Яндекс

Отказоустойчивость сервисов

Павел Пушкарев
Руководитель группы администрирования

Я.Субботник, Санкт-Петербург, 30 июня 2012 года

Зачем всё это нужно?

Зачем всё это?

- Деньги

Зачем всё это?

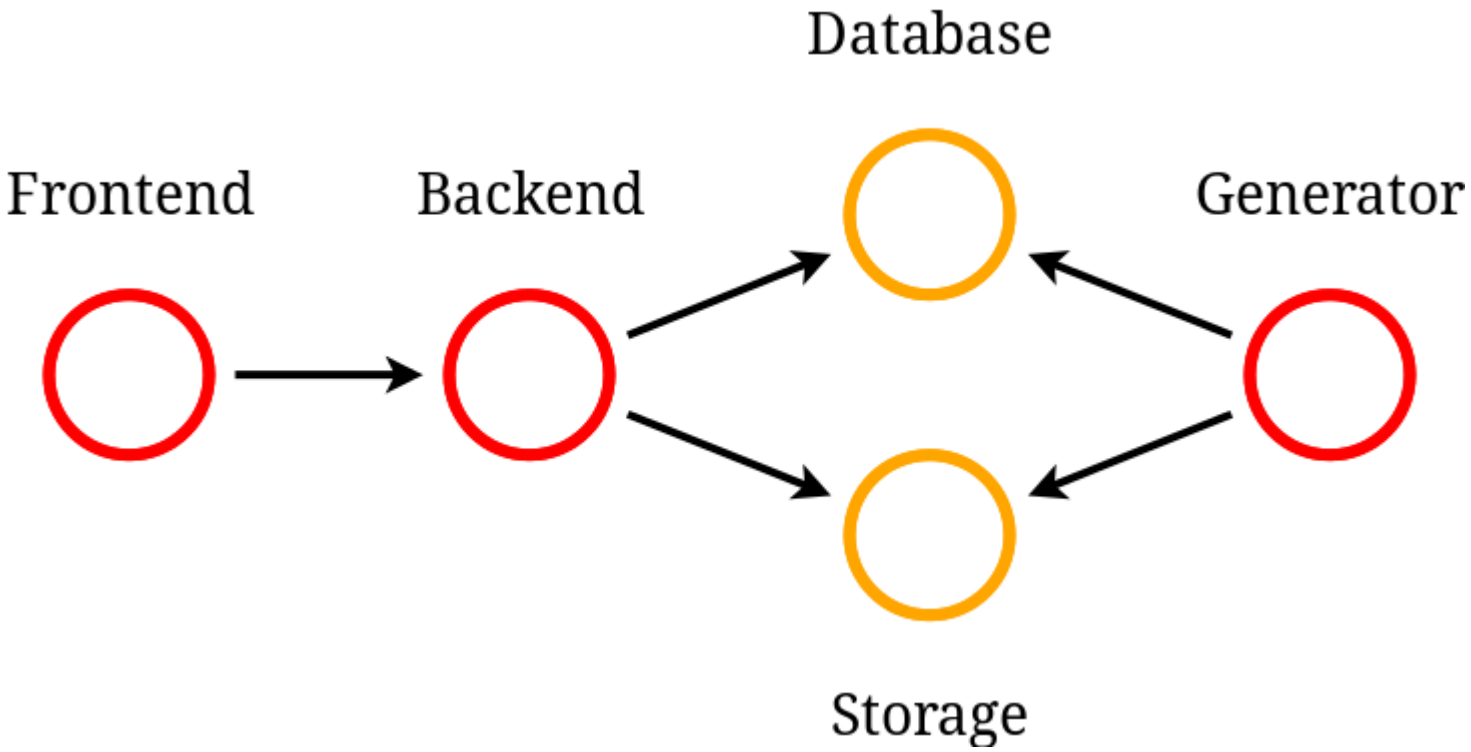
- Деньги
- Имидж

* эта бутылка должна наталкивать на мысли про имидж и жажду

Зачем всё это?

