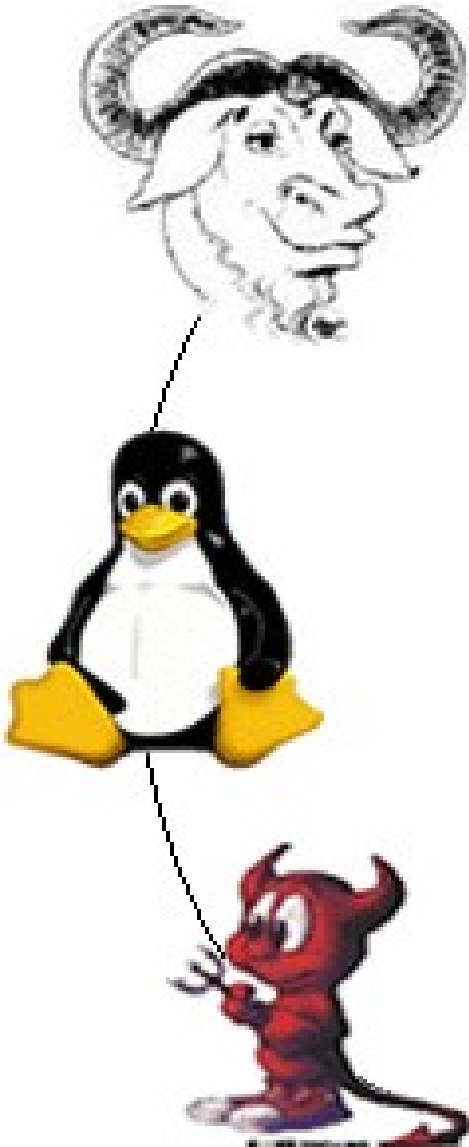


APESOL

Asociación Peruana
de Software Libre



*1eras Jornadas Técnicas
con el Estado*

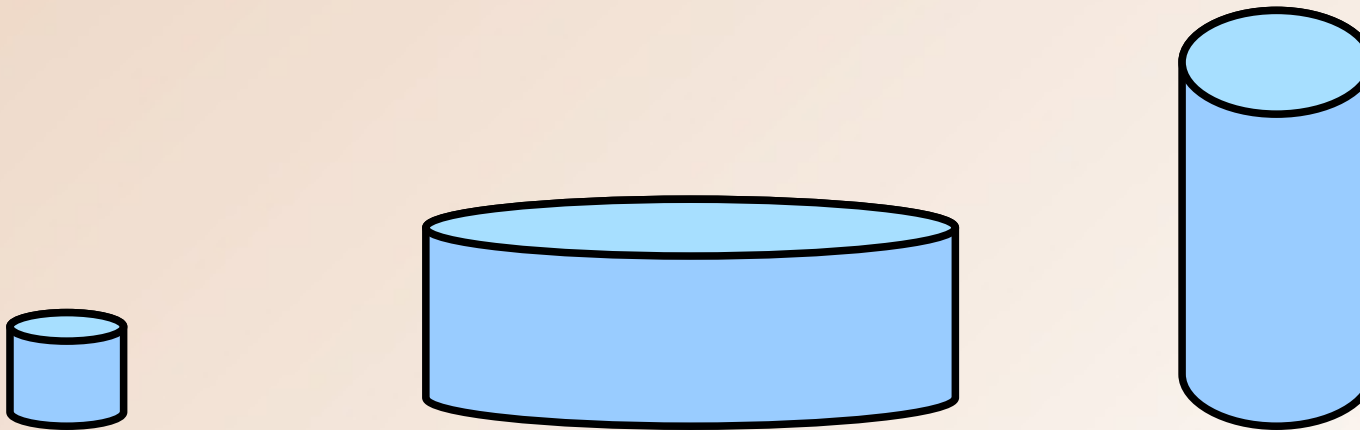
DBMS Libres en entornos de Alta Demanda

*Ernesto Quiñones Azcárate
Presidencia de Apesol
2006-2008
ernestoq@apesol.org*

 **APESOL**
Asociación Peruana de Software Libre

Agenda

1. Porque Mysql y porque PostgreSQL
2. Escalabilidad
3. Conclusiones



1. Porque Mysql y porque PostgreSQL

MySql

- Entre octubre del 2005 y Marzo del 2006 solo la versión 5.0 contabilizo 6 millones de descargas.
- Con modificaciones especiales Google la usa para su buscador (puede indexar 20 Pbytes en 1 solo día)
- Sun compro MySql por 1,000 millones de dólares.

