



Поиск "плохих" запросов



Кто такие Data Egret

- Удаленные DBA PostgreSQL
- Консультанты PostgreSQL
- Постоянно готовим доклады на конференции
- Проводим мастер-классы



Кто наши
плохиши?



```
WITH RECURSIVE fias (  
    SELECT * FROM адреса  
    WHERE level = 1  
    UNION ALL  
    ...  
)  
SELECT r0.id, r1.id, ... r100.id ...  
    FROM таблица1 r0  
    LEFT JOIN справочник1 r1  
    ...  
WHERE r32.id IS NOT NULL  
    AND r56.id IS NOT NULL  
GROUP BY r0.id, r1.id, ... r100.id ...
```

Это плохой запрос?



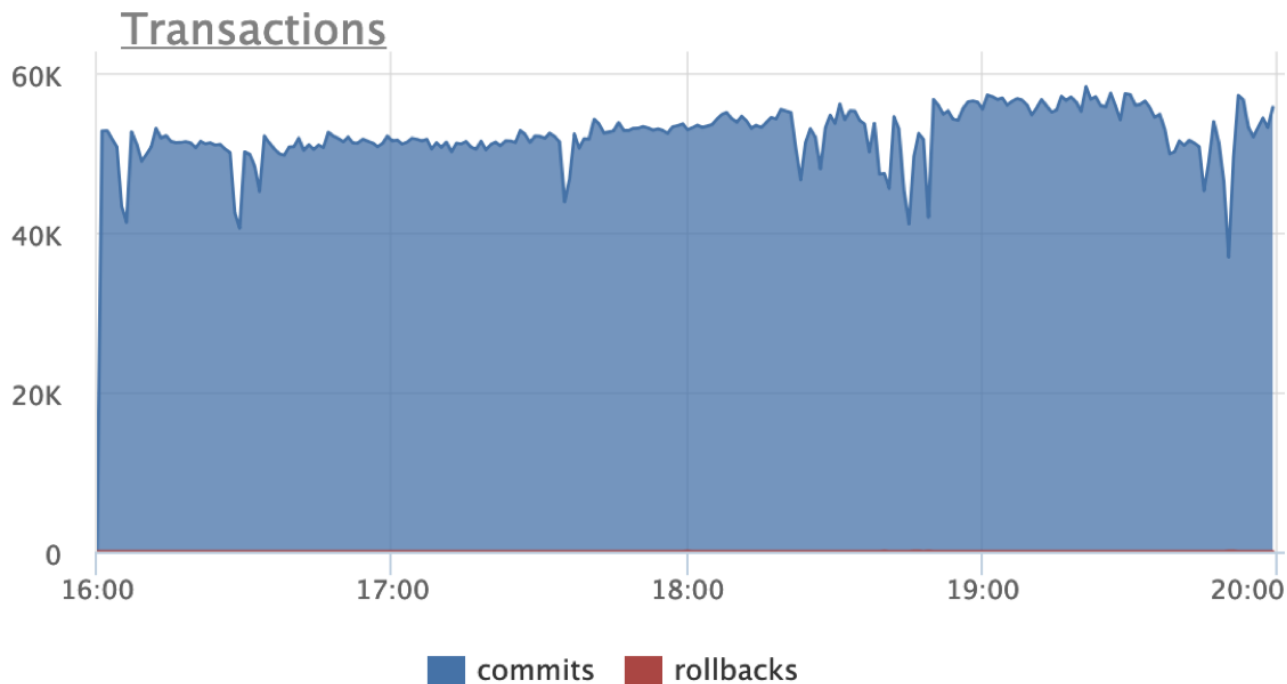
```
SELECT *  
  FROM gender  
 WHERE planet = 'somewherefar'  
    AND alien_type = 'spoke'
```

```
SELECT iu_name  
  FROM dictionary  
 WHERE id = 10
```

