



PostgreSQL глазами инженера по эксплуатации

Смерть эксплуататорам и программистам!

Александр Чистяков,
главный инженер Git in Sky,
2015





Мы

- § A small consulting company
- § Web operations engineers
- § Performance engineers
- § Elephant lovers
- § Знаем английский :)





Вы

- § Инженеры по эксплуатации?
- § Веб(?)-разработчики?
- § Любите слонов
- § Может быть, пони вы тоже любите?





Повестка дня

- § Что такое “хорошо” и что такое “плохо”
- § Случаи из жизни
- § Раздача слонов и материализация views
- § Границы применимости, законы физики и другое занудство
- § PostgreSQL vs. MySQL, PostgreSQL vs. MongoDB, ...





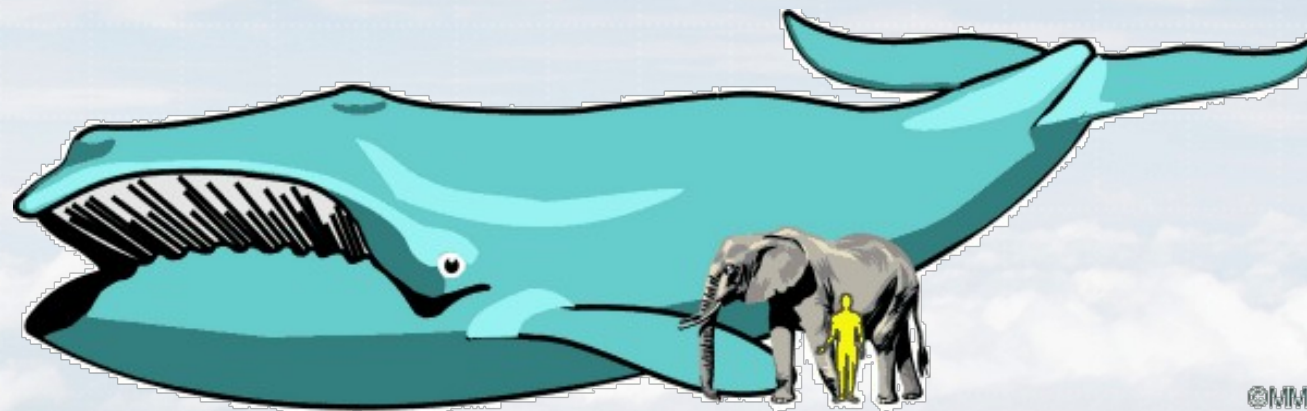
Что мы эксплуатируем

- § В основном — PostgreSQL 9.3.X
- § В недавнем прошлом — 9.1.X и 9.2.X
- § Уже начали использовать 9.4



Составляющие процесса эксплуатации

- § Sizing & capacity planning
- § Установка и настройка
- § Репликация
- § Бэкап
- § Мониторинг
- § Соблюдение законов физики
- § Обновление схемы БД и версий



©MMI



PostgreSQL

Sizing & capacity planning

- § PostgreSQL — современная база данных
- § Очень производительная!
- § Поэтому sizing и capacity planning мы делать не будем
- § И обсуждать тоже не будем
- § Зачем время тратить?





Установка и настройка

- § Как правило — ничего удивительного
- § Однажды мы запустили PostgreSQL в Docker
- § Было не очень удобно,
- § Однажды мы запустили PostgreSQL в Amazon
- § Но про capacity planning мы решили молчать



- § Бывает:
- § «Встроенная» асинхронная
- § «Встроенная» синхронная
- § «Внешняя» триггерная
- § Не бывает:
- § Мастер-мастер (хотя именно ее все и просят)
- § Пока не пробовали: logical decoding





Встроенная репликация

- § Работает вот уже примерно два года
- § Настраивается в < 10 простых шагов
- § Ничего удивительного
- § Лучше настраивать сразу, пока база еще маленькая
- § Если база большая — возьмите WALs с бэкапа





Внешняя триггерная репликация

§ Не нужна*



- § Многих славный путь:
- § pgdump =>
- § barman =>
- § отрицание, гнев, торг, депрессия, смирение
- § Космодемьянский на мой доклад не пришел,
поэтому ZFS!