PostgreSQL

- La versión 8 a registrado a Julio del 2006 mas de un millón de descargas
- Con modificaciones especiales Yahoo reclamo el poseer la DB viva mas grande del mundo 2Pbytes.
- IBM acaba de invertir 10 millones de dólares en una versión comercial de PostgreSQL.

DB4

- Mas de 200 millones de implementaciones
- Embebida
- Oracle compro SleepyCat el proveedor de la herramienta.

1. Porque Mysql y porque PostgreSQL

MySql

- Sun esta migrando su estrategia de apoyo a PostgreSQL (pero no abandonándolo) por MySql, el soporte certificado de Solaris y el hardware de Sun puede ser su gran despegue.
- Cuenta con mucho apoyo de las distribuciones Enterprise (RedHat especialmente)

PostgreSQL

- Ya cuenta con muchas soluciones de alta demanda certificadas por Sun.
- IBM a mirado hacia PostgreSQL como su posible alternativa de ingresar a los DBMS libres, pero piensa liberar DB2.
- Fujitsu a certificado infraestructura para alta demanda con PostgreSQL.

1. Porque Mysql y porque PostgreSQL

- El mercado tiene bastante know-how en MySQL y empezando a crecer en PostgreSQL.
- PostgreSQL a entrado fuerte en soluciones GIS en el mercado local.
- Se adapta al parque nacional de computadoras.
- Soporte local o cercano es posible de contratar, ¿oficial?, si tienes el presupuesto SI.

2. Escalabilidad

U
S
U
A
R
I
O
S

- Una ventaja en la escalabilidad de las dbms libres es que los productos para “crecer” son libres también y hay varias opciones, los proyectos suelen recomendar algunas soluciones.
- Si necesitas crecer, la inversión en infraestructura (hardware) que debes planificar debe tener el mismo nivel de inversión de una solución privativa, las dbms libres consumen menos recursos pero no son una forma de reciclar hardware viejo.

DATA

2. Escalabilidad

MySql

- Mas rápido leyendo información.
- Aspira a ser ACID certificado.
- Ofrece diferentes engines para administrar la data, diferentes niveles de integridad de los datos.
- Concurrencia de usuarios buena solo si no procesan consultas pesadas.

- Mejor en Linux.
- Incluye en su versión base herramientas para escalar.

- Soporte comercial del mismo proyecto.

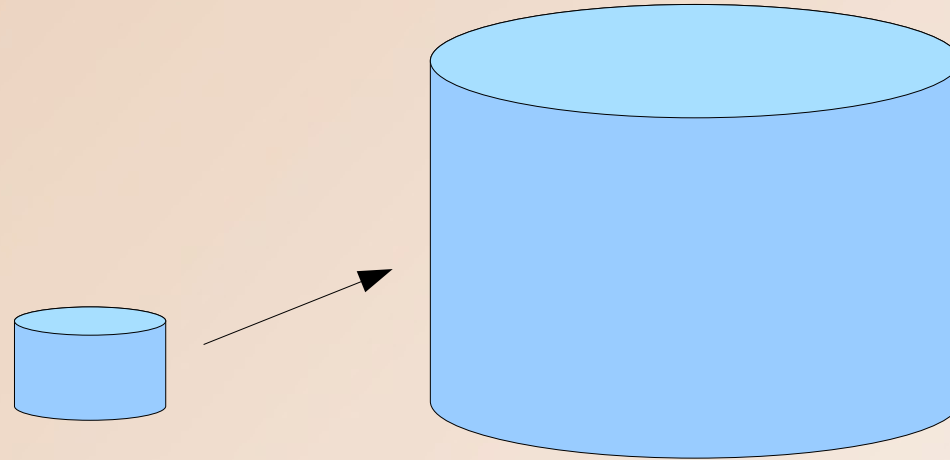
PostgreSQL

- Mas rápido en transacciones.
- ACID certificado.
- Un solo engine, integridad de los datos mayor.