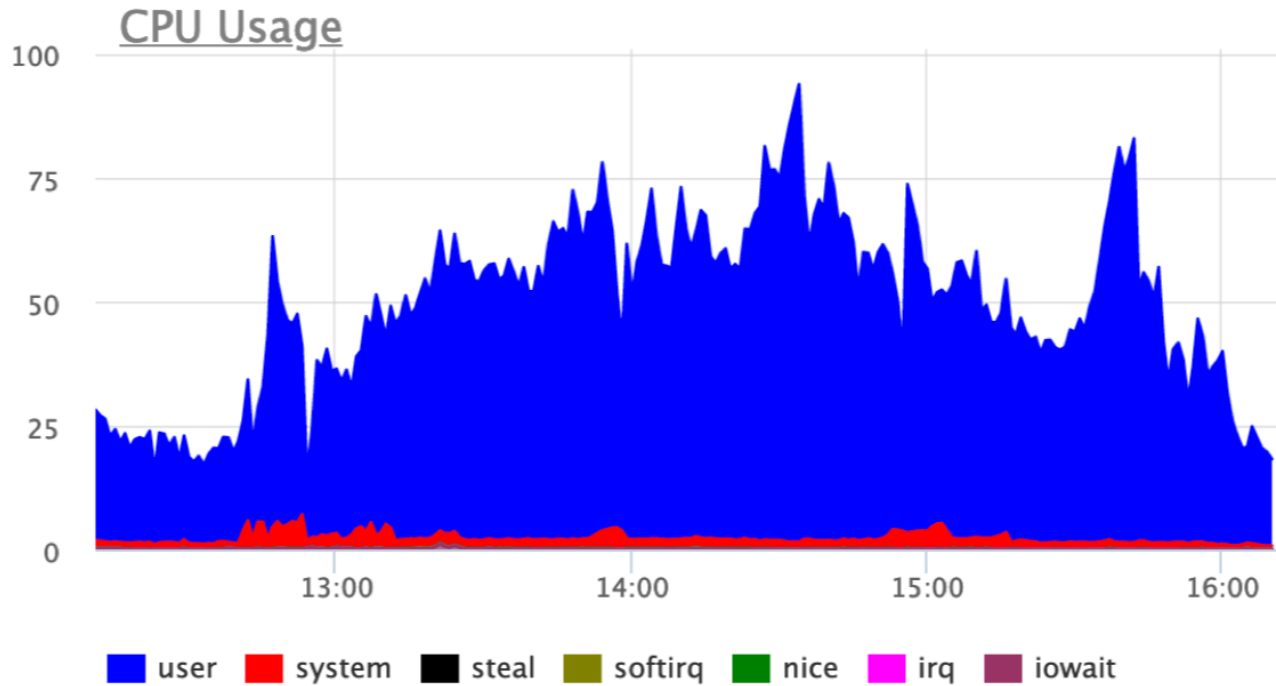
Это плохой запрос?



