



Кровавый энтерпрайз, или как мы приручали слона





Онас

- РТ Лабс, группа компаний «Ростелеком»
- Самые известные проекты связаны с электронным правительством и региональной медициной
- Haш High Load по основным базам:
 - Электронное правительство 180k RPS
 - Региональная медицинская информационная система (РМИС) 43к RPS



План презентации

- 1. Стандартизация
- 2. Мониторинг и бекап
- 3. Автоматизация
- 4. Региональная медицинская информационная система
- 5. Вопросы, которые пришлось решать



План презентации

- 1. Стандартизация
- 2. Мониторинг и бекап
- 3. Автоматизация
- 4. Региональная медицинская информационная система
- 5. Вопросы, которые пришлось решать



Внутренний стандарт по развертыванию PostgreSQL

- В поддержке более 150 экземпляров БД PostgreSQL
- Сокращение времени на установку
- Исключение ошибок при начальной настройке
- Позволил автоматизировать большую часть задач



План презентации

- 1. Стандартизация
- 2. Мониторинг и бекап
- 3. Автоматизация
- 4. Региональная медицинская информационная система
- 5. Вопросы, которые пришлось решать



Мониторинг

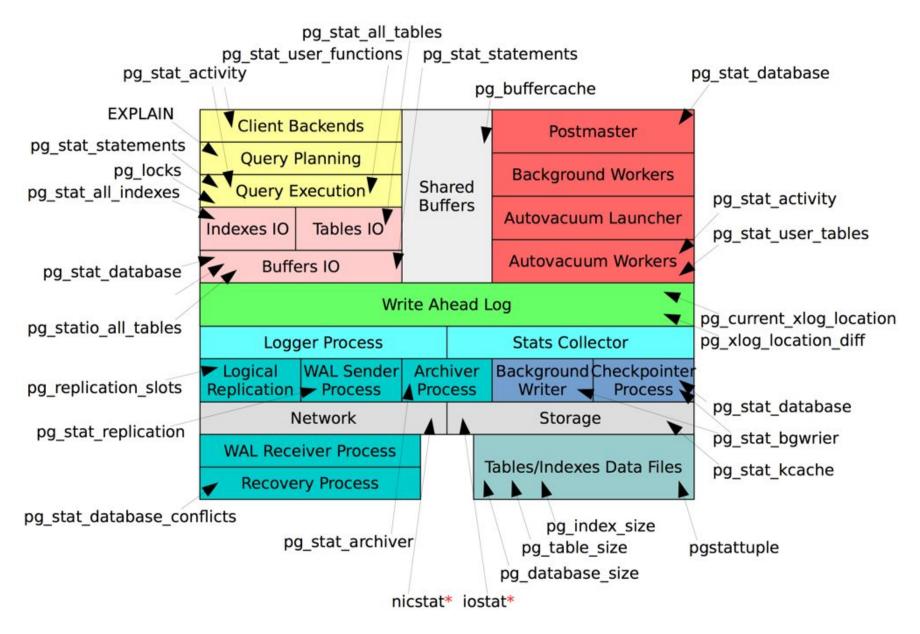




Мониторинг, наши системы

- BMC
- Zabbix
- Vmware log insight
- Grafana





https://www.youtube.com/watch?v=nBSF3rHXM38



Мониторинг, наше расширение pg_eyes

- eyes.get_activity()
 - Pg_stat_activity
 - Pg_stat_database
 - Pg_bgwriter
 - Replication
 - Wal
 - Self response time: clock_timestamp current_timestamp
- eyes.get_activity(p_stat_group)
 - Запросы в таблице eyes.get_activity