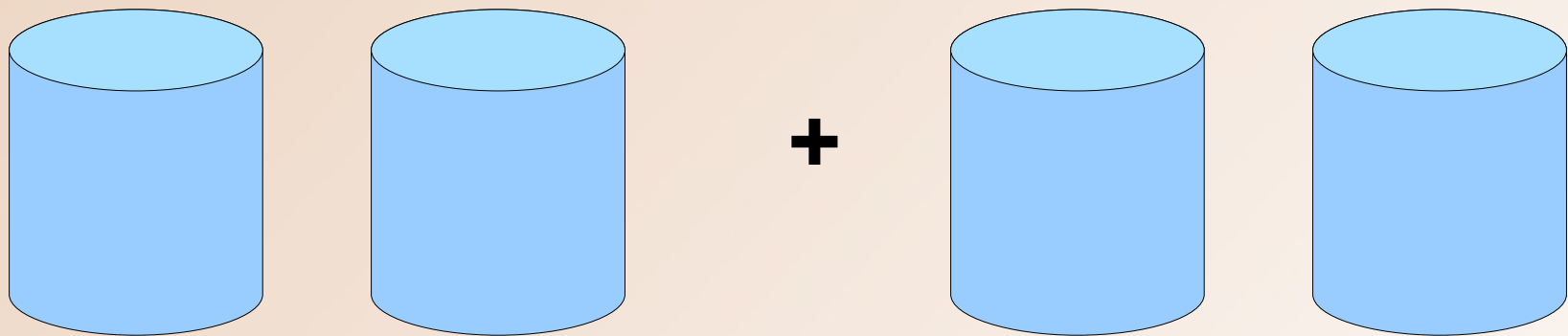
- Buena administración de concurrencia de usuarios incluso en transacciones/consultas pesadas.
- Mejor en Unix (libres o privativos).
- No incluye herramientas para escalar propias del proyecto, todas son proporcionadas por otros proyectos.
- El proyecto no da soporte comercial, o al menos no lo ofrece directamente.