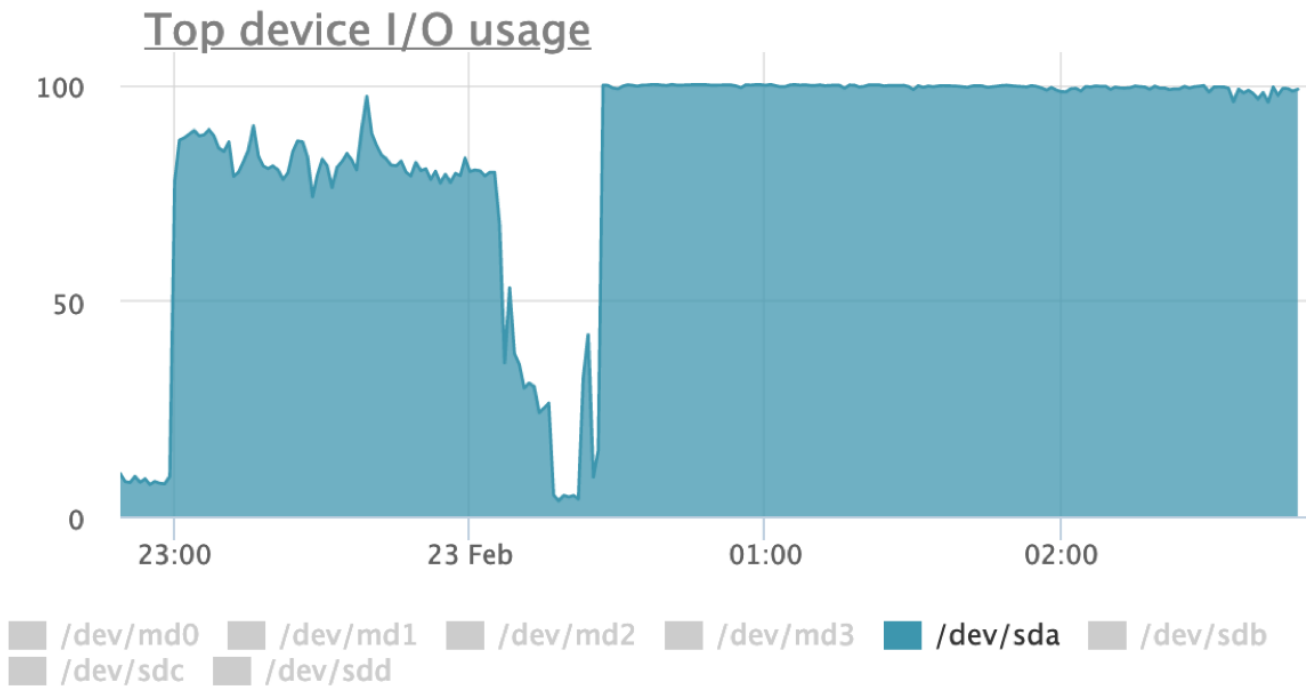
Количество транзакций



Нагрузка на процессор



Диски



pgBadger и
постгресовый
ЛОГ



Минусы

- Специальный конфиг для логов
- Дополнительная нагрузка на диски
- Анализ запросов постфактум
- Пример: 50к запросов в секунду средний размер 100, получим 4Мб/с или 337Гб/день



Плюсы

- Полная информация о всех запросах
- Действительно полная с их параметрами
- Красивые картинки в pgBadger

<http://pgbadger.darold.net/samplev7.html#global-stats>



А как быть если
нужна
статистика
онлайн?



pg_stat_statements

- Стандартное расширение PostgreSQL
- Собирает информацию по всем запросам
- Группирует одинаковые запросы
- Можно сбрасывать статистику
- Оверхед на выполнение запроса 2-5%
- Можно сделать свой отчет на его основе



```
total time: 15:56:10 (IO: 6.73%)
total queries: 40,727,541 (unique: 1,059)
report for all databases, version 0.9.5 @ PostgreSQL 11.5
tracking top 10000 queries, utilities on, logging 1000ms+ queries
=====
pos:1    total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)
        calls: 1,171,859 (2.11%)    avg_time: 31.92ms (IO: 0.0%)
user: super_db    db: super_db    rows: 1,171,859 (1.57%)    query:
SELECT DISTINCT r0.id, r0.Name FROM пользователи r0
LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL
=====
pos:2    total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)
        calls: 19,150,090 (47.02%)  avg_time: 0.17ms (IO: 0.0%)
user: all    db: all    rows: 125,506,633 (20.15%)  query:
other
```



total time: 15:56:10 (IO: 6.73%)

total queries: 40,727,541 (unique: 1,059)

report for all databases, version 0.9.5 @ PostgreSQL 11.5

tracking top 10000 queries, utilities on, logging 1000ms+ queries

=====

pos:1 total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)

calls: 1,171,859 (2.11%) avg_time: 31.92ms (IO: 0.0%)

user: super_db db: super_db rows: 1,171,859 (1.57%) query:

