



Эволюция использования PostgreSQL в справочном API 2GIS

Иванов Денис

2gis.ru

2GIS WebAPI

Данные

- 8 стран

Данные

- 8 стран
 - 300 городов
 - 50 тыс населенных пунктов

Данные

- 8 стран
 - 300 городов
 - 50 тыс населенных пунктов
- 5 млн организаций

Данные

- 8 стран
 - 300 городов
 - 50 тыс населенных пунктов
- 5 млн организаций
- 50 млн гео-объектов

Данные

- 8 стран
 - 300 городов
 - 50 тыс населенных пунктов
- 5 млн организаций
- 50 млн гео-объектов

И еще много всего

Нагрузка

- 30 млн. пользователей в месяц

Нагрузка

- 30 млн. пользователей в месяц
- 2000 RPS

Нагрузка

- 30 млн. пользователей в месяц
- 2000 RPS
- 3 датацентра

Нагрузка

- 30 млн. пользователей в месяц
- 2000 RPS
- 3 датацентра
- 75-100k Tuples Per Second

Нагрузка

- 30 млн. пользователей в месяц
- 2000 RPS
- 3 датацентра
- 75-100k Tuples Per Second
- 3-4k TPS

Цель

2010

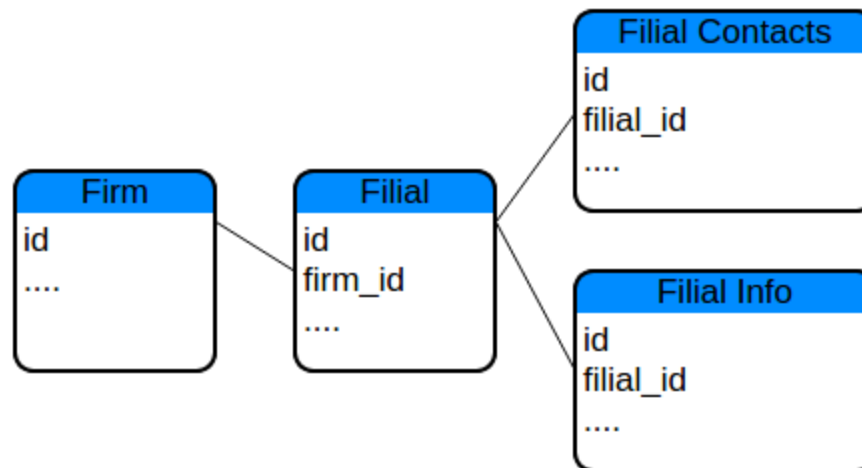
Запуск проекта

Postgresql 9.1

+

Postgis

Нормализованная структура

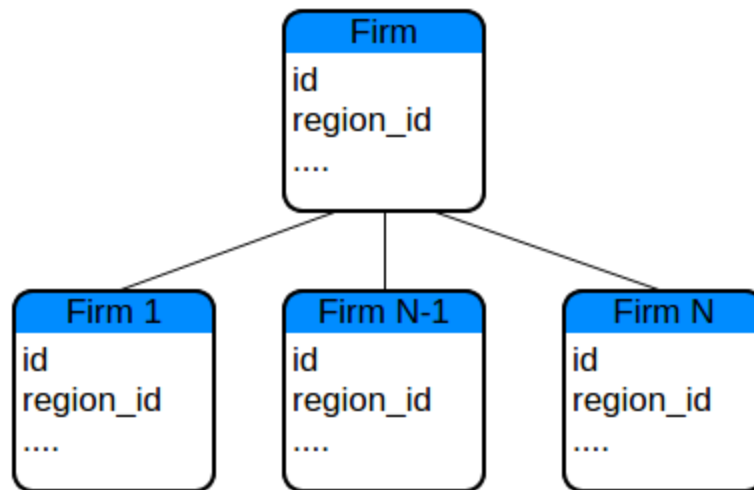


1 Мастер, 1 Слейв
Не было партиципирования

2011

Данных стало больше

30 партиций



1 Мастер, 2 Слейва

2012

Postgresql 9.2

Расширение географии

3 ДЦ
1 Мастер, 6 Слейвов

Денормализация. Json как текст + V8

CREATE OR REPLACE FUNCTION

```
_2gis_get_firm_rubrics_ids(j text)
```

```
RETURNS bigint[]
```

```
LANGUAGE plv8 IMMUTABLE
```

```
AS $$
```

```
01. j = JSON.parse(j);
02. var rubrics = new Array();
03. if (typeof(j) === 'object' && j !== null) {
04.   for (var i in j) {
05.     var rubric = j[i];
06.     rubrics[i] = rubric['id'];
07.   }
08. }
09. return rubrics;
```

```
_2gis_get_firm_rubrics_ids(rubrics) && '{1,2,3,4,5,6,...}'
```

```
CREATE INDEX firm_2gis_get_normalize_contacts_email
ON firm USING gin (
    _2gis_get_normalize_contacts(firm_contacts, 'email')
);
```

Проблемы

Проблемы

- Просели на вставке

Проблемы

- Просели на вставке
- Отставание реплик

Проблемы

- Просели на вставке
- Отставание реплик
- ssh gzipped tunnel - частичное решение

Проблемы

- Просели на вставке