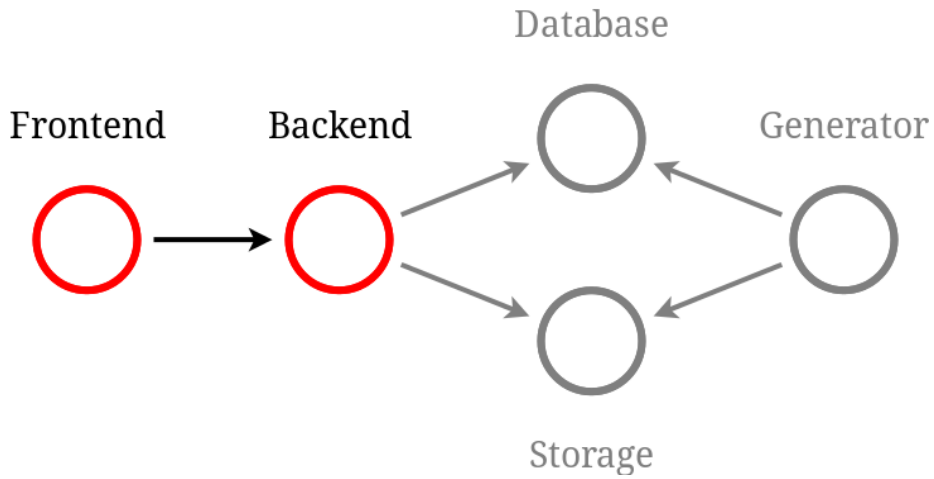
- Деньги
- Имидж
- Тренировка ;-)

Конкретизируем сервис



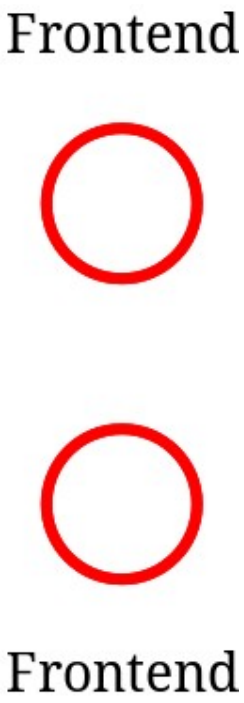
Наша цель — сделать сервис совсем без точек отказа!