SELECT DISTINCT r0.id, r0.Name FROM пользователи r0

LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL

=====

pos:2 total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)

calls: 19,150,090 (47.02%) avg_time: 0.17ms (IO: 0.0%)

user: all db: all rows: 125,506,633 (20.15%) query:

other



total time: 15:56:10 (IO: 6.73%)

total queries: 40,727,541 (unique: 1,059)

report for all databases, version 0.9.5 @ PostgreSQL 11.5

tracking top 10000 queries, utilities on, logging 1000ms+ queries

=====

pos:1 total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)

calls: 1,171,859 (2.11%) avg_time: 31.92ms (IO: 0.0%)

user: super_db db: super_db rows: 1,171,859 (1.57%) query:

SELECT DISTINCT r0.id, r0.Name FROM пользователи r0

LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL

=====

pos:2 total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)

calls: 19,150,090 (47.02%) avg_time: 0.17ms (IO: 0.0%)

user: all db: all rows: 125,506,633 (20.15%) query:

other



total time: 15:56:10 (IO: 6.73%)

total queries: 40,727,541 (unique: 1,059)

report for all databases, version 0.9.5 @ PostgreSQL 11.5

tracking top 10000 queries, utilities on, logging 1000ms+ queries

=====

pos:1 total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)

calls: 1,171,859 (2.11%) avg_time: 31.92ms (IO: 0.0%)

user: super_db db: super_db rows: 1,171,859 (1.57%) query:

SELECT DISTINCT r0.id, r0.Name FROM пользователи r0

LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL

=====

pos:2 total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)

calls: 19,150,090 (47.02%) avg_time: 0.17ms (IO: 0.0%)

user: all db: all rows: 125,506,633 (20.15%) query:

other



total time: 15:56:10 (IO: 6.73%)

total queries: 40,727,541 (unique: 1,059)

report for all databases, version 0.9.5 @ PostgreSQL 11.5

tracking top 10000 queries, utilities on, logging 1000ms+ queries

=====
pos:1 total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)

calls: 1,171,859 (2.11%) avg_time: 31.92ms (IO: 0.0%)

user: super_db db: super_db rows: 1,171,859 (1.57%) query:

SELECT DISTINCT r0.id, r0.Name FROM пользователи r0

LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL

=====
pos:2 total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)

calls: 19,150,090 (47.02%) avg_time: 0.17ms (IO: 0.0%)

user: all db: all rows: 125,506,633 (20.15%) query:

