

PostgreSQL: Μέθοδοι για Data Replication

FOSSCOMM 2013

Δημήτρης Αγγελάκος

@vyruss / vyruss@hellug.gr

PostgreSQL: Μέθοδοι για Data Replication

- 1) Εισαγωγή στην PostgreSQL
- 2) Data Replication
- 3) Λύσεις
- 4) Σύγκριση Τεχνολογιών
- 5) Εναλλακτικές

Εισαγωγή στην PostgreSQL

- “Η πιο εξελιγμένη open source database”
 - ACID (Atomicity, Consistency, Isolation, Durability)
 - SQL:2008
 - Procedural languages
 - User-defined data types & objects
 - GiST: PostGIS, Full-text search, etc.
 - Database events & Asynchronous Notifications

Φρέσκια PostgreSQL

<http://www.postgresql.org/download>

Binary Packages

- Linux, BSD, Solaris, Mac OS X, Windows
- Repositories:
 - Red Hat (RHEL, CentOS, Fedora, SL)
 - Debian / Ubuntu
- Ελλάδα: <http://PostgreSQL.gr/>

PostgreSQL: Μέθοδοι για Data Replication

- 1) Εισαγωγή στην PostgreSQL
- 2) **Data Replication**
- 3) Λύσεις
- 4) Σύγκριση Τεχνολογιών
- 5) Εναλλακτικές

Data Replication

- Failover
- Load Balancing
- Data Warehousing

Τύποι Data Replication (I)

- Shared Storage (π.χ. NAS)
 - Failover χωρίς απώλεια δεδομένων & overhead
 - Slaves ανενεργοί
 - SPOF
- Filesystem Replication (π.χ. DRBD)
 - Slaves ανενεργοί
 - Πιθανή απώλεια δεδομένων (async)

Τύποι Data Replication (II)

- Log-shipping Replication
 - Slaves ενεργοί
 - Πιθανή απώλεια δεδομένων (async)
- Trigger-based Replication
 - Overhead
 - Επίπεδο πίνακα
 - Data partitioning

PostgreSQL: Μέθοδοι για Data Replication

- 1) Εισαγωγή στην PostgreSQL
- 2) Data Replication
- 3) **Λύσεις**
- 4) Σύγκριση Τεχνολογιών
- 5) Συγκριτική ανάλυση
- 6) Εναλλακτικές

PostgreSQL Native (I)

- Log-shipping (Streaming) Replication
 - Τι είναι τα WALs
 - Write-Ahead Logging
 - Online Backup
 - Point-in-time Recovery (PITR)
- Standby Servers
- File-based / Streaming

PostgreSQL Native (II)

- Σε 5 λεπτά ;
- Master :
 - postgresql.conf :

```
listen_address = '*'  
wal_level = hot_standby  
max_wal_senders = 3  
archive_mode = on  
archive_command = 'cp -i %p /tmp/archive/%f'
```
 - pg_hba.conf :

```
host    replication    all        10.0.0.0/24    trust
```

PostgreSQL Native (III)

- Slave :
 - postgresql.conf :
 hot_standby = on
 - recovery.conf : (όχι στο config dir αλλά στο data dir)
 standby_mode = 'on'
 primary_conninfo = 'host=10.0.0.1'
 archive_mode = on
 restore_command = 'cp -i /tmp/archive/%f %p'

PostgreSQL Native (IV)

- Master :

```
SELECT pg_start_backup('backup',true);
```

```
cd /var/lib/postgresql/9.1/
```

```
sudo tar cpJf /tmp/backup.txz main
```

```
rsync -av /tmp/backup.txz ubuntu@10.0.0.5:/tmp
```

- Slave :

```
cd /var/lib/postgresql/9.1/
```

```
sudo tar xf /tmp/backup.txz
```

- Master :

```
SELECT pg_stop_backup();
```

Slony-I (I)

