápidamente se tuvo un negocio en funcionamiento

# La elección: MySQL

- Problemas de integridad de datos
- Escalabilidad
- MySQL hace muy fácil hacer las cosas mal
  - Tipo de tabla incorrecta
  - Silenciosa "corrección" de datos

Si me aprobaron un préstamo de \$1200, ¿por qué solo me dan \$200?

# A elegir otra vez

- MS SQL Server: No trabajaba bien con Rails
- DB2: Muy cara
- Oracle: Insanamente cara
  - \$100mil solo para empezar!

# La nueva elección: PostgreSQL

- Puede ser frustrante al inicio
- PostgreSQL "no perdona"
- Encontrar gente que sepa no es tan fácil

# Lecciones aprendidas

- Probar todos los SQL
- Usar Foreign Keys
- Si tienes éxito, tendrás problemas de escalabilidad... estén listos
- Eventualmente quedarás atado a la tecnología que estas usando

# ¿Donde está Enova hoy?

- La base OLTP de US tiene 1.5TB
- En promedio 640 tps
- 4000 tps en horas pico
- El conjunto de datos en el que se trabaja normalmente esta entre 100 y 200GB
- 2 réplicas para reportes

# ¿Donde está Enova hoy?

El hardware puede ayudar

Sept 2007 – 300GB, 4x dual-core, 32GB

Oct 2008 – 800GB, 4x 4-core, 96GB

Nov 2009 – 1TB, 4x 6-core, 192GB

Nov 2010 – 1.5TB, 8x 6-core, 512GB

# Cosas que Enova usa de PostgreSQL

- herencia
- restricciones CHECK y triggers (para integridad de datos)
- plperl
- PgQ

# Cosas que Enova quiere mejorar de PostgreSQL

- restricciones CHECK y triggers (para integridad de datos)
- Financio a Command Prompt para desarrollar CHECK NOT VALID

Jim Nasby - 2011

.^All databases still suck in different ways... but it's damn nice using a database that is created by an open community where you can actually get things that suck changed to suck less!"

# Tabla de Contenido

1 Introducción

2 Enova Financial

3 Mind Candy

# ¿Qué hace Mind Candy?

Mind Candy es una de las principales compañías en el mundo de juegos sociales y multi-jugador, ayudando a los niños (y los niños grandes!) en todo el mundo a jugar y comunicarse.

# ¿Qué hace Mind Candy?



# PostgreSQL en Mind Candy

- Fundada en 2003
- Empezo a desarrollar Moshi Monsters en 2007
- Crecio de unos cuantos miles a más de 50 millones de usuarios en solo 3 años
- ... de los que casi 30 millones están en Reino Unido

# El problema

- Imaginen 50 millones de niños regresando de la escuela
- ... chequeando el estado de mascota virtual, alimentandola, compartiendo, etc

# La solución

- mejor hardware, probado
- particionamiento mas eficiente
- mejores índices

# Los discos al ataque

- Quien a medido el tiempo de respuesta de consultas en PostgreSQL sabe que a veces aun las consultas mas rapidas se ponen lentas
- Los checkpoints deberían aliviar la carga de I/O no empeorarla
- ... pero la etapa final de un checkpoint puede ser un problema
- agregar mas memoria solo empeora el problema, agregar mas discos también

# Soporte que si resuelve el problema

- ¿Cual es el problema?
- Modifiquemos PostgreSQL para que nos de más información
- ...hey! esto es útil que tal si lo contribuimos!
- Oh! es el fsync al final del checkpoint
- ...modifiquemos PostgreSQL para que distribuya eso
- ...hey! esto es útil que tal si lo contribuimos!

# Los beneficios

- Mind Candy hace a felices a millones de niños
- El negocio es rentable
- Muchos otros usuarios de PostgreSQL se beneficiarán de las mejoras en 9.1

# Preguntas

¿Preguntas?