other



```
total time: 15:56:10 (IO: 6.73%)
total queries: 40,727,541 (unique: 1,059)
report for all databases, version 0.9.5 @ PostgreSQL 11.5
tracking top 10000 queries, utilities on, logging 1000ms+ queries
=====
pos:1      total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)
      calls: 1,171,859 (2.11%)      avg_time: 31.92ms (IO: 0.0%)
user: super_db  db: super_db      rows: 1,171,859 (1.57%)      query:
SELECT DISTINCT r0.id, r0.Name FROM пользователи r0
      LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL
=====
pos:2      total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)
      calls: 19,150,090 (47.02%)  avg_time: 0.17ms (IO: 0.0%)
user: all      db: all      rows: 125,506,633 (20.15%)  query:
other
```



```
total time: 15:56:10 (IO: 6.73%)
total queries: 40,727,541 (unique: 1,059)
report for all databases, version 0.9.5 @ PostgreSQL 11.5
tracking top 10000 queries, utilities on, logging 1000ms+ queries
=====
pos:1    total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)
        calls: 1,171,859 (2.11%)    avg_time: 31.92ms (IO: 0.0%)
user: super_db    db: super_db    rows: 1,171,859 (1.57%)    query:
SELECT DISTINCT r0.id, r0.Name FROM пользователи r0
LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL
=====
pos:2    total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)
        calls: 19,150,090 (47.02%)    avg_time: 0.17ms (IO: 0.0%)
user: all    db: all    rows: 125,506,633 (20.15%)    query:
other
```



```
total time: 15:56:10 (IO: 6.73%)
total queries: 40,727,541 (unique: 1,059)
report for all databases, version 0.9.5 @ PostgreSQL 11.5
tracking top 10000 queries, utilities on, logging 1000ms+ queries
=====
pos:1    total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)
        calls: 1,171,859 (2.11%)    avg_time: 31.92ms (IO: 0.0%)
user: super_db    db: super_db    rows: 1,171,859 (1.57%)    query:
SELECT DISTINCT r0.id, r0.Name FROM пользователи r0
LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL
=====
pos:2    total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)
        calls: 19,150,090 (47.02%)    avg_time: 0.17ms (IO: 0.0%)
user: all        db: all        rows: 125,506,633 (20.15%)    query:
other
```



```
total time: 15:56:10 (IO: 6.73%)
total queries: 40,727,541 (unique: 1,059)
report for all databases, version 0.9.5 @ PostgreSQL 11.5
tracking top 10000 queries, utilities on, logging 1000ms+ queries
=====
pos:1      total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)
           calls: 1,171,859 (2.11%)      avg_time: 31.92ms (IO: 0.0%)
user: super_db   db: super_db      rows: 1,171,859 (1.57%)      query:
SELECT DISTINCT r0.id, r0.Name FROM пользователи r0
  LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL
=====
pos:2      total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)
           calls: 19,150,090 (47.02%)   avg_time: 0.17ms (IO: 0.0%)
user: all   db: all      rows: 125,506,633 (20.15%)  query:
other
```



```
total time: 15:56:10 (IO: 6.73%)
total queries: 40,727,541 (unique: 1,059)
report for all databases, version 0.9.5 @ PostgreSQL 11.5
tracking top 10000 queries, utilities on, logging 1000ms+ queries
=====
pos:1    total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)
        calls: 1,171,859 (2.11%)    avg_time: 31.92ms (IO: 0.0%)
user: super_db    db: super_db    rows: 1,171,859 (1.57%)    query:
SELECT DISTINCT r0.id, r0.Name FROM пользователи r0
LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL
=====
pos:2    total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)
        calls: 19,150,090 (47.02%)  avg_time: 0.17ms (IO: 0.0%)
user: all        db: all        rows: 125,506,633 (20.15%)  query:
other
```



```
total time: 15:56:10 (IO: 6.73%)
total queries: 40,727,541 (unique: 1,059)
report for all databases, version 0.9.5 @ PostgreSQL 11.5
tracking top 10000 queries, utilities on, logging 1000ms+ queries
=====
pos:1      total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)
           calls: 1,171,859 (2.11%)      avg_time: 31.92ms (IO: 0.0%)
user: super_db  db: super_db  rows: 1,171,859 (1.57%)      query:
SELECT DISTINCT r0.id, r0.Name FROM пользователи r0
LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL
=====
pos:2      total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)
           calls: 19,150,090 (47.02%)   avg_time: 0.17ms (IO: 0.0%)
user: all      db: all      rows: 125,506,633 (20.15%)  query:
other
```




```
total time: 15:56:10 (IO: 6.73%)
total queries: 40,727,541 (unique: 1,059)
report for all databases, version 0.9.5 @ PostgreSQL 11.5
tracking top 10000 queries, utilities on, logging 1000ms+ queries
=====
pos:1    total time: 10:23:19 (65.2%, CPU: 69.9%, IO: 0.0%)
        calls: 1,171,859 (2.11%)    avg_time: 31.92ms (IO: 0.0%)
user: super_db    db: super_db    rows: 1,171,859 (1.57%)    query:
SELECT DISTINCT r0.id, r0.Name FROM пользователи r0
LEFT JOIN транзакции r1 ON r1.usr_id = r0.id WHERE r1.id IS NOT NULL
=====
pos:2    total time: 00:52:40 (26.8%, CPU: 26.8%, IO: 37.8%)
        calls: 19,150,090 (47.02%)  avg_time: 0.17ms (IO: 0.0%)
user: all    db: all    rows: 125,506,633 (20.15%)    query:
other
```