Отказоустойчивость регулярных узлов



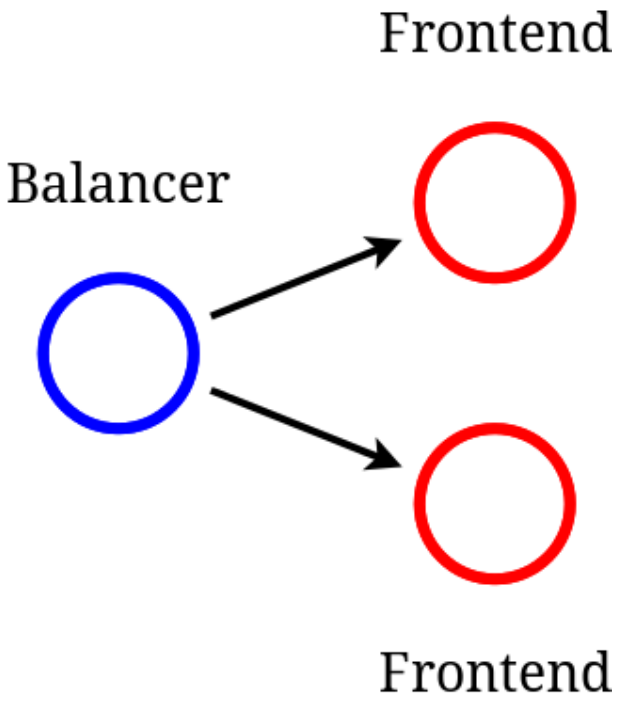
Собрать балансировщик

- Добавим компьютеров



Собрать балансировщик

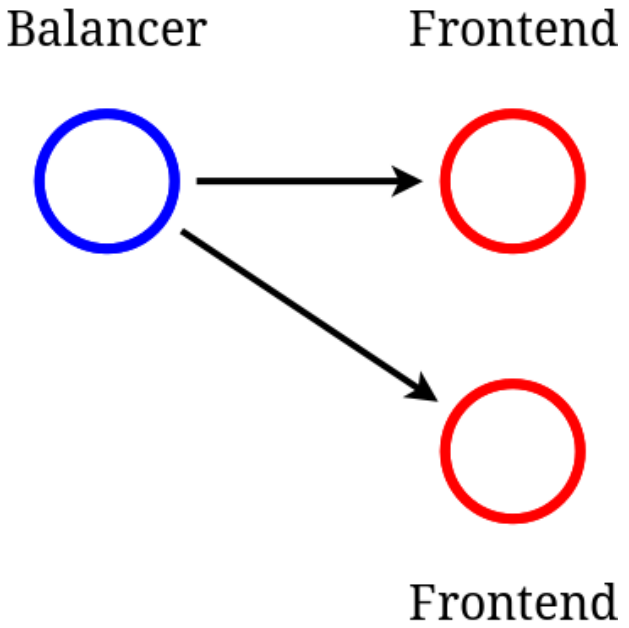
- Добавим компьютеров
- Соберем балансировщик нагрузки



