PalominoDB

Proven Database Excellence

PostgreSQL 9.1
FOSDEM 2012

Presentado por:
Emanuel Calvo @3manuek
DBA
www.palominodb.com

About the author and the talk:

- Operational DBA at PalominoDB
- Consultor Principal en Helpame
- DBA Mysql y Oracle en Aedgency.

- Twitter: @3manuek
- Mail: emanuel@palominodb.com

- The main source of this presentation is the following article:
    - http://palominodb.com/blog/2011/09/28/postgresql-91-part-1-general-features

New features

- Sychronous replication
- Foreign Data Wrappers (AKA SQL/MED)
- Collation per column
- Serializable Snapshot Isolation
- Unlogged tables
- Writable Common table Expressions
- KNN for GiST and GIN
- SE-LINUX support
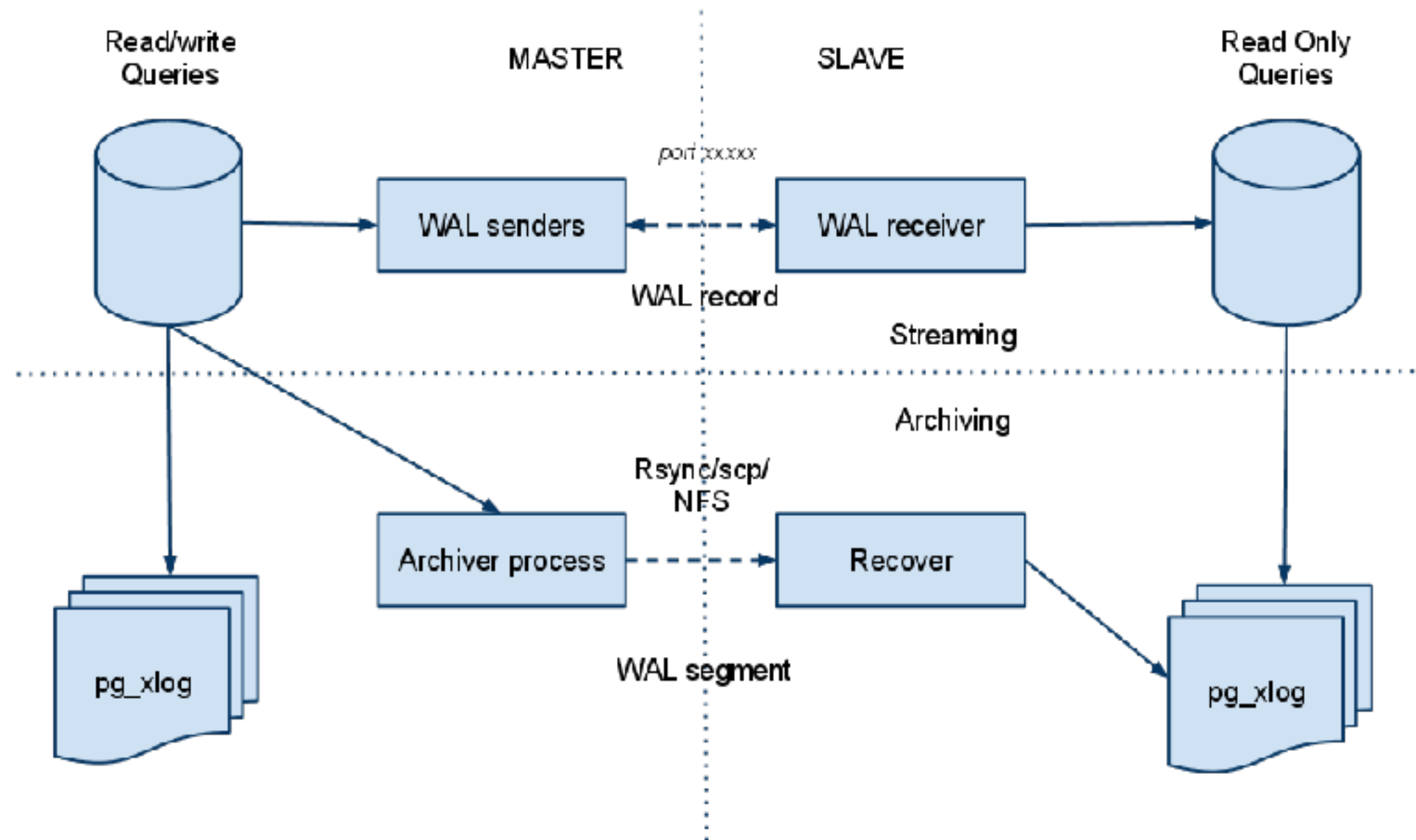- Upgrade Python version server side
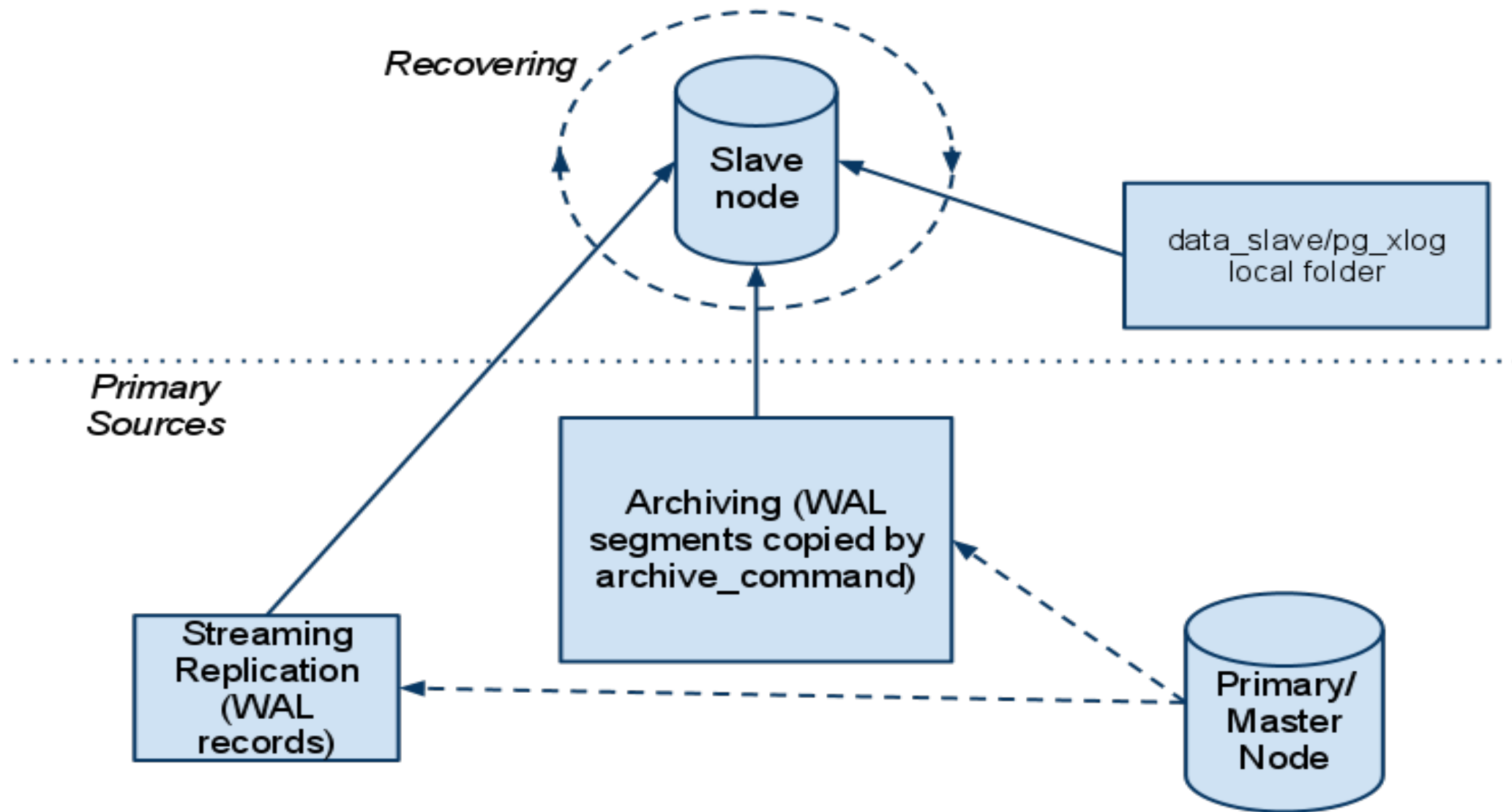- PGXN
- More stuff...
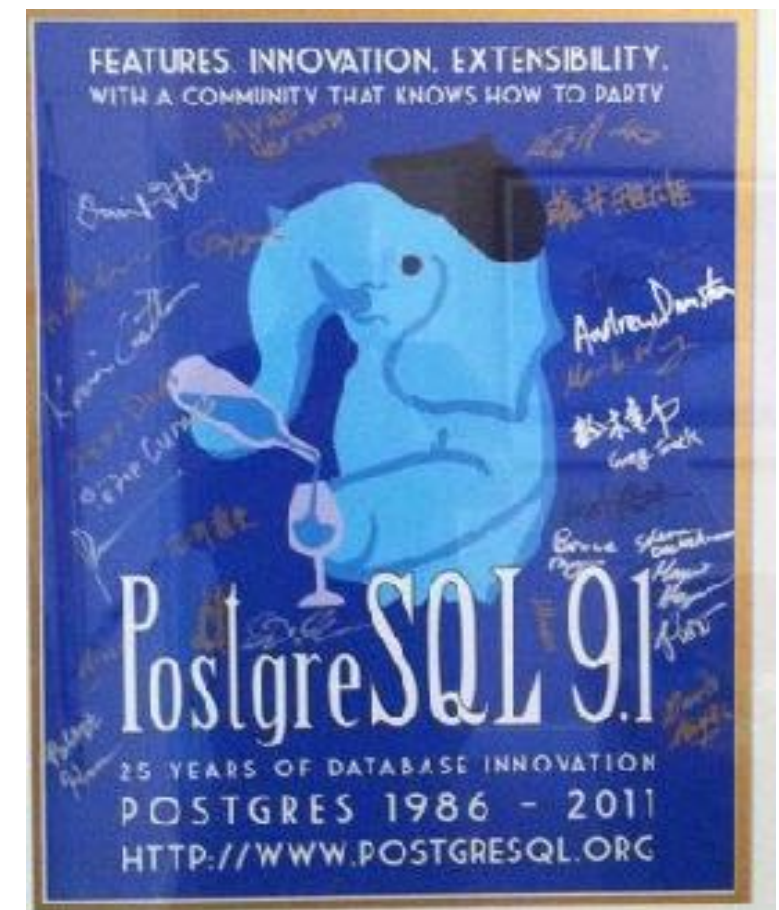
Synchronous Replication