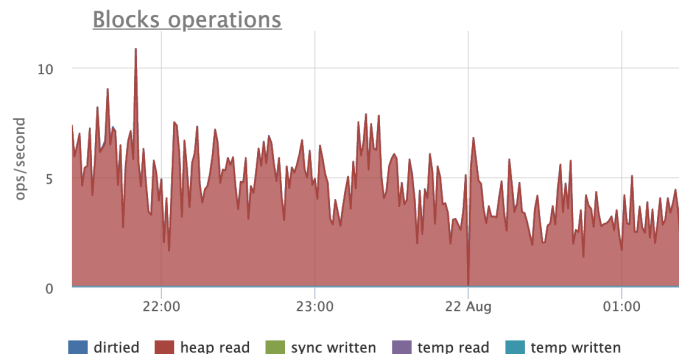
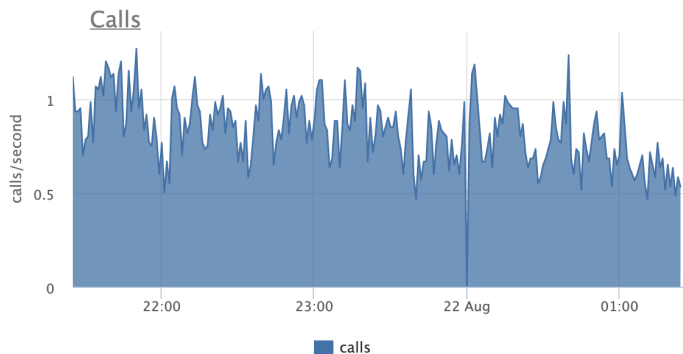
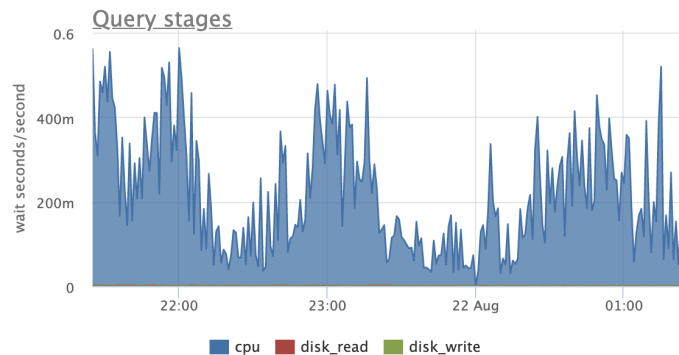
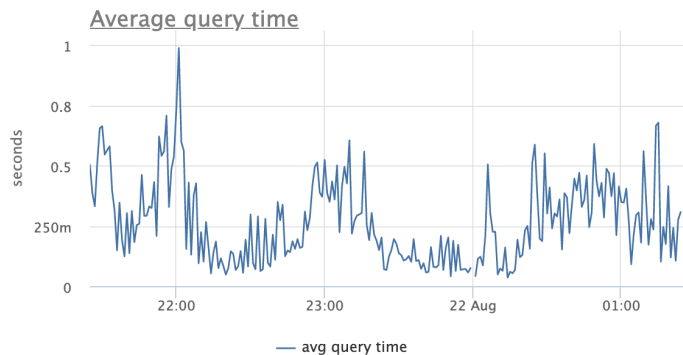


Аномалии по отчету

- Не пришел отчет
- Появились новые запросы
- Запрос стал выполняться чаще
- Запрос стал потреблять больше CPU
- Запрос стал потреблять больше дисков