<http://en.wikipedia.org/wiki/ACID>

2. Escalabilidad



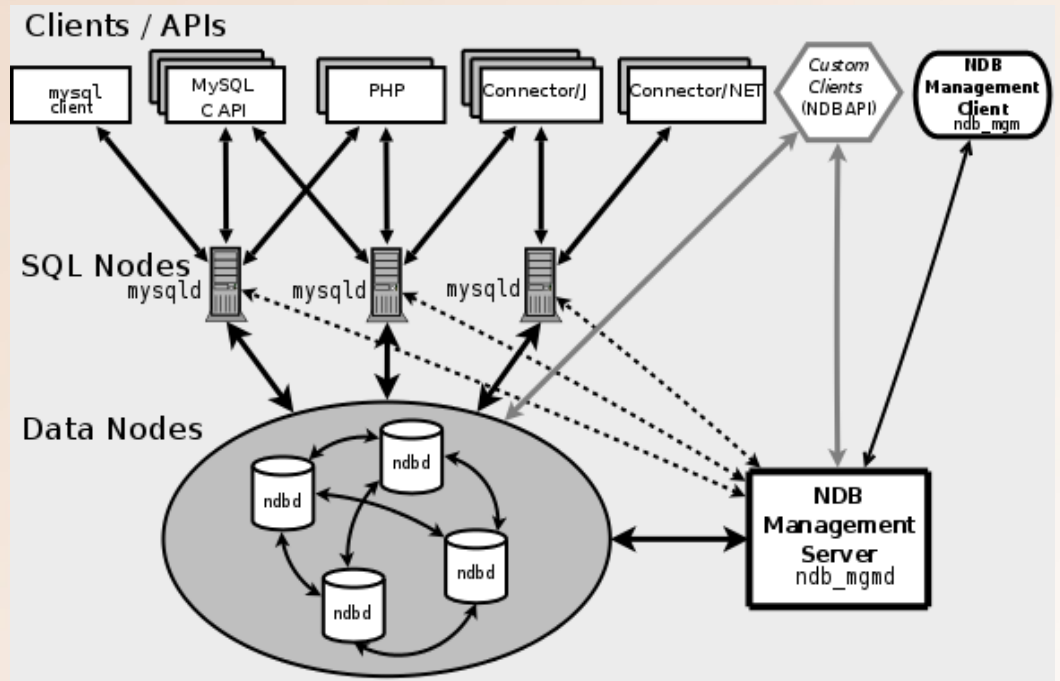
Escalar Verticalmente



Escalar Horizontalmente

2. Escalabilidad MySql

- La escalabilidad en MySQL es siempre asociada con la capacidad de escalar una infraestructura web.
- Para alta cantidad de peticiones a ciertas tablas considerar usarlas con el motor “MEMORY STORAGE” para que estas suban a memoria física totalmente (obviamente ponerle mucha ram al servidor)
- MySQL provee su propio esquema de clusterización, provee del engine NDB especialmente diseñado para estos casos.
- En un esquema más básico podemos replicar bases de datos y balancear la carga desde la aplicación.
- Cuanto más poderoso sea el servidor conviene ajustar más los valores del servidor para aceptar más usuarios y asignar la cantidad exacta de recursos a nuestras peticiones.



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

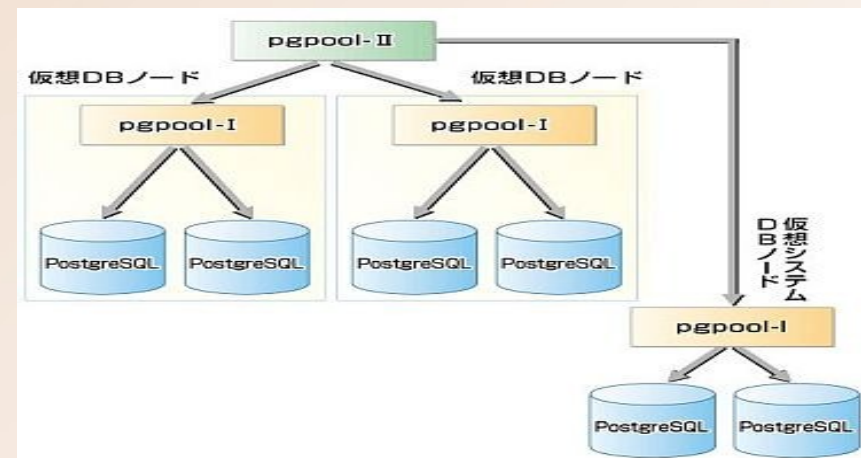
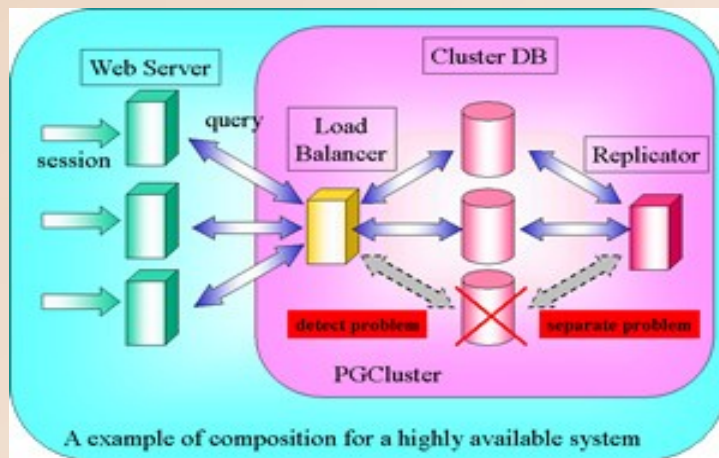
<http://dev.mysql.com/doc/refman/5.1/en/mysql-cluster.html>