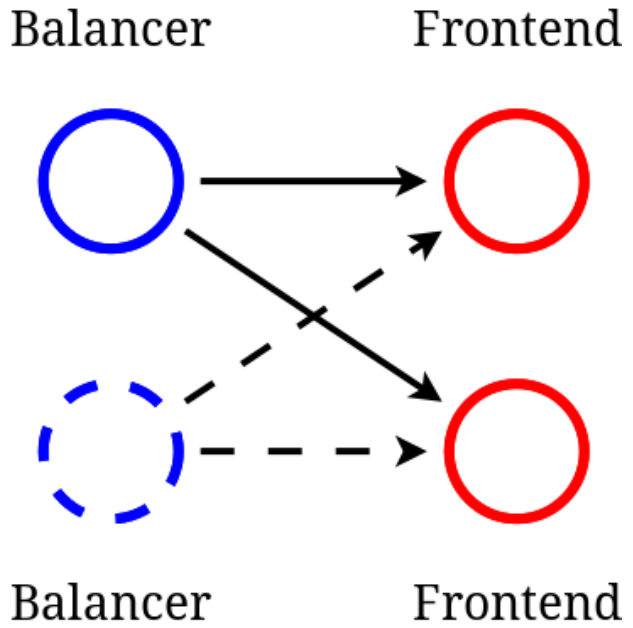
Популярные балансировщики

nginx	http://nginx.org
haproxy	http://haproxy.1wt.eu
ipvs	http://kernel.org

Точка отказа осталась



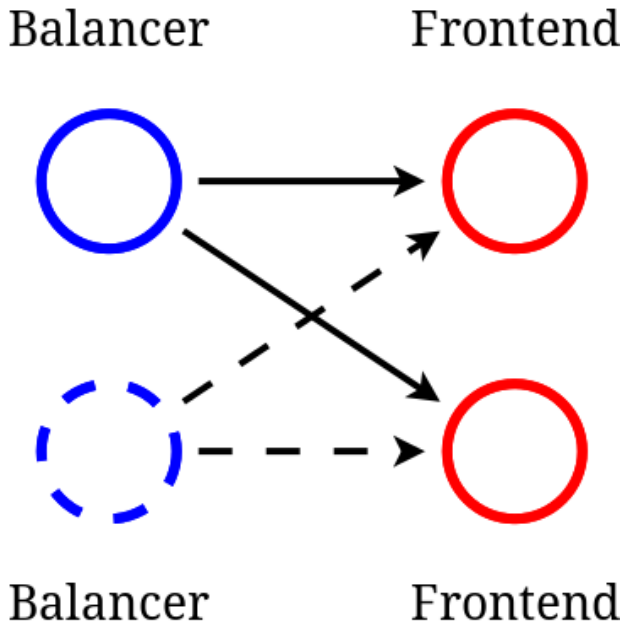
Точка отказа осталась



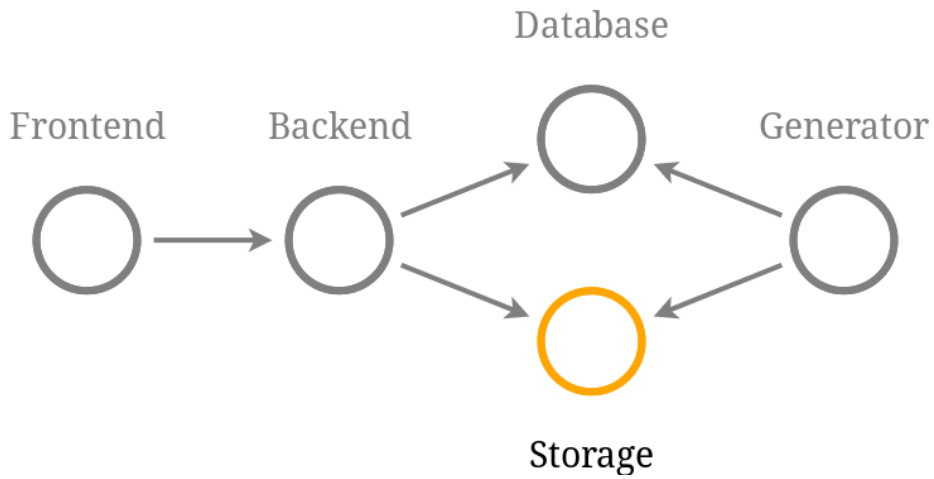
Точка отказа осталась

Нужен умный протокол

- heartbeat
- ospf



Отказоустойчивость хранилищ



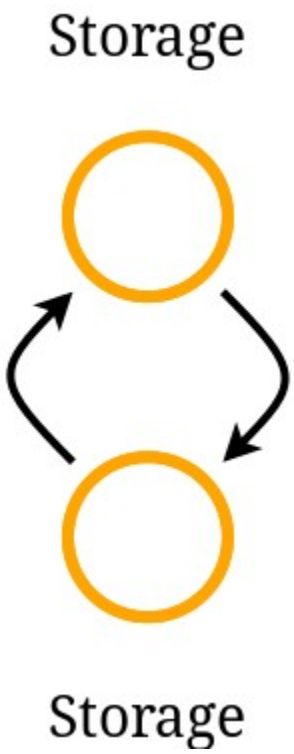
Что это и зачем это?

Храним большие объемы данных:

- тексты
- картинки
- легальную ;-) музыку
- ролики ~~и лыжи~~

Как хранить?

- Самый простой способ — в файликах
- Добавить способ синхронизации

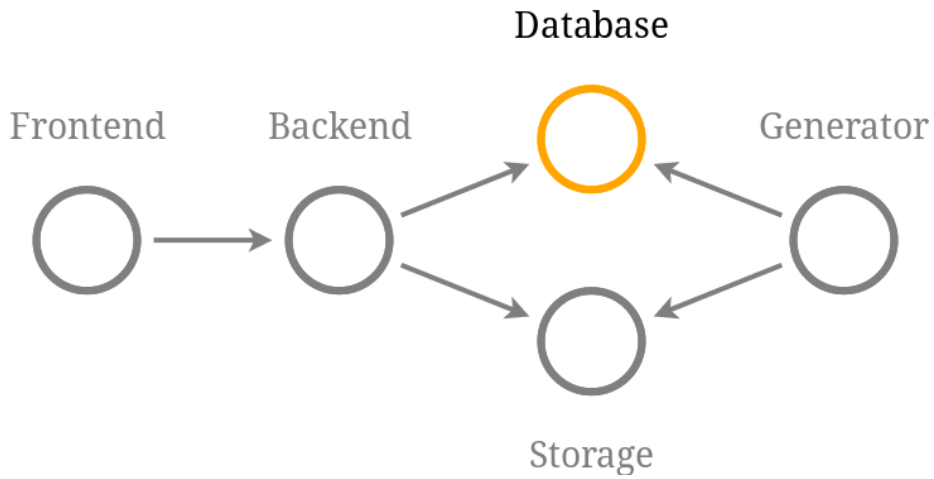


Как еще хранить?