- Отставание реплик
- ssh gzipped tunnel - частичное решение
- ДЦ в Сантьяго не взлетел

Итоги

- PostgreSQL 9.2
- 3 ДЦ
- Денормализация
- Json
- V8

2013

Postgresql 9.3

Text -> Json

3 Мастера, 6 Слейвов

80 партиций

```
01. if(not exists(
        select 1 from pg_indexes
        where indexname=itable || '_name'
    )) then
02.     execute 'CREATE INDEX ' || itable || '_name on '
        || itable || ' USING btree (name)';
03. end if;
```

Итоги

- PostgreSQL 9.3
- Text -> Json
- 3 Мастера, 6 Слейвов
- 80 партиций

**2015
и сейчас**

Postgresql 9.4

Json -> Jsonb

JsQuery

Было

01. `SELECT data->aaa, data->bbb`

02. `ПОТОМ`

03. `_2gis_js_object('aaa', data->aaa, 'bbb', data->bbb)`

Стало

```
json_build_object(...)
```

Было

```
_2gis_get_firm_rubrics_ids(rubrics) && '{1,2,3,4,5,6,...}'
```

Стало

```
rubrics @@ '#.id IN (1,2,3,4,5)'
```

Было

```
WHERE firm_contacts::TEXT ILIKE '%schedule%'  
      AND rubrics::text != '[]'
```

Стало

```
firm_contacts @@ '#(schedule = *)'
```

Партиципирование

Партицирование

```
01. if(not exists(  
    select 1 from pg_indexes  
    where indexname=itable || '_name'  
)) then  
02.     execute 'CREATE INDEX ' || itable || '_name on '  
        || itable || ' USING btree (name);'  
03. end if;
```


Partition Magic

```
01. CREATE SEQUENCE "news_id_seq" START WITH 1 ...;
02. CREATE TABLE news (
03.     id BIGINT DEFAULT nextval('news_id_seq'::regclass),
04.     category_id INT,
05.     title TEXT,
06.     data TEXT
07. );
08. ALTER TABLE ONLY "news" ADD CONSTRAINT "pk_news"
    PRIMARY KEY ("id");
```

```
01. CREATE TABLE news_1 (  
02.     CHECK ( category_id = 1 )  
03. ) INHERITS (news)
```

```
_2gis_partition_magic('news', 'category_id')
```

https://github.com/2gis/partition_magic

Итоги

- PostgreSQL 9.4
- 3 мастера, 9 слейвов (по 3 в ДЦ)
- 120 партиций
- Jsonb
- JQuery
- Partition magic

Что дальше?

Что дальше?

- PostgreSQL 9.5

Что дальше?

- Postgresql 9.5
- Upsert

Что дальше?

- Postgresql 9.5
- Upsert
- Foreign data wrapper

Вопросы?

Иванов Денис

dv.ivanov@2gis.ru