<http://dev.mysql.com/doc/refman/6.0/en/replication.html>

<http://dev.mysql.com/doc/refman/6.0/en/optimizing-the-server.html>

2. Escalabilidad PostgreSQL

- A diferencia de MySQL el proyecto PostgreSQL no ofrece por “defecto” herramientas para escalar, sin embargo “da fe” de muchos proyectos de buen que apoyan a la escalabilidad (especialmente horizontal) como PI/Proxy, PgCluster, Slony y PgPool
- Un tendencia interesante es que entre cambios de versiones en la rama 8.x la mejora en el rendimiento a crecido drásticamente, lo cual hace conveniente que PostgreSQL siempre este actualizado, por ejemplo entre la versión 8.2 a las 8.3 la velocidad de procesamiento subió en un 30% en un mismo hardware.
- Sun publica documentación sobre entornos de multicores y multiterabytes con PostgreSQL.



http://wiki.postgresql.org/wiki/Replication,_Clustering,_and_Connection_Pooling

http://blogs.sun.com/jkshah/resource/pgcon_problems.pdf

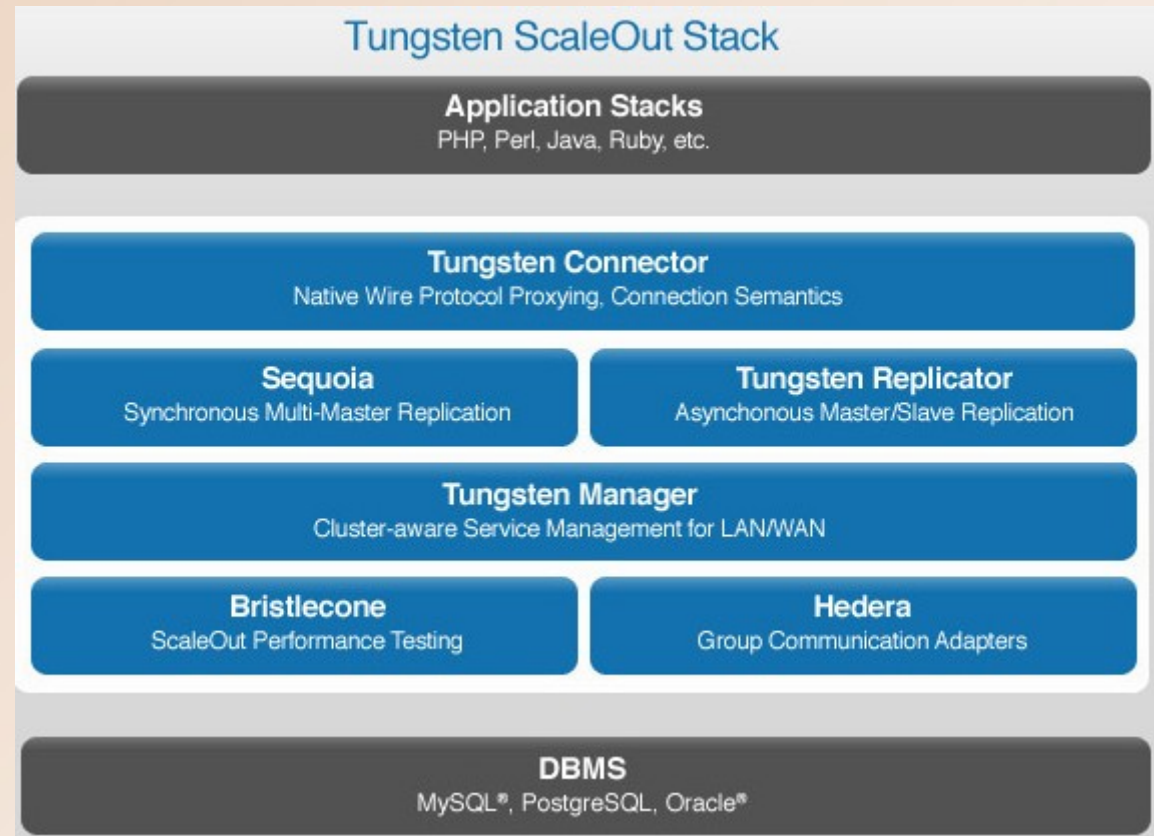
http://blogs.sun.com/jkshah/entry/glassfish_v2_and_postgresql_8