- Control synchronicity per transaction switching synchronous_commit.
- Improvements on the 9.0 replication.
- Administrative functions to resume, pause and show state of the slave
    - pg_xlog_replay_pause()
    - pg_is_xlog_replay_paused()
    - pg_xlog_replay_resume()
    - pg_last_xact_replay_timestamp()
    - Pg_last_xact_replay_location()
- pg_ctl now supports "promote" for the slave
- Clone your slave without downtimes or locking (pg_basebackup)
- Information about slaves into the master (pg_stat_replication).
- pg_stat_database_conflicts Shows information related queries and vacuum conflicts between master and slave.

Replication basics
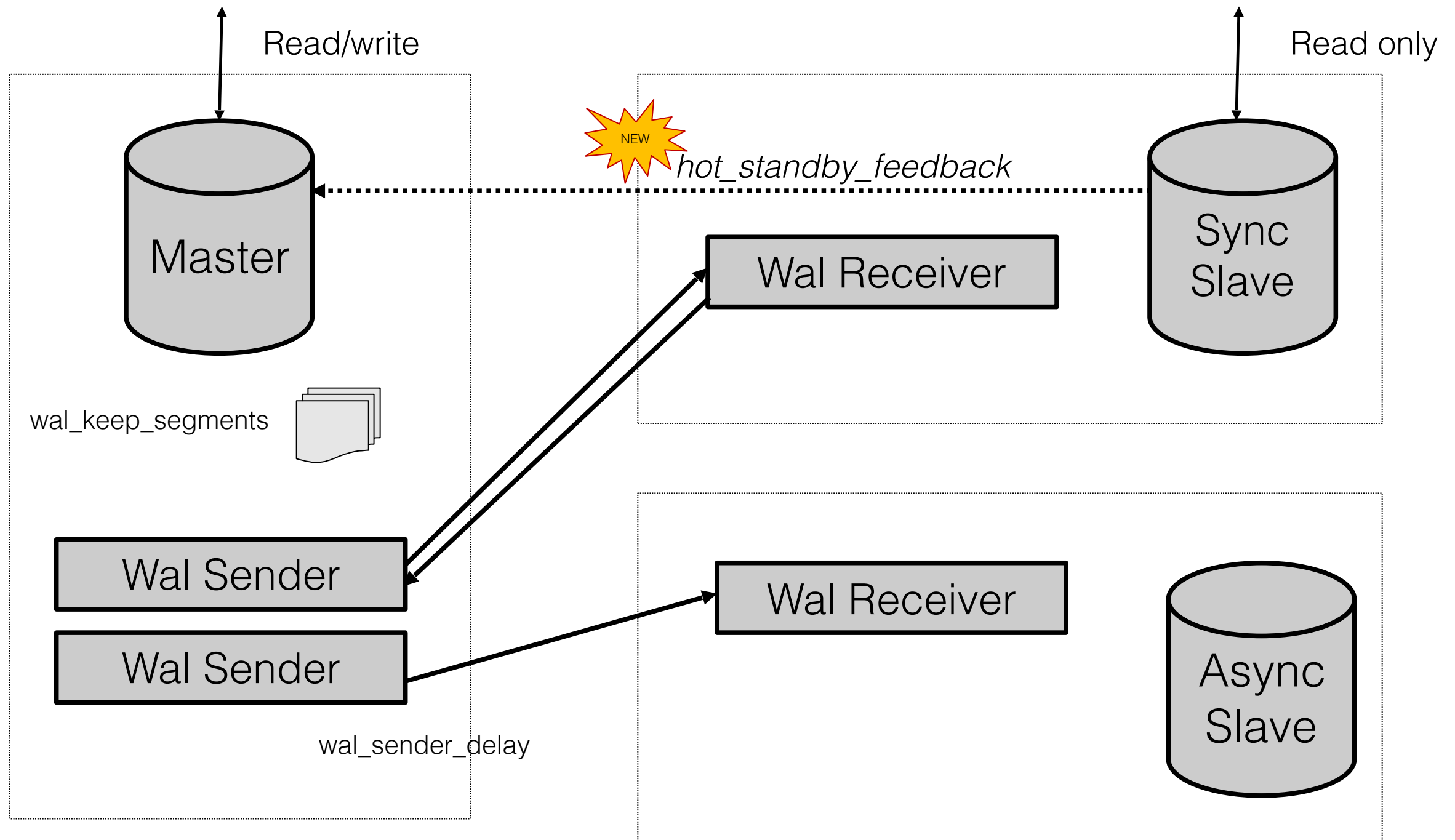
Standby sources

Synchronous replication (deeper scope)

Read/write

Read only

NEW

hot_standby_feedback

Master

Sync
Slave

Wal Receiver

wal_keep_segments

Wal Sender

Wal Sender

Wal Receiver

Async
Slave

wal_sender_delay

Replication limits

- Only one server can be synchronous.
- Don't support multi master.
- Don't support partial replication.
- Don't support distributed data.