- Trigger-based Replication
 - Επιλέγουμε τους πίνακες που μας ενδιαφέρουν
 - Πολλαπλές πηγές πίνακα (Master)
 - Υποστηρίζει διαφορετικές :
 - Εκδόσεις PostgreSQL
 - Αρχιτεκτονικές
 - Λειτουργικά συστήματα
 - Για αργό δίκτυο

Slony-I (II)

- Σε 5 λεπτά ;
 - Πακέτα :
 - postgresql-9.1-slony1-2
 - slony1-2-bin
 - Config :
 - pg_hba.conf:
host mydb_slony slony 10.0.0.0/24 md5
 - /usr/share/doc/slony1-2-bin/examples/slon_tools.conf-sample.gz

Slony-I (III)

- /etc/slony1/slon_tools.conf :

```
add_node(node => 1,  
         host  => '10.0.0.1',  
         dbname => 'mydb_slony',  
         port  => 5432,  
         user   => 'slony',  
         password => '12345');  
  
add_node(node => 2,  
         parent => 1,  
         host  => '10.0.0.4',  
         dbname => 'mydb_slony',  
         port  => 5432,  
         user   => 'slony',  
         password => '12345');
```


Slony-I (IV)

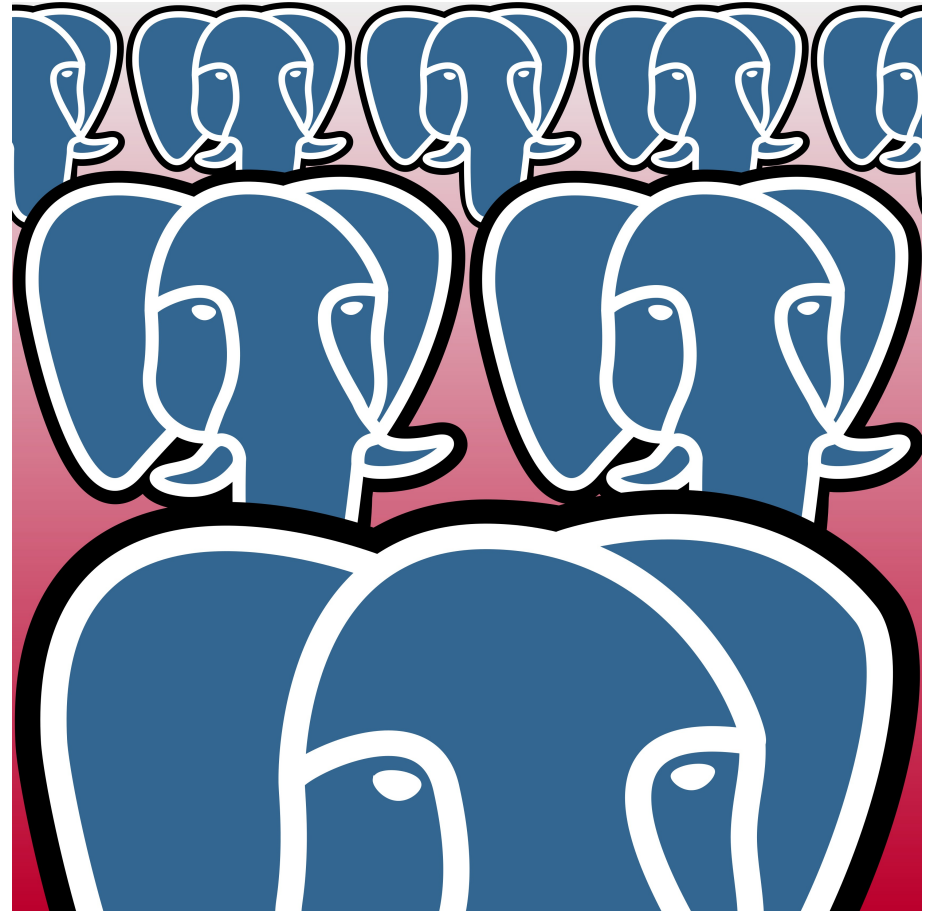
- /etc/slony1/slon_tools.conf :
"set1" => {
 "set_id" => 1,
 "pkeyedtables" => [
 'mytable',
],
},
- slonik_init_cluster | slonik
- sudo slon_start 1
- sudo slon_start 2
- slonik_create_set 1 | slonik
- slonik_subscribe_set 1 2 | slonik

pgpool-II (I)