oostgres=# select * FROM eyes.get_activity(); stat_name	stat_value
pg_stat_activity.total	79
pg_stat_activity.active	1
pg_stat_activity.active_1s	0
pg_stat_activity.active_time	0
pg_stat_activity.idle	78
pg_stat_activity.idle_in_tr	0
pg_stat_activity.idle_in_tr_1s	0
pg_stat_activity.idle_in_tr_time	0
pg_stat_activity.xact_time	0
pg_stat_activity.wait_lock	0
pg_stat_activity.wait_lock_1s	0
pg_stat_activity.wait_lock_time	0
pg_stat_activity.autovacuum	0
pg_stat_activity.autovacuum_time	0
pg_stat_activity.dba_task	0
pg_stat_activity.dba_task_time	0
pg_stat_database.backends_pct	16
pg_stat_database.xact_total	4419616598
pg_stat_database.xact_commit	4415809946
pg_stat_database.xact_rollback	3806652
pg_stat_database.blks_read	6659104375
pg_stat_database.blks_hit	64684615464
pg_stat_database.tup_returned	2165156597924
pg_stat_database.tup_fetched	12962182159
pg_stat_database.tup_inserted	1420062183
pg_stat_database.tup_updated	1677700996
pg_stat_database.tup_deleted	615422757
pg_stat_database.conflicts	0
pg_stat_database.temp_files	3675
pg_stat_database.temp_bytes	31975560425
pg_stat_database.deadlocks	0
pg_stat_bgwriter.checkpoints_timed	36538
pg_stat_bgwriter.checkpoints_req	12440
pg_stat_bgwriter.checkpoint_write_time	5382678416
pg_stat_bgwriter.checkpoint_sync_time	3902439
pg_stat_bgwriter.buffers_checkpoint	2441008208
pg_stat_bgwriter.buffers_clean	184248642
pg_stat_bgwriter.maxwritten_clean	565298
pg_stat_bgwriter.buffers_backend	1117984270
pg_stat_bgwriter.buffers_backend_fsync	0
pg_stat_bgwriter.buffers_alloc	2957745620
wal_written_b	22961242936384
replication.streaming_p00pguquaddbstb01_b_lag	256
replication.is_in_recovery	0
replication.ms_lag	0
response_time	20
46 rows)	





Мониторинг, наше расширение pg_dbwa

- Аналог oracle statspack
- Можно настроить сбор статистики со стендбай баз
- dbwa.show_top_queries
- dbwa.show_stat_query
- Для работы необходимо расписение в cron
- pg_dbwa, https://goo.gl/7Cnw0M

```
# pg_dbwa
00 * * * * psql -d postgres -c "SELECT dbwa.get_stat('local');"
00 * * * * psql -d postgres -c "SELECT dbwa.get_stat('db02');"
00 * * * * psql -d postgres -c "SELECT dbwa.get_stat('db03');"
```



