



# Архитектура Яндекс.Маркета вчера и сегодня

Роман Гершкович  
Юрий Манушкин

---

Я.Субботник, Екатеринбург, 20.09.14

# КТО МЫ

- группа администрирования
- разработки выдачи и индексации
- frontend
- MBI, MBO, IR, PERS

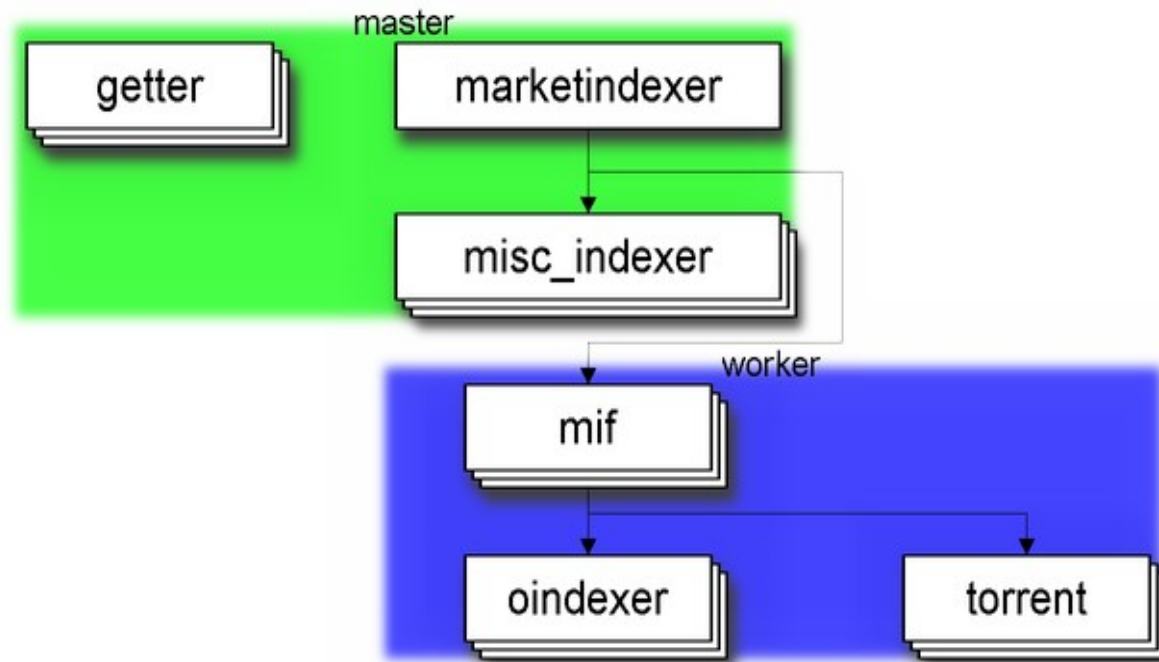
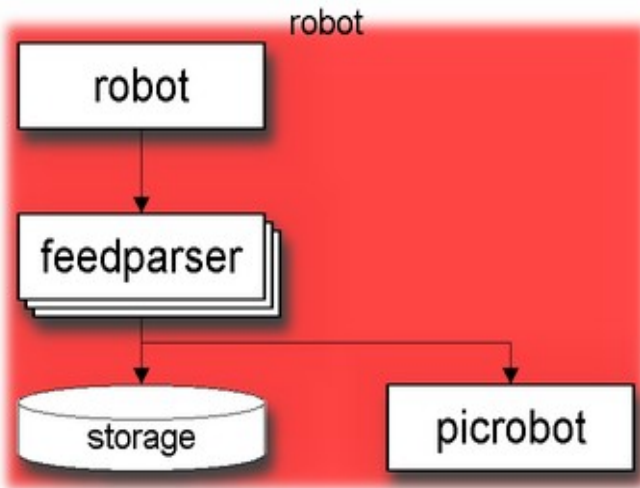
# Про что

- подготовить данные (индексация)
- разложить данные на бекенды (раскладка)
- выдача

# Индексация: задачи

- качать предложения магазинов
- строить поисковый индекс
- считать различные статистики (#средняя цена)
- проксировать другие данные