pg_stat_statements в картинках



Полезные ссылки

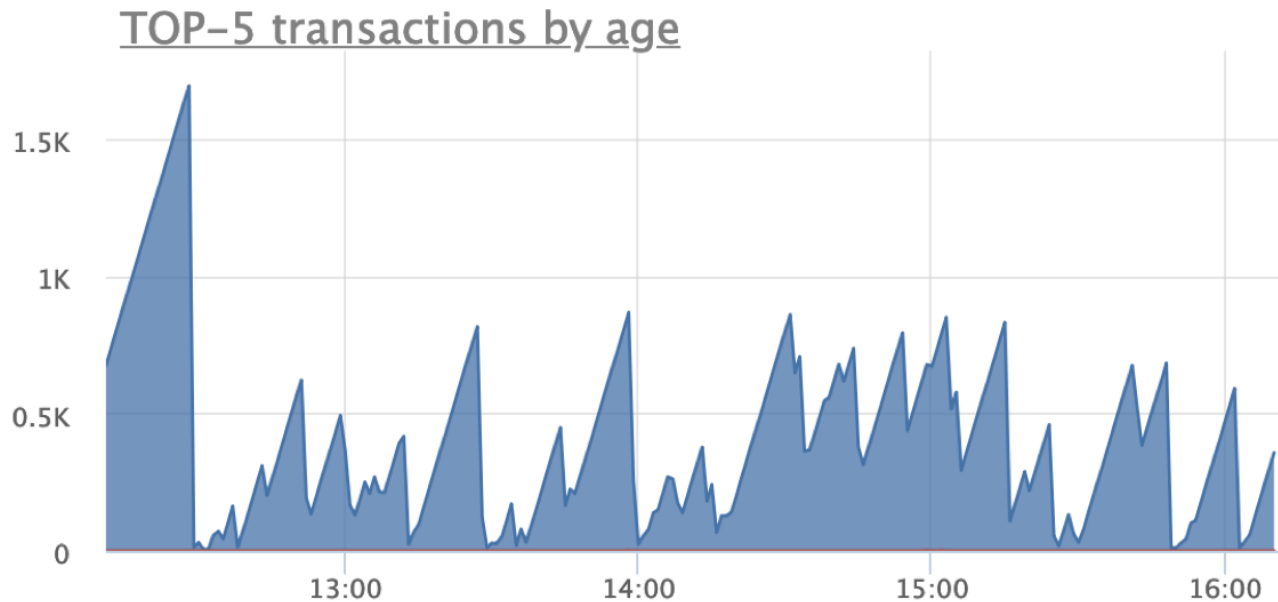
- Набор полезных скриптов и утилит
 - <https://github.com/dataegret/pg-utils>
- Отчет по pg_stat_statements
 - https://github.com/dataegret/pg-utils/blob/master/sql/global_reports/query_stat_total.sql
- Компактор баз данных
 - <https://github.com/dataegret/pgcompactable>