Мониторинг, расширение pg_dbwa

```
SELECT
  2
         username,
         calls, calls pct,
  3
         round (calls/total time::numeric,3) AS avg ms,
         total time::real,total time pct,
         blk rw time pct,
  6
         regexp_replace(query_text, E'[\n\r]+', ' ', 'g' )::varchar(50) query,
           calls/
  8
                   EXTRACT ( SECOND FROM last_time -first_time) +
  9
                   EXTRACT ( MINUTE FROM last time -first time) *60 +
 10
                                        FROM last time -first time)*60*60
                    EXTRACT ( HOUR
 11
 12
                ) AS tps
     FROM dbwa.show top queries('local','2017-07-01 10:55:00'::timestamp, '2017-07-01 12:02:00'::timestamp,8);
         Query History >
                                  calls
                                             calls_pct
                                                                      total_time
                                                                                  total_time_pct
                                                                                                 blk_rw_time_pct
           username
                                                           avg_ms
                                                                                                                                  query
                                                                                                                                                              tps
                                  283477.0
                                                                      259964.75
                                                                                         56.76
                                                                                                                                                        7.195988482...
quadruple
                                                  11.02
                                                               1.09
                                                                                                            70.16
                                                                                                                      update users_json
                                                                                                                                           set has b ...
                                  240608.0
                                                   9.35
                                                              5.544
                                                                      43398.828
                                                                                          9.47
                                                                                                            9.25
                                                                                                                      select * from users_json where us...
                                                                                                                                                        7.426248269...
quadruple
                                                              7.144
                                                                      26397.895
                                                                                          5.76
                                                                                                            4.23
                                                                                                                                                        4.787383300...
quadruple
                                  188593.0
                                                   7.33
                                                                                                                      delete from user_banners where...
quadruple
                                   97845.0
                                                    3.8
                                                              3.855
                                                                      25379.457
                                                                                          5.54
                                                                                                              2.1
                                                                                                                      insert into user_banners (user_id,...
                                                                                                                                                        2.483769382...
                                                   4.07
                                                               4.74
                                                                      22103.801
                                                                                          4.82
                                                                                                            5.37
                                  104783.0
                                                                                                                      select * from user_rule_notif wher...
                                                                                                                                                        2.659888672...
quadruple
quadruple
                                   43003.0
                                                   1.67
                                                                      16759.004
                                                                                                            3.84
                                                                                                                      update users ison
                                                                                                                                            set data j...
                                                              2.566
                                                                                          3.65
                                                                                                                                                        1.091619753...
quadruple
                                   43558.0
                                                   1.69
                                                              2.709
                                                                      16076.883
                                                                                          3.51
                                                                                                                      select * from rules where is active...
                                                                                                                                                        0.930726472...
                                                                                                              0.0
quadruple
                                   35421.0
                                                   1.37
                                                              3.571
                                                                      9917.7695
                                                                                          2.16
                                                                                                            2.23
                                                                                                                      select banner mnemonic from clo...
                                                                                                                                                        0.899152687...
```



Мониторинг, Zabbix

=)

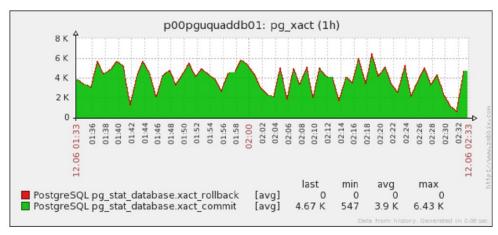
- Хорошо известен всем
- Легко настраивается
- Есть уведомления
- •Дашборды (Screens)
- Оптимизированное хранение
 - History--> Trends

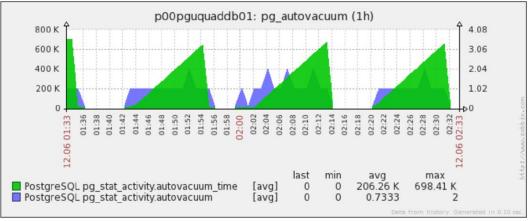
=(

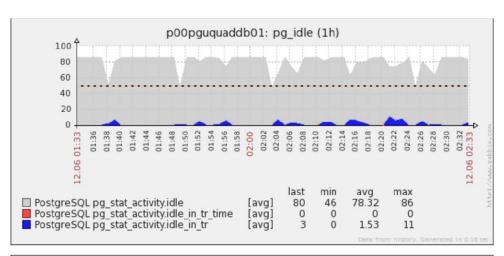
- Не справляется с нашими нагрузками
- Не смогли автоматизировать создание дашбордов

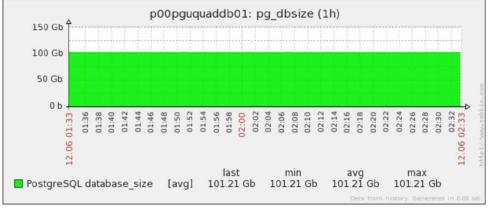


Мониторинг, zabbix











Мониторинг, zabbix

Ссылка на шаблон и описание по установке

https://goo.gl/24TM4m



Мониторинг, collectd graphite grafana

=)