# Индексация: устройство



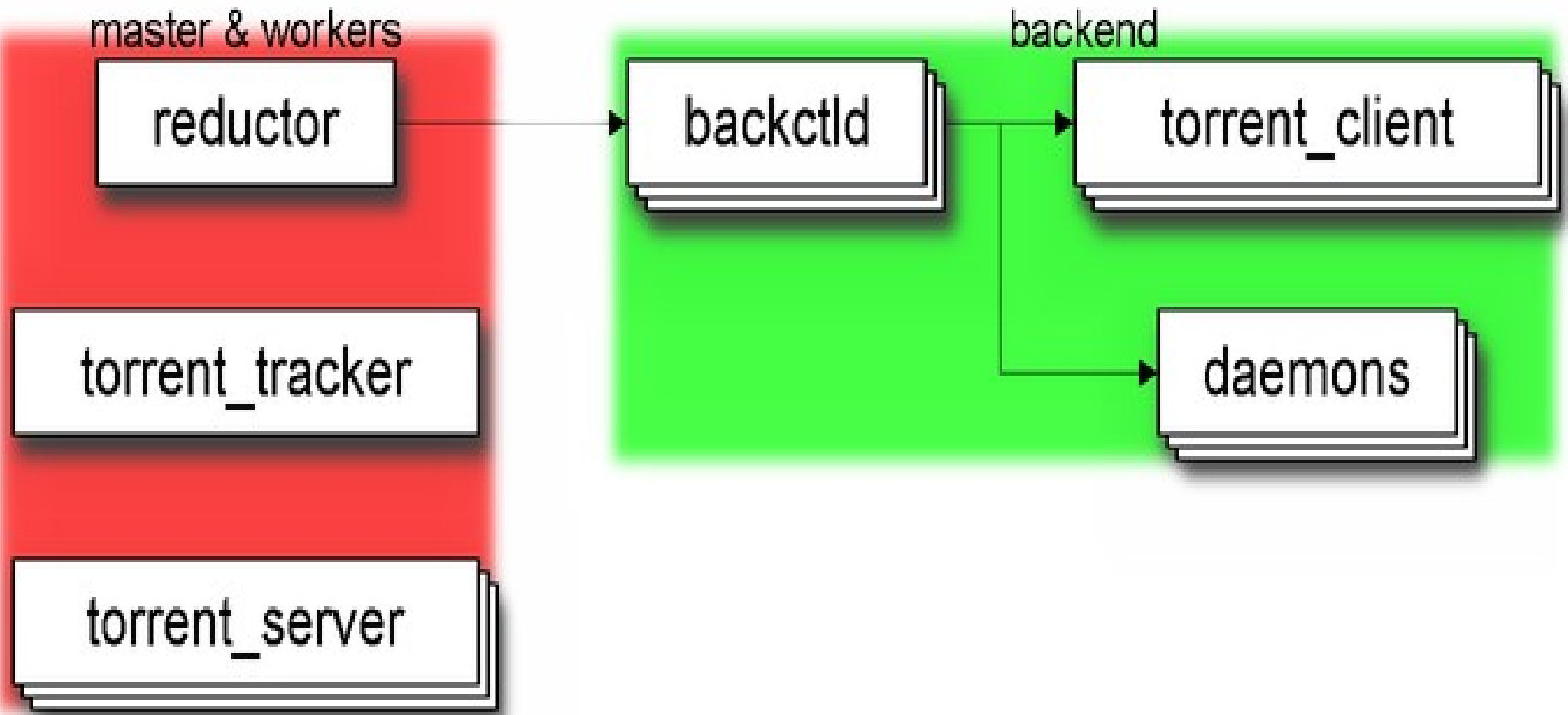
# Масштабирование и отказоустойчивость

- мастер + 8 воркеров (b&c)
- 2 индексатора в разных ДЦ (zk)
- что делать, если сломался мастер/воркер?
- расчет статистик в hadoop
- MySQL => hbase (in progress)

# Раскладка: задачи

- разложить данные (#поисковый индекс)
- плавный restart/reload компонент на бекендах
- устойчивость к отказу части машин (отваливается ethernet, разваливается raid, ломается switch)

# Раскладка: устройство





# Раскладка: числа

- индексация 2 часа
- 200GB данных
- раскладка 1 час
- 500 машин
- асинхронность
- эффективно использовать канал