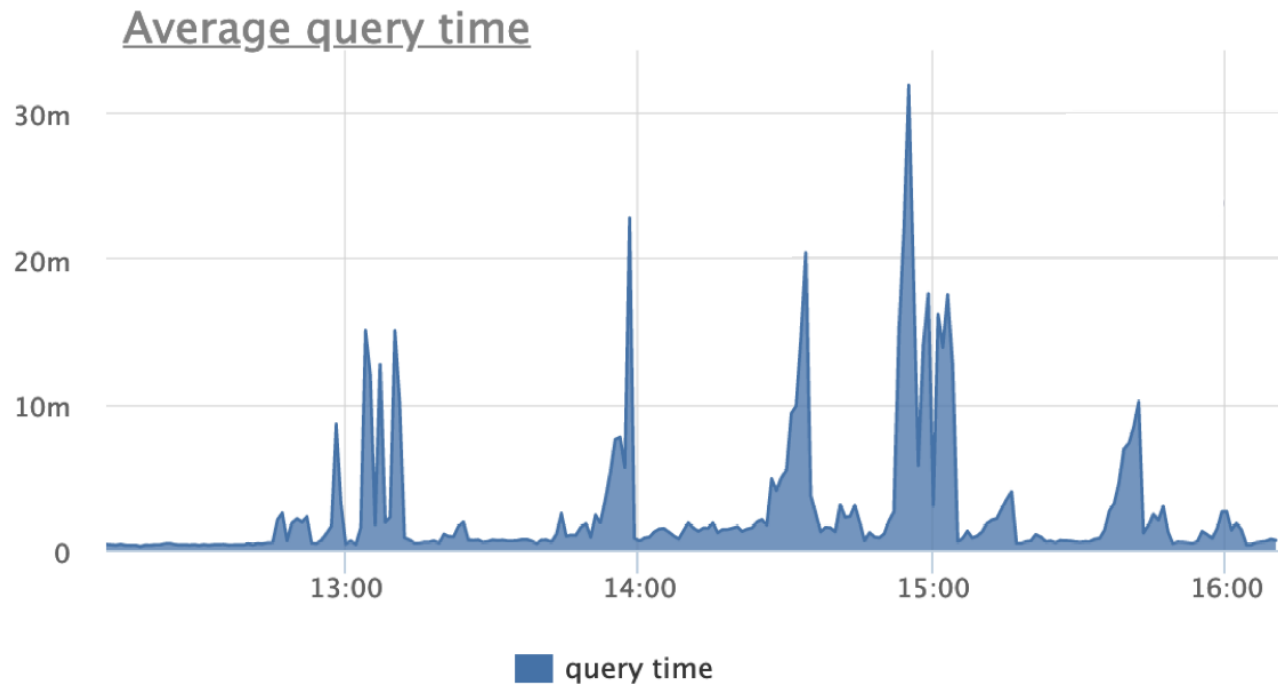
Бонус



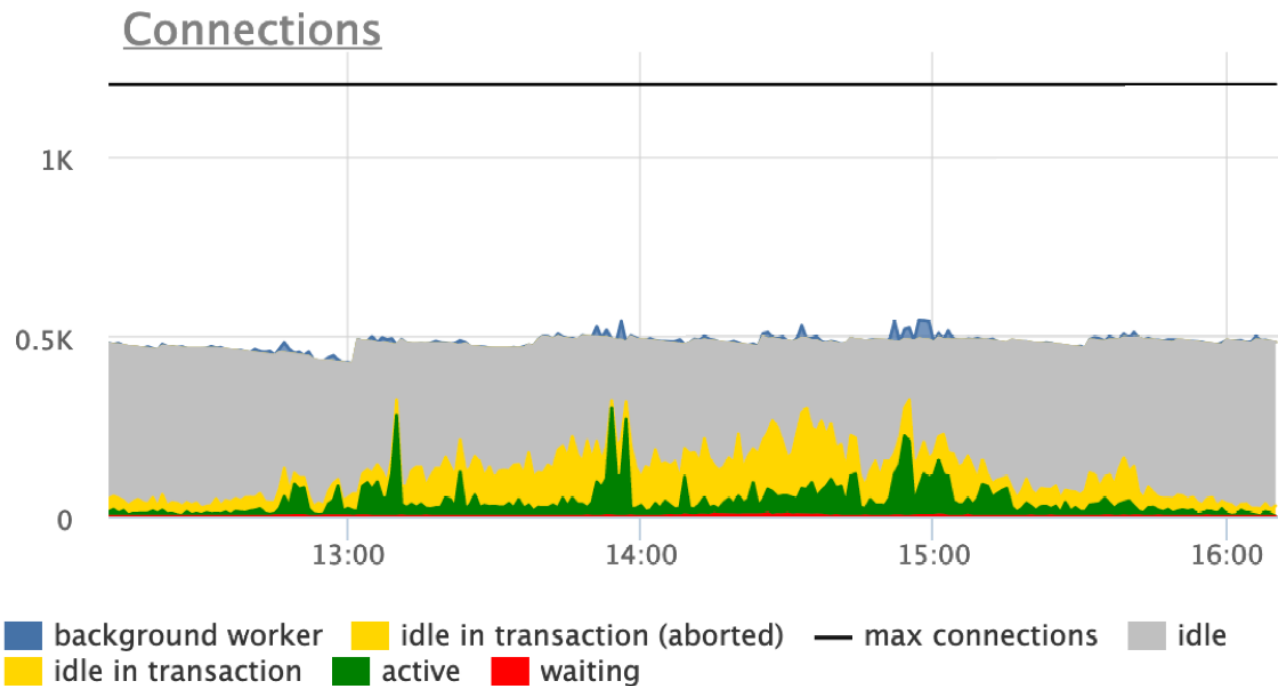
Длительные транзакции в базе данных



Длительные транзакции в базе данных



Длительные транзакции в базе данных





Спасибо за внимание!

PGDAY '20 RUSSIA



10.07.2020