- Statement-based Cluster Middleware
 - Connection Pooling
 - Load Balancing
 - Query Cache στη μνήμη
 - Failover
 - Replication (Ενσωματωμένο sync ή Εξωτερικό)
 - High Availability (watchdog)
 - Connection Queueing
 - Parallel Queries

pgpool-II (II)

- Η συνεισφορά μου:



pgpool-II (III)

- Σε 5 λεπτά ;
- Πακέτα:
 - pgpool2
 - postgresql-9.1-pgpool2
- /etc/pgpool2/pgpool.conf :

```
backend_hostname0 = '10.0.0.1'  
backend_port0 = 5432  
backend_weight0 = 1  
backend_data_directory0 = '/var/lib/postgresql/9.1/main'  
backend_flag0 = 'ALLOW_TO_FAILOVER'  
backend_hostname1 = '10.0.0.2'  
backend_port1 = 5432  
backend_weight1 = 1  
backend_data_directory1 = '/var/lib/postgresql/9.1/main'  
backend_flag1 = 'ALLOW_TO_FAILOVER'  
pool_passwd = ''  
replication_mode = on  
load_balance_mode = on
```

Χρειάζομαι όντως Replication;

- Παλιά καλά DBLINK functions!
- Ξέρω καλύτερα τη βάση μου
- Επιλεκτικό
- Προχωρημένες λειτουργίες (Business Logic)

Χρειάζομαι όντως Replication;

```
rsql := 'SELECT id, blah
        FROM mytable
        WHERE id > ' || last_inserted_id ||
        ORDER BY id ASC';
PERFORM dblink_connect('conn', conn_string);
PERFORM dblink_open('conn', 'cur', rsql);
LOOP
INSERT INTO toptable (remote_id, blah)
SELECT id, blah
FROM dblink_fetch('conn', 'cur', 500)
IF NOT FOUND THEN EXIT; END IF;
END LOOP;
```

PostgreSQL: Μέθοδοι για Data Replication

- 1) Εισαγωγή στην PostgreSQL
- 2) Data Replication
- 3) Λύσεις
- 4) Σύγκριση Τεχνολογιών**
- 5) Εναλλακτικές

Σύγκριση Τεχνολογιών (I)

- Scope:
 - Native, Pgpool : Ολόκληρη η βάση
 - Slony : Επίπεδο πίνακα
- Conflict Resolution
 - Native, Slony : Δε χρειάζεται
- Πιθανότητα απώλειας δεδομένων
 - Όχι όταν έχουμε Synchronous (άρα όχι Slony)

Σύγκριση Τεχνολογιών (II)

- Πολλαπλοί Master
 - Native, Slony : Όχι
 - Pgpool : Ναι
- Multimaster Replication
 - Άλλος τύπος replication

PostgreSQL: Μέθοδοι για Data Replication

- 1) Εισαγωγή στην PostgreSQL
- 2) Data Replication
- 3) Λύσεις
- 4) Σύγκριση Τεχνολογιών
- 5) **Εναλλακτικές**

Εναλλακτικές

- Postgres-XC
 - Παρακλάδι, τακτικά merge (έρχεται το Μάιο)
 - Write-scalable, shared-nothing
- Bucardo
 - Ασύγχρονο Multimaster
 - Trigger-based
 - Load Balancing
- Londiste

PostgreSQL: Μέθοδοι για Data Replication

Σας ευχαριστώ!

<http://PostgreSQL.gr/>