- § Как часто делать базовый бэкап — time/space trade-off
- § Ориентируйтесь на размер WALs и скорость восстановления
- § Немного нарушает принцип наименьшего удивления - хранит WALs в сжатом виде без расширения .gz
- § Даже будучи установлен локально, хочет забирать WAL'ы по ssh



Обновление схемы БД

§ ALTER TABLE ADD COLUMN DEFAULT





Обновление версий

- § Задача, отсутствующая в мире MySQL
- § В мире PostgreSQL до версии 9.4 — триггерная репликация
- § Bucardo





Bucardo

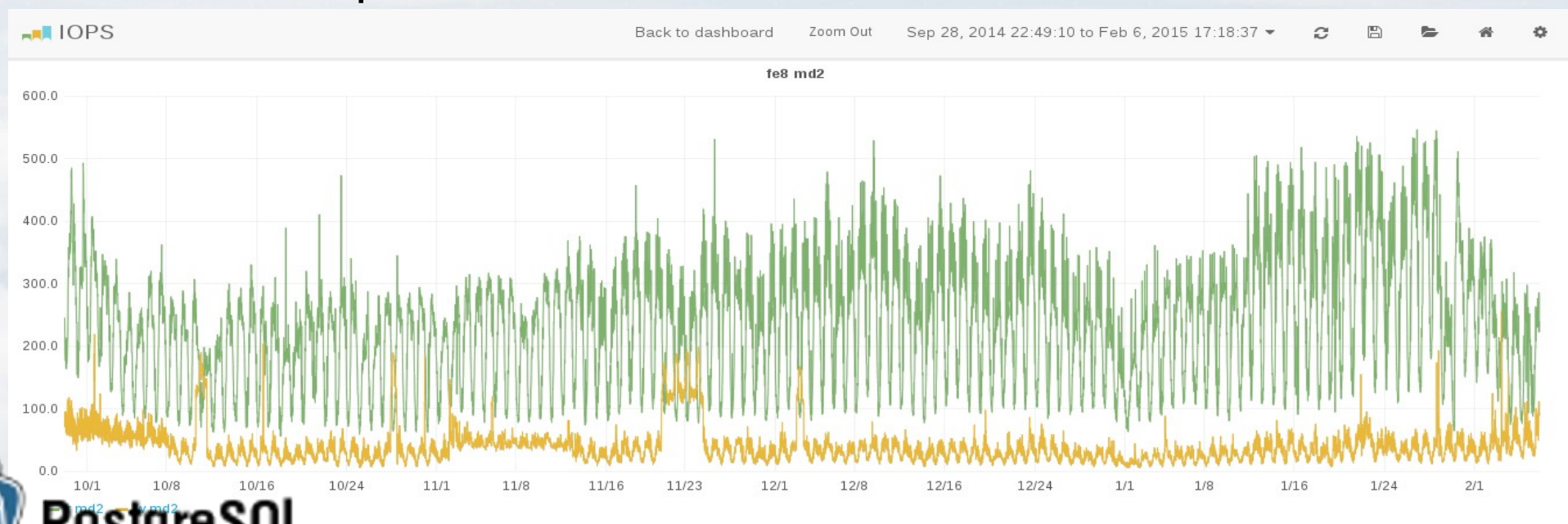
- § Устанавливается почти не вызывая удивления
- § В каком каталоге .pgpass?
- § Работает почти не вызывая удивления
- § <https://gist.github.com/alexclear/a11deed1be3b1d045187>
- § Если в бизнес-логике используется TRUNCATE — беда