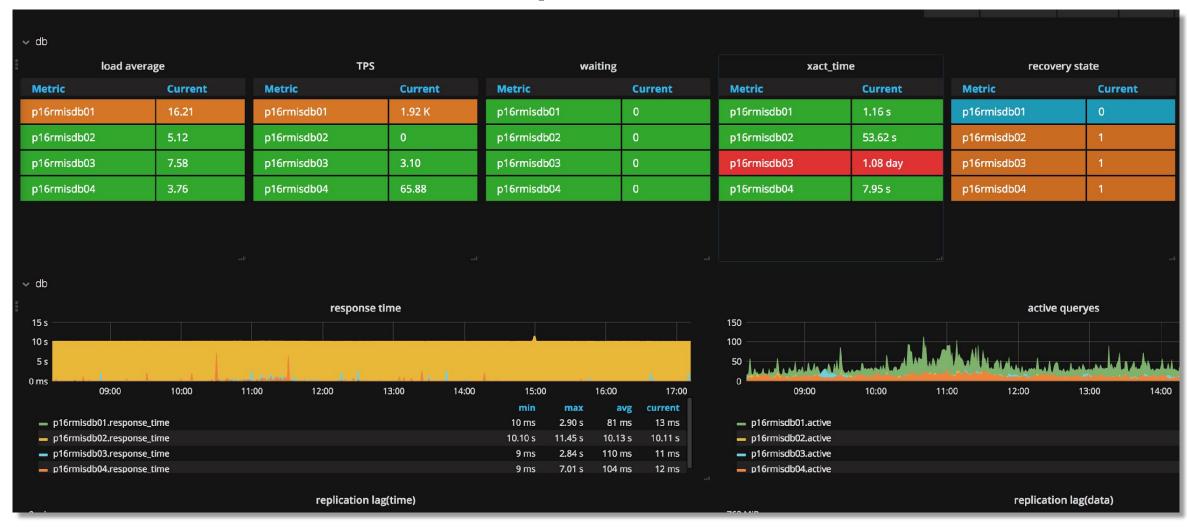
- Держит большие нагрузки
- Можно автоматизировать создание дашбордов для новых систем
- Гибко настраивается
- Оптимизированное хранение
 - History--> Trends

=(

- •До версии 4 не было функционала оповещений
- Сложен в начале



Мониторинг, grafana





Мониторинг, grafana

Ссылка на описание по установке и автоматизации создания дашбордов:

https://goo.gl/TSAZy3



=)

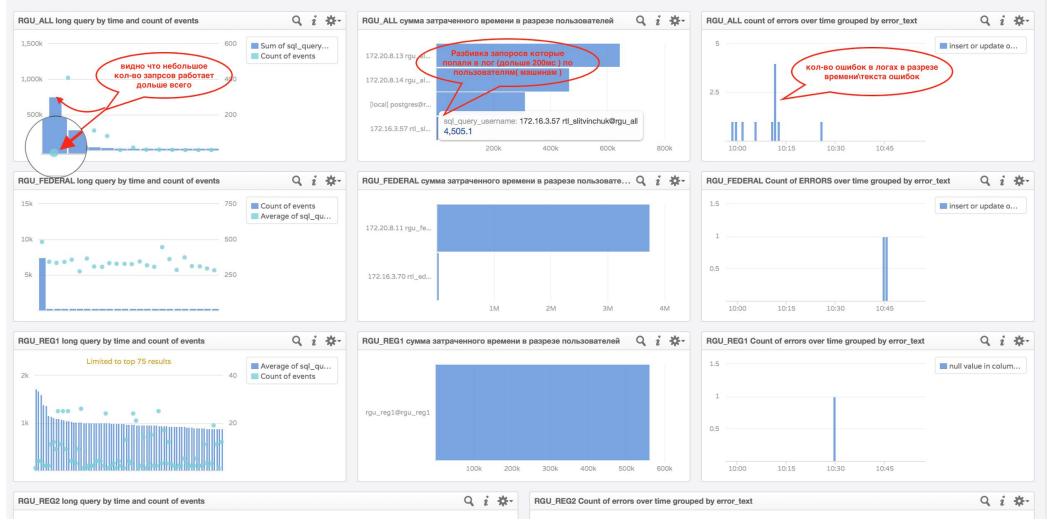
• Интуитивно понятен

- Легко устанавливать и настраивать
- Хорошо масштабируется
- Можно настроить оповещения по хитрым условиям