Использовать отказоустойчивое хранилище:

- elliptics
- HDFS
- GPFS
- mulca

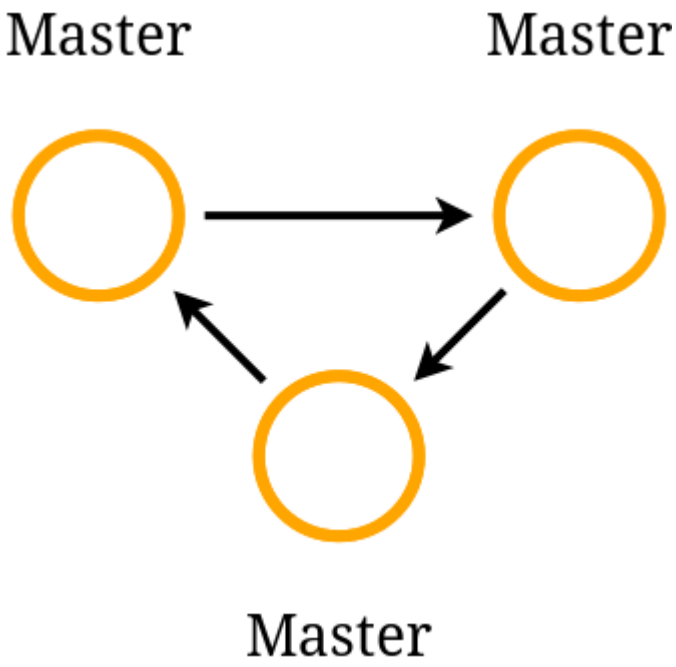
Отказоустойчивость баз данных



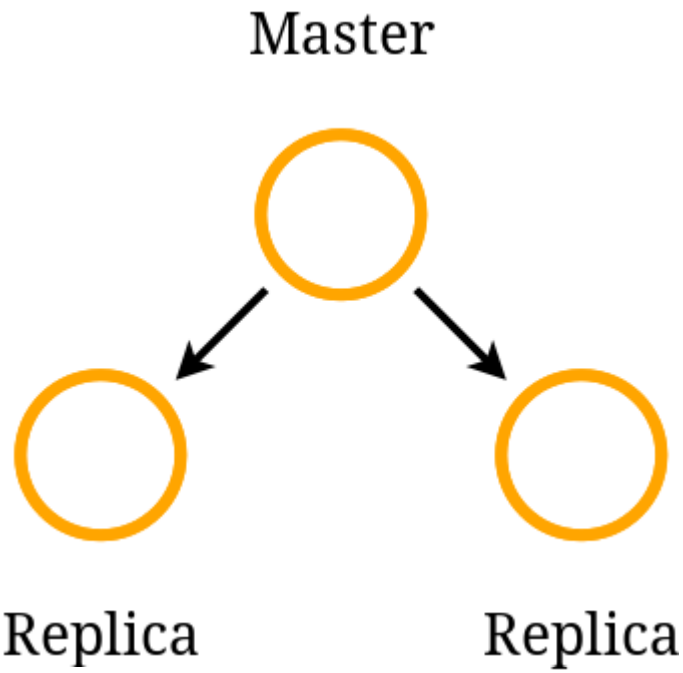
Проблема баз данных

Не предназначены для обеспечения отказоустойчивости

Стандартная репликация MySQL



Переключаемый мастер



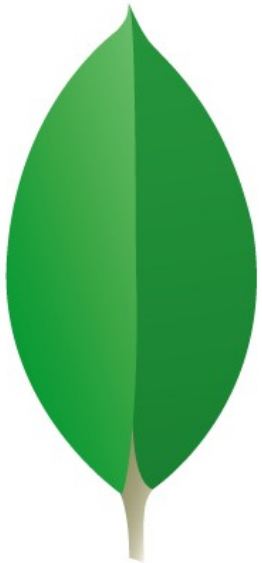
Как переключать?