§ 9.3 — 9.4, 2.0 Tb — 1.7 Tb

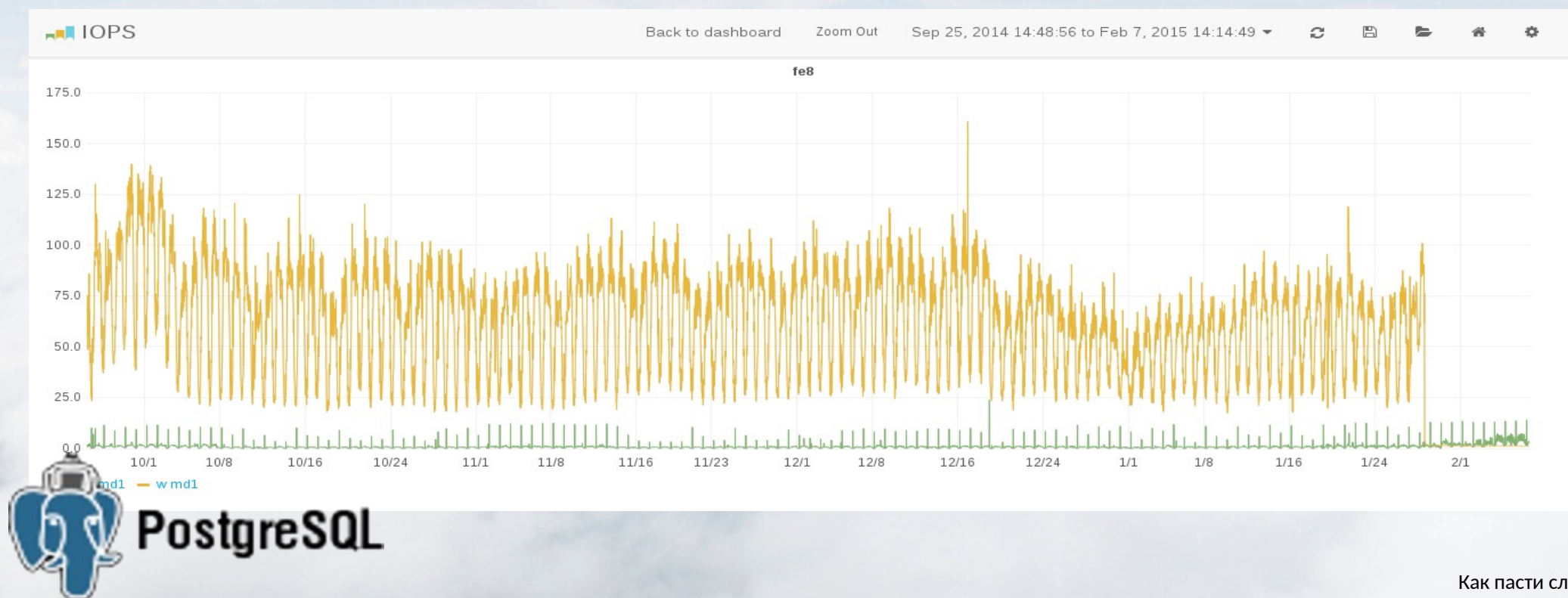
```
root@fe8 ~ # bucardo status fe8fe10
=====
Last good          : Jan 30, 2015 02:24:29 (time to run: 93h 16m 12s)
Rows deleted/inserted : 0 / 605,796,434
Sync name          : fe8fe10
Current state       : Sync public.content (KID 3323)
Source relgroup/database : fe8fe10 / binstore
Tables in sync      : 13
Status              : Active
Check time          : None
Overdue time         : 00:00:00
Expired time         : 00:00:00
Stayalive/Kidsalive : Yes / Yes
Rebuild index        : No
Autokick             : Yes
Onetimecopy          : No
Post-copy analyze    : Yes
Last error:          :
=====
```

§ В PostgreSQL масса всяких параметров, ни один из которых мне не интересен



PostgreSQL

§ Индексы должны помещаться в память!*





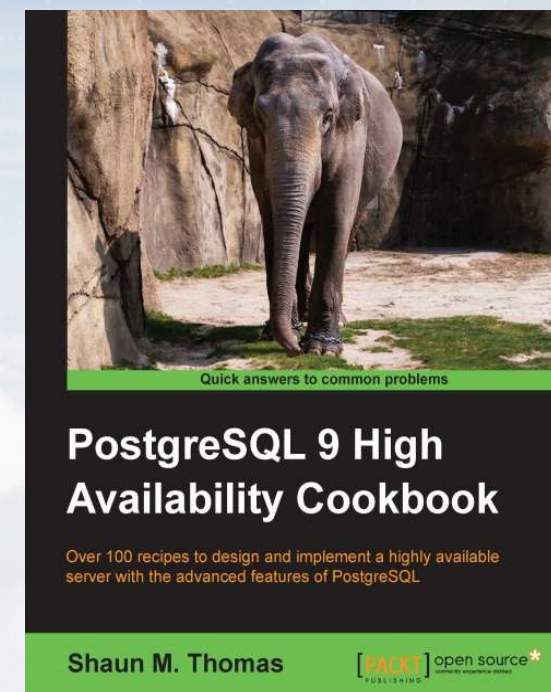
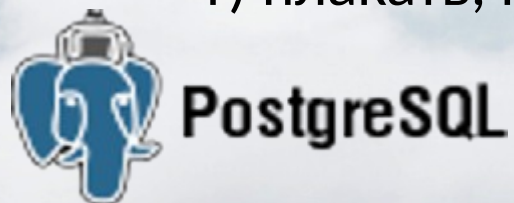
Мониторинг

- § `pg_stat_statements`
- § POWA
- § Настройка над `pg_stat_statements`, которая умеет хранить историю
- § `shared_preload_libraries = 'powa,pg_stat_statements'`
- § При апгрейде базы не забыть пересобрать
- § Не работает*