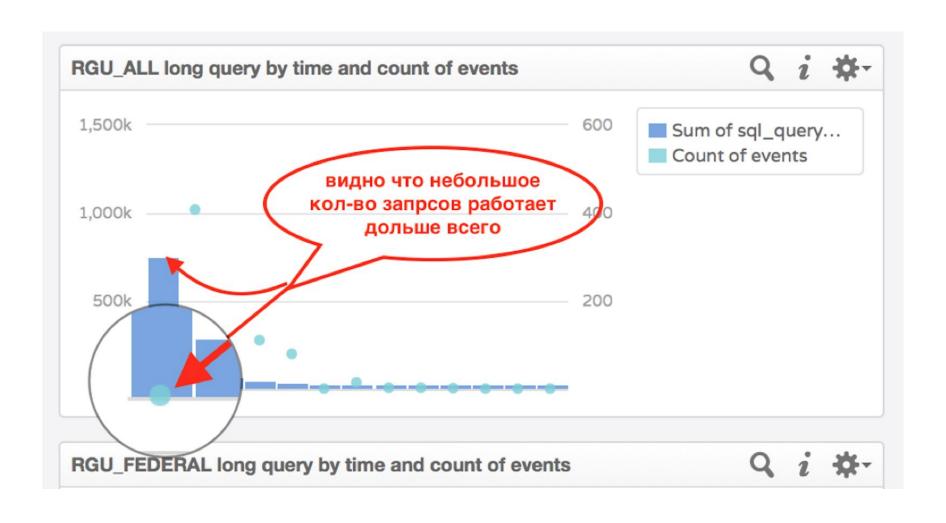
=(

• платный

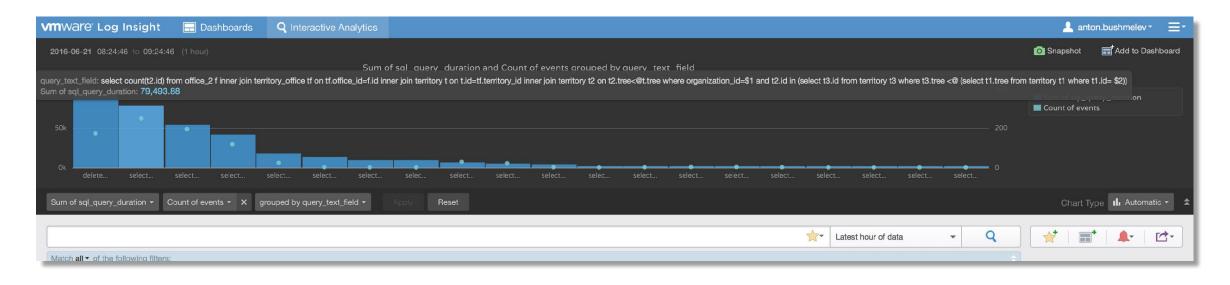


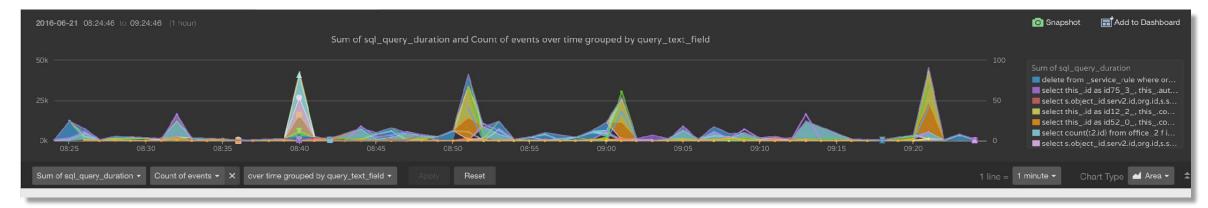














Optional. Thes	e notes are included in the notification message when the alert is fired.
	The state of the s
ify:	
	ss(es) separated by commas
	eparated by spaces
oend to vRealize Op	perations Manager (Contact your admin to enable.)
ID TEST ALERT	
se an alert:	
On any match	
	een for the first time in the last 12 Hours
When an event is se	
When an event is so	
When more than	0 matches are found in the last 5 Minutes
When an event is so When more than When Sum of sql_q	0 matches are found in the last 5 Minutes
When an event is so When more than When Sum of sql_q	v 0 matches are found in the last 5 Minutes v usery_duration in a single group is more than v 0 in the last
When an event is so When more than When Sum of sql_q	wery_duration in a single group ismore thanv0 in the last
When an event is so When more than When Sum of sql_q	wery_duration in a single group ismore thanv0 in the last
When an event is so When more than When Sum of sql_q	matches are found in the last 5 Minutes vuluery_duration in a single group is more than v 0 in the last minute and will only alert once for the defined threshold above.
When an event is so When more than When Sum of sql_q 5 Minutes	matches are found in the last 5 Minutes vuluery_duration in a single group is more than v 0 in the last minute and will only alert once for the defined threshold above.



Ссылка на описание по установке и настройке

https://goo.gl/v1BU4j



по мониторингу

- Заббикс+pnet отвечает за оповещения и нужны для дежурных
- Grafana и pg_dbwa для анализа проблем
- Vmware log insight или bgBadger для анализа долгих запросов и ошибок.