Foreign Data Wrapper
(MED – Management of External Data)

```
CREATE EXTENSION file_fdw;

CREATE SERVER pglog FOREIGN DATA WRAPPER file_fdw;

CREATE FOREIGN TABLE gaiyin (
    nombre text,
    numero integer,
    letra char(1)
) SERVER pglog
OPTIONS ( filename 'c:\temp\file.csv', format 'csv' );

postgres=# explain select * from gaiyin;
                QUERY PLAN
------------------------------------------------------------
 Foreign Scan on gaiyin  (cost=0.00..1.10 rows=1 width=44)
   Foreign File: c:\temp\file.csv
   Foreign File Size: 36
(3 filas)
```

http://wiki.postgresql.org/wiki/SQL/MED

PALOMINODB   Proven Database Excellence

Foreign Data Wrapper - mysql_fdw

apt-get install libmysqld-dev

* Compilar como un contrib (make; make install)

CREATE EXTENSION mysql_fdw;

CREATE SERVER mysql_srv
FOREIGN DATA WRAPPER mysql_fdw
OPTIONS (address 'localhost', port '3306');

CREATE USER MAPPING
 FOR PUBLIC SERVER mysql_srv
OPTIONS (username 'root',password 'postgres');

CREATE FOREIGN TABLE prueba_my
(texto text, numero int, floaty float)
SERVER mysql_srv OPTIONS (table 'palomino.prueba');

CREATE FOREIGN TABLE prueba_my (
id integer,
name text,
address text)
SERVER mysql_svr
OPTIONS (table 'palomino.prueba');

Fuente: http://pgsnake.blogspot.com/2011/08/postgresql-91-meet-mysql.html

Foreign Data Wrappers (other plugins)

Oracle
Mysql
OBDC
Redis
CouchDB
Filetext
Twitter
LDAP

…

http://wiki.postgresql.org/wiki/Foreign_data_wrappers
http://repositorio.siu.edu.ar/trac/postgresql/wiki/Dblink

Per column collation

CREATE TABLE prueba_texto (
  _id serial PRIMARY KEY,
  texto_en text COLLATE 'en_US',
  texto_es text COLLATE 'es_ES'
);

SELECT
  texto_es < ('rendevouz' COLLATE"es_ES")
  FROM prueba_texto;

SELECT
  texto_es COLLATE "en_EN" < texto_en
  FROM prueba_texto;

CREATE INDEX ON prueba
  (texto COLLATE "en_US);

- < 8.4 the collation was strongly configured with the SO, being less flexible.
- In this version we can declare type columns and specify a different collation for each one.
- You can create indexes with a specific collation
  - Be careful, if you use another collation in the querie, the index will not be used.

PALOMINODB   Proven Database Excellence

SSI

- Serializable Snapshot Isolation
- Postgres supports now, 3 transactional levels:
    - READ COMMITED
    - REPETEABLE READ (old SERIALIZABLE)
        - Called SNAPSHOT ISOLATION
    - SERIALIZABLE
- No more "select for update"
- Simplify your code
- It has cost, use predictive locking and could increase the number of rollbacks due conflicts between transactions.

http://wiki.postgresql.org/wiki/SSI
http://wiki.postgresql.org/images/4/4f/SSI-PGConfEU2011.pdf

Unlogged tables

- These tables cannot be restored from the WAL in case of crash.
- More performance due of less impact of writing to disk.
- No replication for this tables.
- Truncated in case of crash.
- CLOG contention article: http://rhaas.blogspot.com/2011/11/hint-bits.html