2. Escalabilidad ¿Qué mas hay?

<http://www.continuent.org/>

Continuent es un proyecto que esta enfocado en crear productos que permitan una alta escalabilidad en bases de datos heterogeneas.

Toda la arquitectura desarrollada en Java.

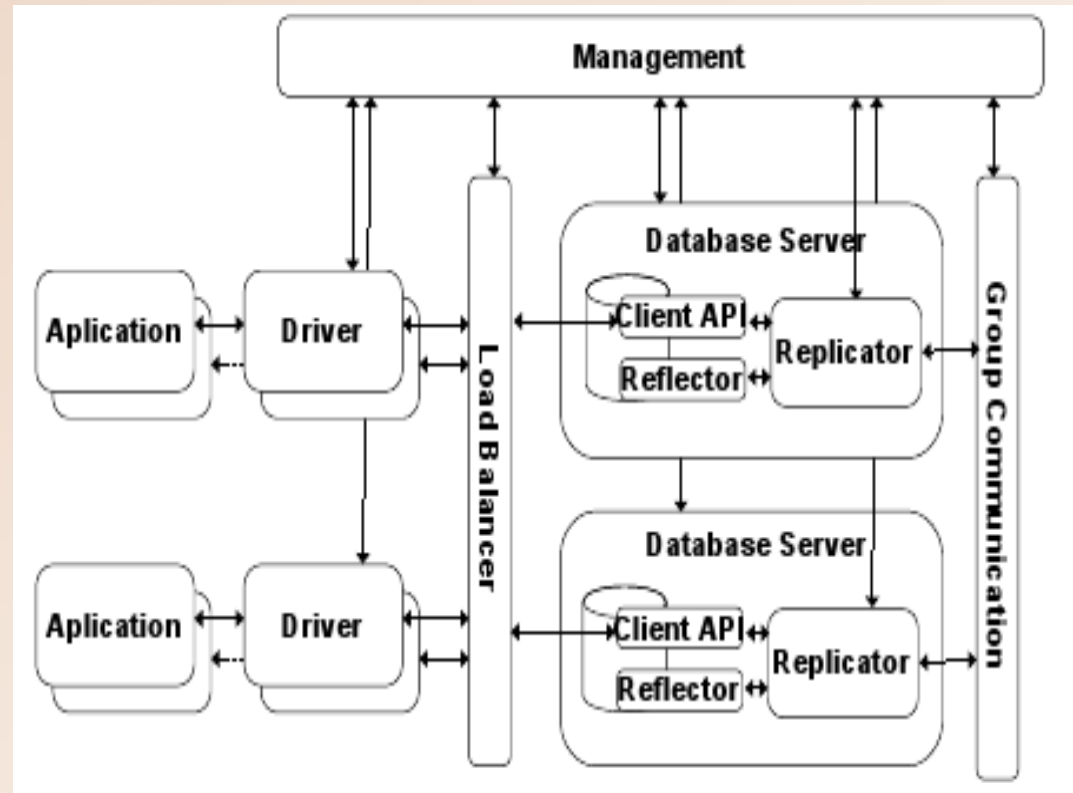


2. Escalabilidad ¿Qué mas hay?

<http://escada.sourceforge.net/>

Escada es un sistema de replicación de alta performance pensado en escenarios donde la infraestructura se encuentra distribuida en lugares remotos como prevención de desastres.