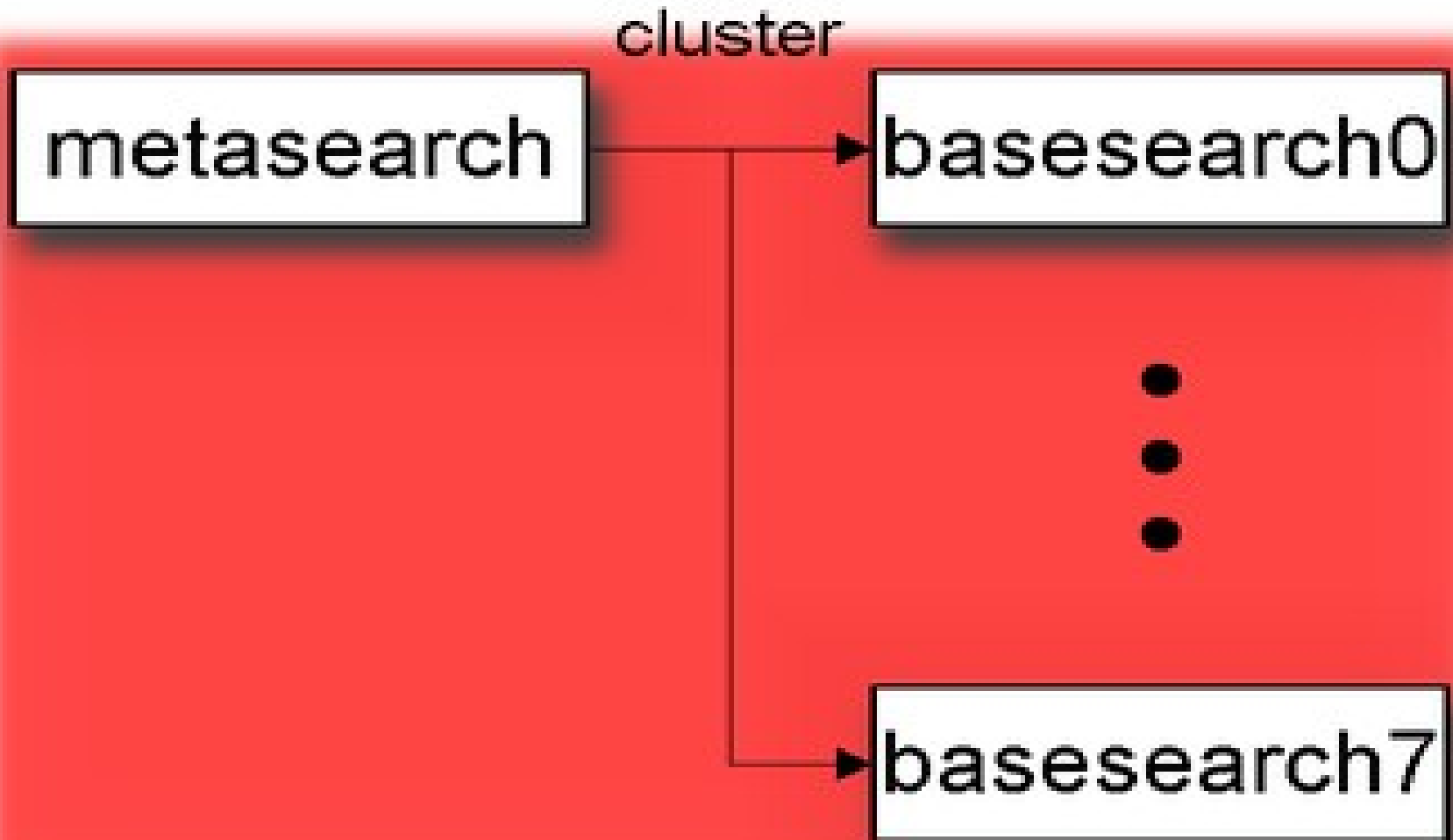
# Выдача

- CORBA
  - много приложений
  - малая и средняя нагрузка
  - десятки машин
- search
  - только поиск
  - большая нагрузка
  - сотни машин

# Поиск: числа

- кластер = 8 машин
  - 2 процесса поиска на машине (утилизация CPU, экономия памяти)
  - RPS
    - mrate: 20K
    - serp: 8K
    - market: 1.5K
-

# Поиск: устройство



# Что дальше

- 200М предложений
- обновление в реальном времени
- улучшать MIR

# Спасибо!



Рома: [rgersh@yandex-team.ru](mailto:rgersh@yandex-team.ru)

Юра: [manushkin@yandex-team.ru](mailto:manushkin@yandex-team.ru)