Резервное копирование

- Наши требования
 - Возможность отслеживать статус в системе мониторинга
 - Простое, настраиваемое
 - Инкрементальные бекапы
 - Проверять валидность бекапа
- Наш скрипт: pg bkpman.sh
- Pg_dump



```
task_id
                      6264
                      p66rmisdbbkp01
task_host
task_db_host
                      db03
task_db_port
                      5432
task start time
                      2017-07-02 01:30:02.222389+05
task_end_time
                      2017-07-02 11:40:22.46573+05
task_type
                      full backup
task status
                      completed
                      /u01/postgres/log/pg_bkpman/pg_bkpman.sh.20170702.013001.log
task_log
bkp_start_time
                      2017-07-02 01:30:35.19125+05
bkp_end_time
                      2017-07-02 04:00:21.873671+05
bkp_status
                      completed
bkp_location
                      /u01/postgres/9.5/main/recovery/backup/20170702.013001
bkp size
                      1.5T
bkp label
bkp_unix_time
                      1498941002
retention
                      0
retention_wal
                      1
cleanup
bkp deleted
bkp_deleted_task_id
bkp deleted time
wal deleted
wal_deleted_task_id
wal deleted time
external_storage
                      Y
external_cmd
                      /u01/postgres/9.5/main/recovery/scripts/start_fullbackup.sh
external_start_time |
                      2017-07-02 04:00:21.972061+05
external_end_time
                      2017-07-02 11:38:27.695819+05
external_status
                      completed
```



План презентации

- 1. Стандартизация
- 2. Мониторинг и бекап
- 3. Автоматизация
- 4. Региональная медицинская информационная система
- 5. Вопросы, которые пришлось решать