Writable Common Table Expressions

- WCTE
- CTE implemented since 8.4
- Advanced info: http://wiki.postgresql.org/images/d/da/Ciolli-debug-wcte-2011.pdf

```
WITH table_temp(i,j,k) AS
(
    DELETE FROM prueba WHERE i =9 RETURNING *
)
INSERT INTO prueba_del
SELECT i,j,k
FROM table_temp;
```

```
WITH RECURSIVE t(a,b) AS (
    VALUES(0,1)
  UNION ALL
    SELECT greatest(a,b), a + b AS a FROM t
    WHERE b < $1
  )
SELECT a FROM t;
```

KNN in GiST and GIN

- Nearest values using indexes
  - Increases drastically the performance for Geographical and complex data.
  - pg_trgm can be used for improve LIKE and ILIKE searches
- Operator: "<->"

Example article:
http://www.palominodb.com/blog/2011/10/12/indexing-text-columns-gist-or-gin-optimize-ilike-using-pgtrgm-postgres-91-part-1
http://wiki.postgresql.org/images/4/46/Knn.pdf

SE-LINUX support

- Integrated with SE-LINUX.
- The SO permissions are more stronger than database ones.
- Must be enabled explicitly .
- Is based in theSECURITY LABELS creation.
- NOTE: It requires SELINUX libs newer than 2.0.93.

http://wiki.postgresql.org/wiki/SEPostgreSQL_Documentation

PALOMINODB   Proven Database Excellence

Minor features (ALTER ENUM)

- ALTER TYPE tipo_enum ADD VALUE new_value [{BEFORE|AFTER} existing_value];

  CREATE TYPE lista AS ENUM ('tipoa','tipob','tipoc');

  ALTER TYPE lista ADD VALUE 'tipoa2' AFTER 'tipoa';

PGXN

https://github.com/pgxn/pgxn-client
http://pgxn.org/faq/

General and compatibility changes

- Composite types casts were disbled
  - *valor_compuesto*.text, text(*valor_compuesto*)
- CREATE EXTENSION
- Auto tuning of *wal_buffers*
- GiST and GIN improvements
- INSTEAD OF for triggers (for updateables views setup)
- Windows 64 compilation support
- *standard_conforming_strings* now is 'on' by default
  - Before: 'O\'Connor' , now you should use 'O''Connor' or E'O\'Connor'
- PL/SQL changes in RAISE without parameters
- Inherated tables improvement (Merge Append node – Order of the inherated tables inside the append node instead outside).
- FULL OUTER JOIN could use now Hash nodes.
- Detailed info: http://www.postgresql.org/docs/9.1/static/release-9-1.html

Changes (II)

- Replication grants (no more superusers for replictaion)
- Improvement in the security to avoid internal attacks in the replication communication.
- Functional dependencies now are recognized by the GROUP BY (only PK)
  - Before: ERROR: column "tabla.col_pk" must appear in the GROUP BY clause or be used in an aggregate function.

Miscellaneous info

- PLV8 (JavaScript V8 procedural language)
- Postgres-XC http://postgres-xc.sourceforge.net/
- PgPool II 3.1
- New versions londiste, plproxy, bucardo and others

PostgreSQL 9.2

Features incorporated:

 * Index only scans
 * Slave replication cascade
 * Oracle style partitioning
 * Materialized Views
 * …

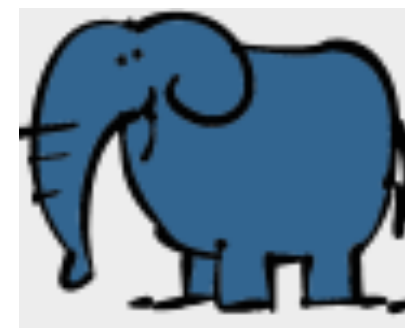http://wiki.postgresql.org/wiki/PostgreSQL_9.2_Development_Plan