Sizing & capacity planning

- § Для богатых:
- § ISBN-10: 1849516960 (вся первая часть книги)
- § Для бедных:
- § а) взять железо, какое было в магазине
- § б) задать целевое значение метрики
- § в) протестировать под нагрузкой
- § г) плакать, проклинать себя, жизнь и индустрию

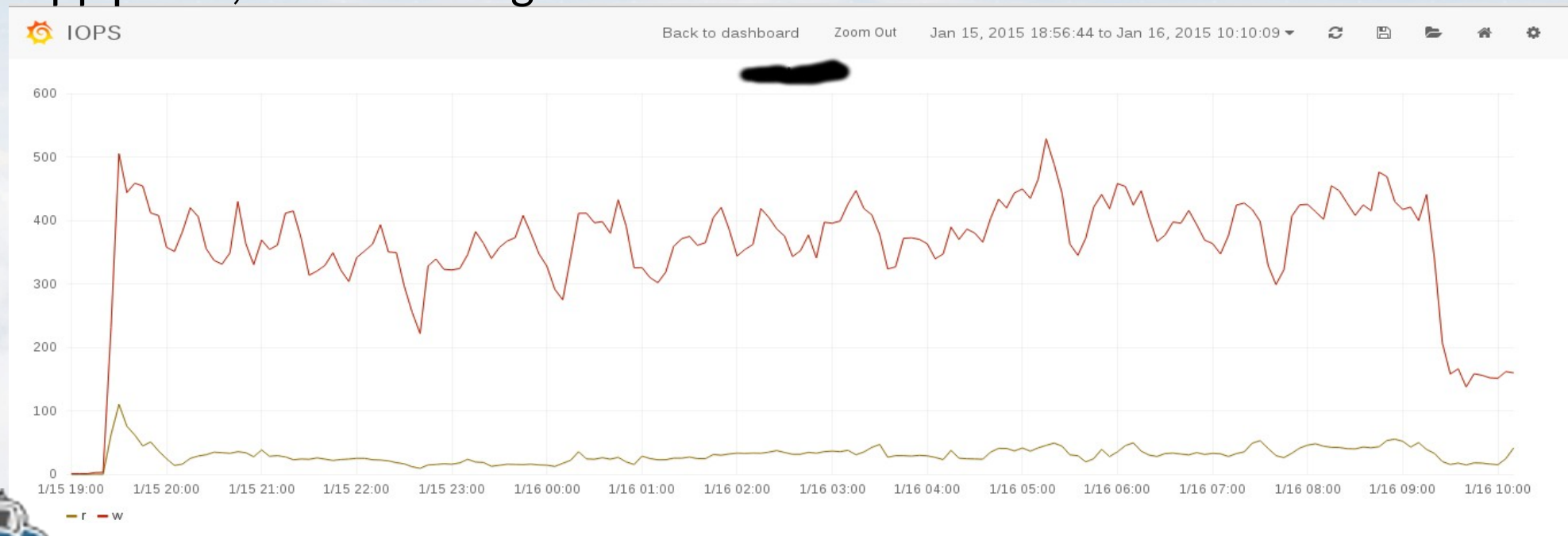


А что делать, если слон устал?

- § У одной моей подруги с ее парнем
- § Исходно было нужно 3K RPS на вставку
- § Требования возросли до 30K RPS за две недели до старта
- § К счастью, есть немного времени подумать
- § и поработать
- § Здесь есть кто-то с 30K вставок в секунду в PostgreSQL?