Автоматизация, Ansible



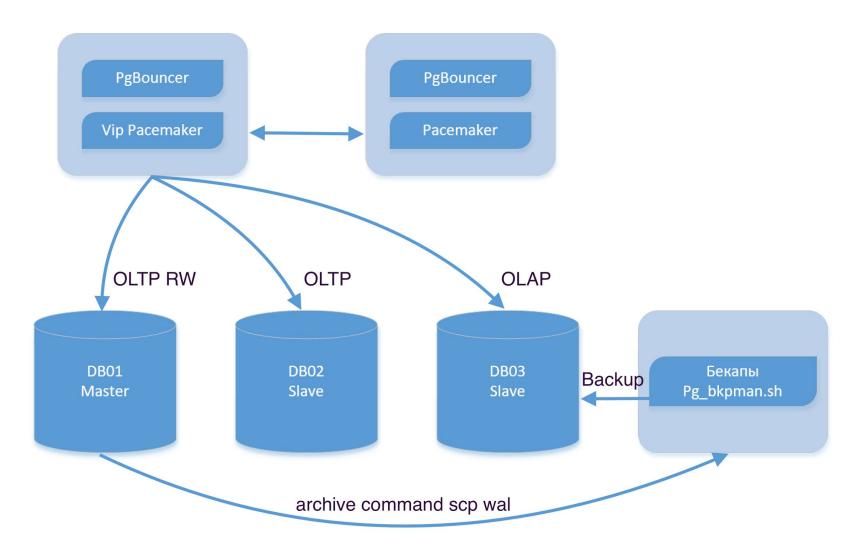


План презентации

- 1. Стандартизация
- 2. Мониторинг и бекап
- 3. Автоматизация
- 4. Региональная медицинская информационная система
- 5. Вопросы, которые пришлось решать



РМИС. типовая архитектура





РМИС, балансировка нагрузки

load average		TPS		waiting		xact_time	
Metric	Current	Metric	Current	Metric	Current	Metric	Current
p56rmisdb01	14.31	p56rmisdb01	3.11 K	p56rmisdb01	0	p56rmisdb01	1.92 min
p56rmisdb02	6.79	p56rmisdb02	463.62	p56rmisdb02	0	p56rmisdb02	9.16 s
p56rmisdb03	2.49	p56rmisdb03	1.86	p56rmisdb03	0	p56rmisdb03	3.63 min
p56rmisdbbkp01	0.03						
p56rmisdbnlb01	1.20						
p56rmisdbnlb02	0.03						
al al al							



План презентации

- 1. Стандартизация
- 2. Мониторинг и бекап
- 3. Автоматизация
- 4. Региональная медицинская информационная система
- 5. Вопросы, которые пришлось решать



Вопросы, которые пришлось решать

- •Оптимальный баланс параметров
- •Vacuum, куда без него ?
- •Репликация и долгие запросы
- •Особенности виртуализации



Итоги

Требования по железу от СЭ тут

Базы данных

Модуль	Объем данных	Кол инстансов
Интеграционный модуль	Сейчса 0	2 (active + backup)
Каталог услуг и интеграция с ФРГУ	?	2 (active + backup)
Модуль вызова внешних сервисов	200Г, после оптимизаци INFORMER и CACHEABLE_SERVICE будет значительно меньше	2 (active + backup)
Модуль уведомлений	< 1Γ	2 (active + backup)
Сервер форм	500M	2 (active + backup) Общий кластер
Сервис периодических задач	< 1M	Общий кластер
Сервис черновиков	Черновики = 575Г + 545Г	Черновики 12 (active + backup),
	Подсказки = 1.3Г	Подсказки 2 (active + backup)
Сервисы конфигурирования и справочная система ЕПГУ	2Γ	Общий кластер
Сервисы личного кабинета	Пользовательские данные = 250Г	Пользовательские данные 8 (active + backup),
	Справочники < 100М	Справочники 2 (active + backup)
Хранилище файлов	120Γ	4 (active + backup) , если шард 100Г
Сервис CMS	< 50M	Общий кластер
Тех портал	< 50M	Общий кластер
Модуль построения отчетов	?	?



Skype: bushmelev_aa

Anton.bushmelev@rtlabs.ru





Полезные ссылки

- http://orabase.org/index.php/pgday2017/postgresql-tips-tricks/
- https://www.youtube.com/watch?v=nBSF3rHXM38
- https://github.com/alexandersamoylov/pg_eyes
- https://github.com/alexandersamoylov/pg_dbwa