Usa internamente la arquitectura propuesta por Gorda (<http://gorda.di.uminho.pt/>) un proyecto enfocado en estandarizar los protocolos de replicación de datos.



2. Escalabilidad ¿Qué mas hay?

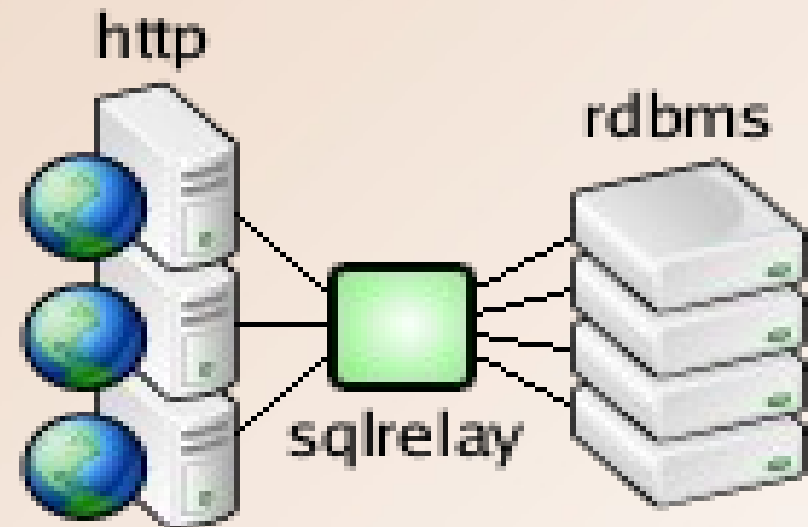
<http://sqlrelay.sourceforge.net/>

SQL Relay es una herramienta que ofrece balanceo de carga y replicación.

Muy orientada a servir en aplicaciones de desarrollo web y en entornos que requieren migración de dbms.

Soporta los siguiente DBMS:

- Oracle
- MySQL
- PostgreSQL
- Sybase
- Microsoft SQL Server 2000 Desktop Engine
- DB2
- Interbase/Firebird
- SQLite
- Microsoft Access



3. Conclusiones

- **No hay razón referida a volumen de datos o concurrencia de peticiones que permita descartar con una arquitectura totalmente libre de base de datos.**
- **Los volúmenes de datos que se manejan a nivel nacional no están siquiera ligeramente cercanos a los volúmenes de datos que pueden manejar las dbms libres presentadas. Yahoo menciona en referencia al uso de PostgreSQL con 2 Petabytes (volumen de datos vivos) soportando 24 mil millones de eventos en un día: “By comparison, large enterprise databases typically grow no larger than the tens of terabytes.” ¿cuantos de los presentes administran una db de decenas de terabytes?.**
- **Existen herramientas para probar las soluciones antes de pensar en “migrar”, y no generan costo de adquisición, solo de investigación.**

<http://tinyurl.com/68avgt>

3. Conclusiones

- **Existe soporte local, no es abundante aún porque la demanda no es masiva, pero eso esta empezando a cambiar. Si manejas una infraestructura tan grande o tan importante que requieres un soporte “clase A++” puedes contratar soporte de desarrolladores de los proyectos en países cercanos (Postgresql-Chile, Mysql-Brasil), ¿puede ofrecerte eso una dbms privativa?**
- **El soporte en dbms libres sigue los costos del mercado, piensa que lo que gastas en un profesional en Oracle, DB2, Informix, etc. Para tu dimensión de infraestructura es lo mismo que vas a tener que gastar en soporte en un profesional de calidad en MySql o PostgreSQL.**
- **Para saber cual es la herramienta correcta asesórate bien, visita los foros, entra a los canales de irc, pregunta, comparte tus dudas, no creas en lo que te dice el vendedor a ojos cerrados.**

Gracias

Toda la información con la que se realizo estas diapositivas puede ser vista aquí:

<http://tinyurl.com/6x9v6a>

APESOL

Website: <http://www.apesol.org>

IRC: server: *irc.freenode.net* sala: *#apesol*

Listas: <http://apesol.org/listas.php>