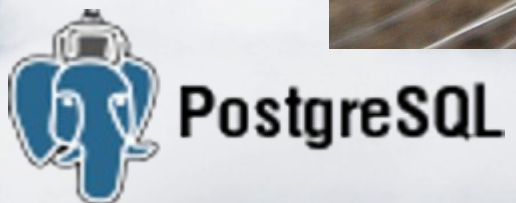


§ Дорогая, я затюнил bgwriter



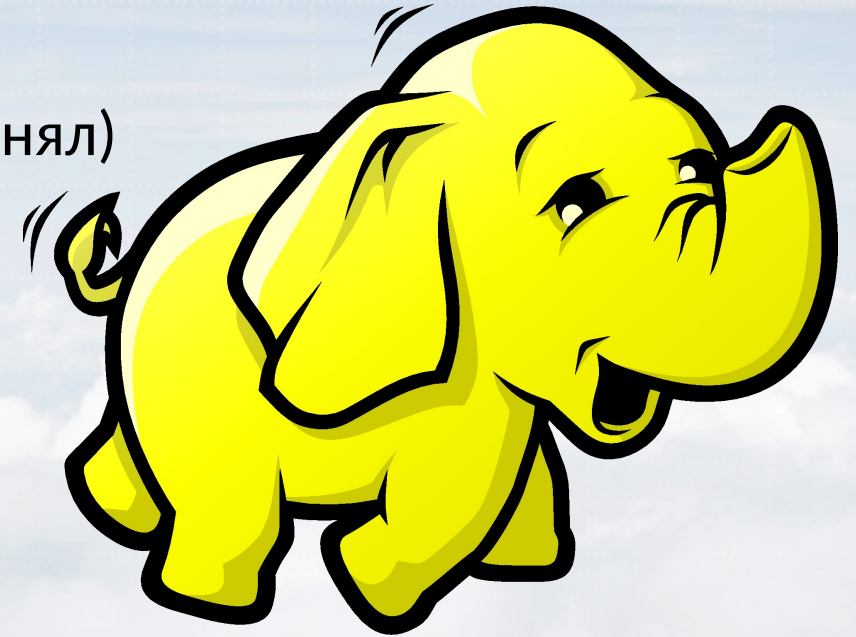


Объективация женщины это плохо!

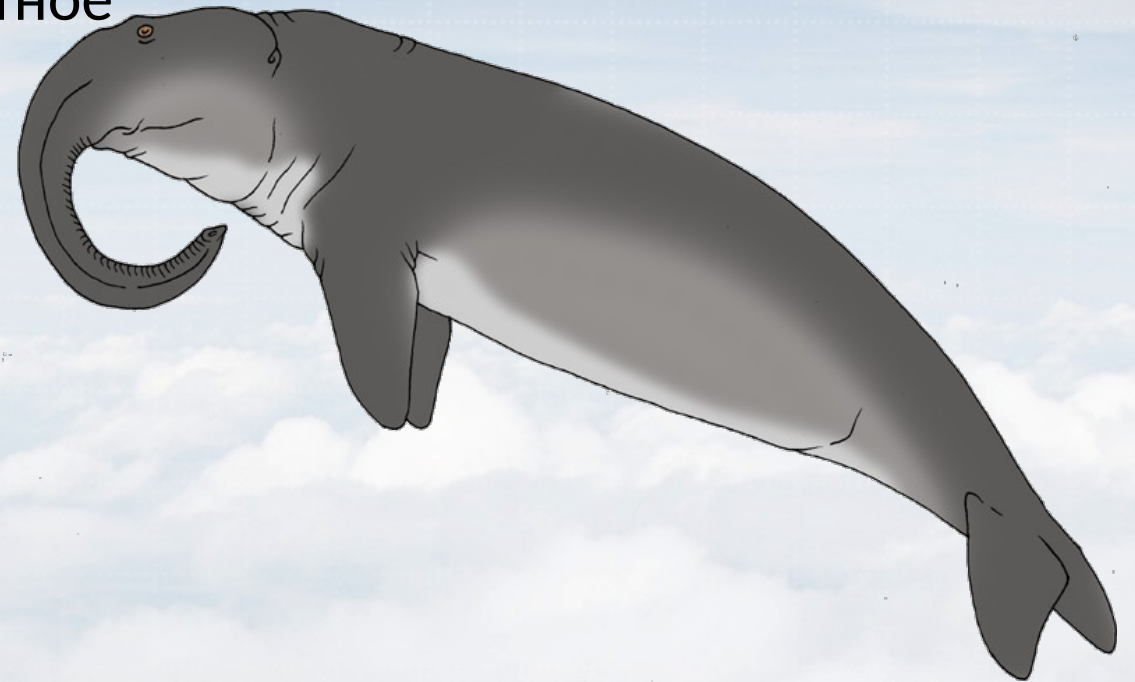


Слон слону — слон!

- § Что такое “база данных” и зачем она нужна?
- § (К слову, зачем нужна MongoDB, я так и не понял)
- § Size does matter — smaller is better!
- § Индексы в PostgreSQL, все остальное — где-нибудь еще



- § Слон — отличное домашнее животное
- § Слоны требуют ухода
- § Не обижайте слонов!





Спасибо за внимание!

Пожалуйста, ваши вопросы.

С вами был Александр Чистяков,
главный инженер Git in Sky

alex@gitinsky.com
<http://gitinsky.com>
<http://meetup.com/DevOps-40>