- mmm — один датацентр
- Собственные механизмы переключения
  - DNS
  - OSPF
  - Метабаза

Другие базы данных?

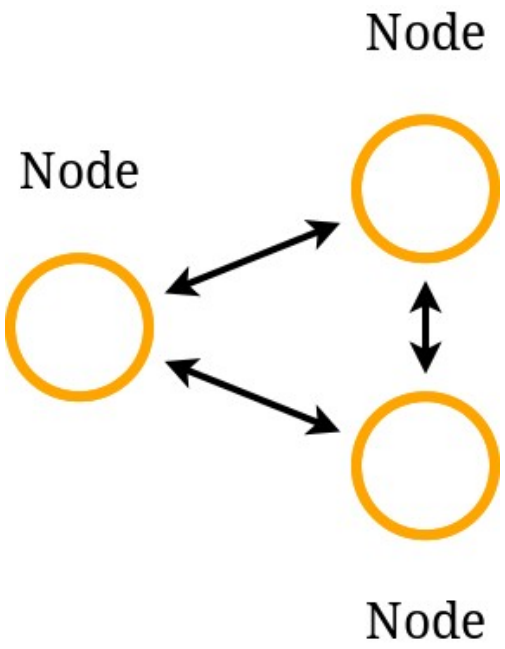
Базы попроще: MongoDB

- умеет сама переключать мастера
- зависит от клиентской библиотеки

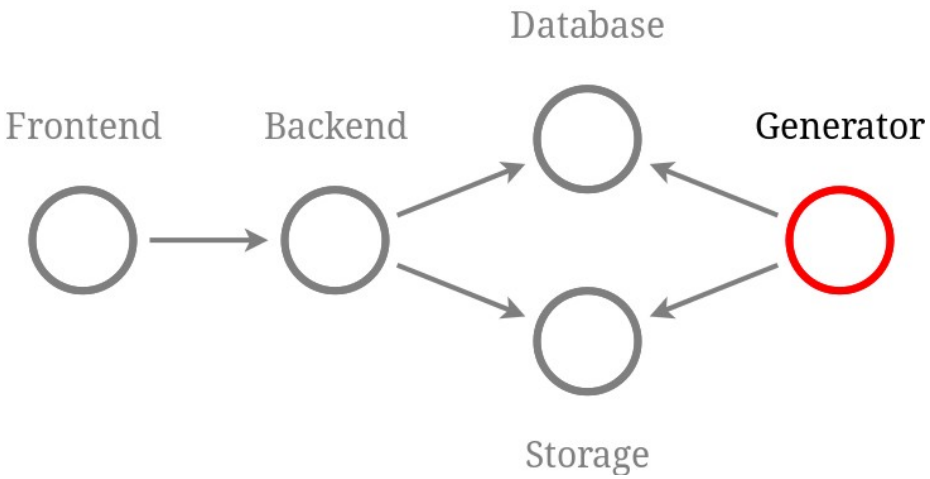


Кластерные базы

- MySQL cluster
- Galera cluster
- Oracle RAC



Отказоустойчивость генераторов



Почему не как фронтенды?

- Задача запускается один раз
- Но при этом хочется, чтобы компьютеры не простаивали :-)

Система синхронизации

Используем распределенную систему синхронизации

- dispofa
- zookeeper

Вопросы?

Павел Пушкарев
Руководитель группы администрирования

paulus@yandex-team.ru
@riarheos