

«Габри» при масштабировании веб-приложения

или «что происходит при росте с 1М до 50М хитов в сутки».

Евгений Коковихин, ведущий разработчик Warstart.

(в конце доклада вам покажут 1000 слонов!)



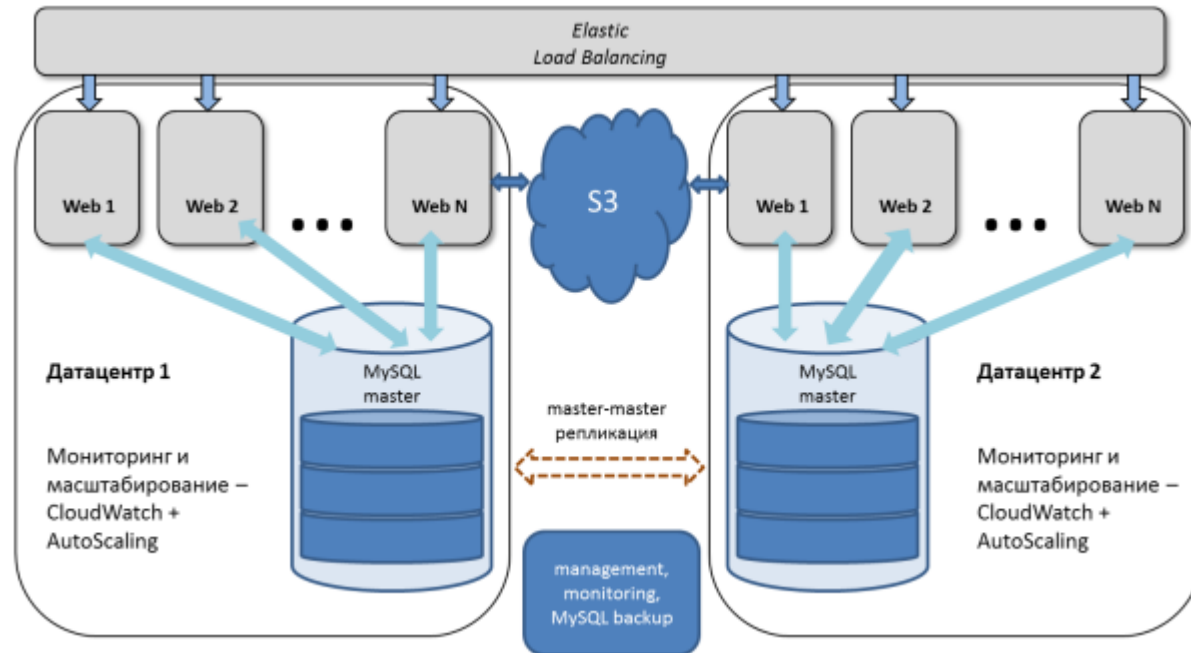
<http://www.devconf.ru>

Кто мы?

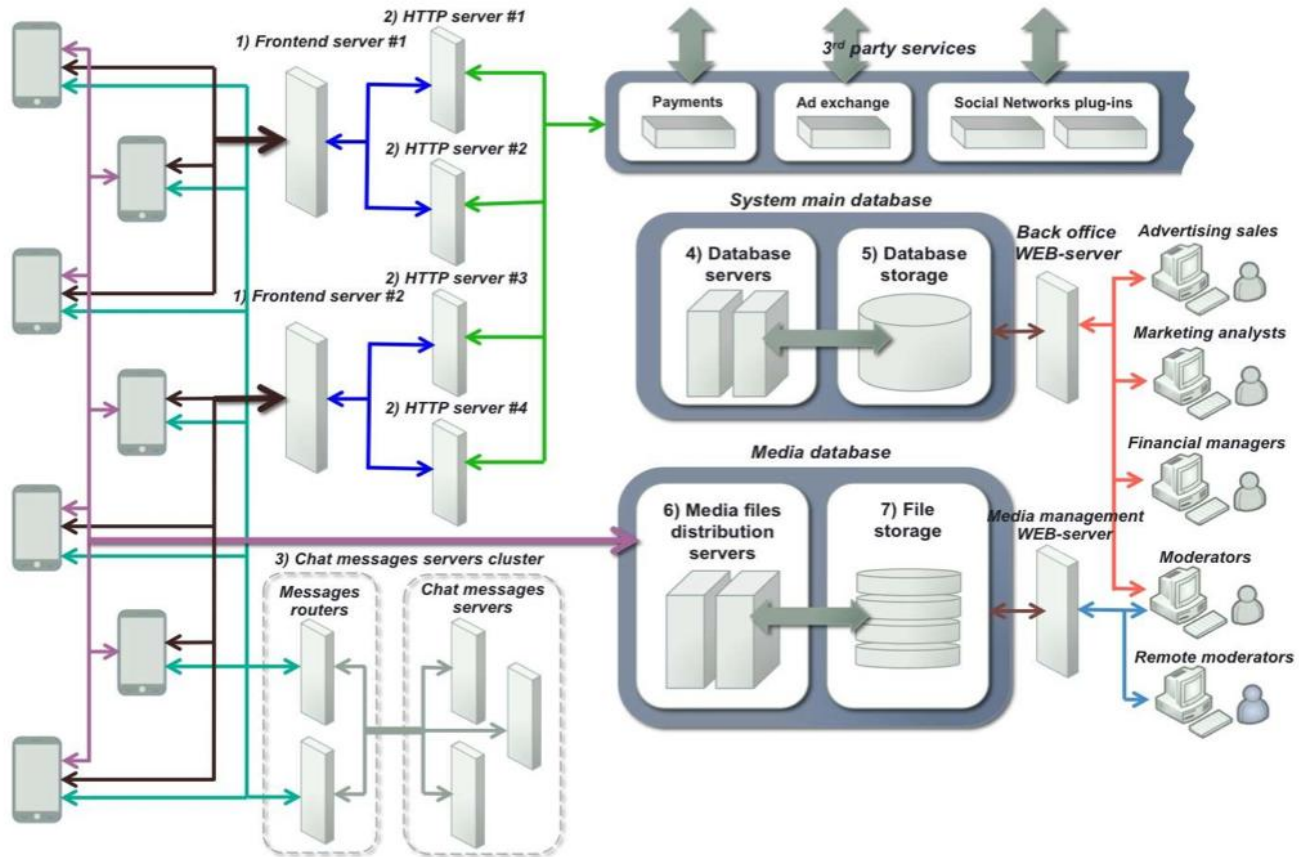
- Компания WapStart - владелец крупнейшей в России мобильной рекламной сети Plus1 WapStart и каталога мобильных сайтов Top WapStart.
- Компания оказывает услуги тысячам издателей мобильных сайтов и приложений.
- Ежемесячная рекламная емкость WapStart - свыше 1,5 млрд. показов, число уникальных посетителей - более 10 миллионов в месяц.
- Компания WapStart владеет передовыми технологиями таргетинга и анализа аудитории, позволяющими проводить рекламные кампании с высокой рентабельностью инвестиций.

© Отдел маркетинга.

Эту картинку я нашел в интернете



... И ЭТУ



Задачи, которые решает plus1.wapstart.ru

- Поиск подходящих баннеров для пользователя.
- Сортировка баннеров по релевантности и стоимости.
- Подсчет денег.
- Администрирование баннеров и рекламных площадок.
- Сбор и анализ статистики.

ПО, которое используется в plus1.wapstart.ru

- PHP 5.3 / php-fpm
- postgresql 9.1
- nginx
- memcache
- FreeBSD (fronts)
- Debian (farms)

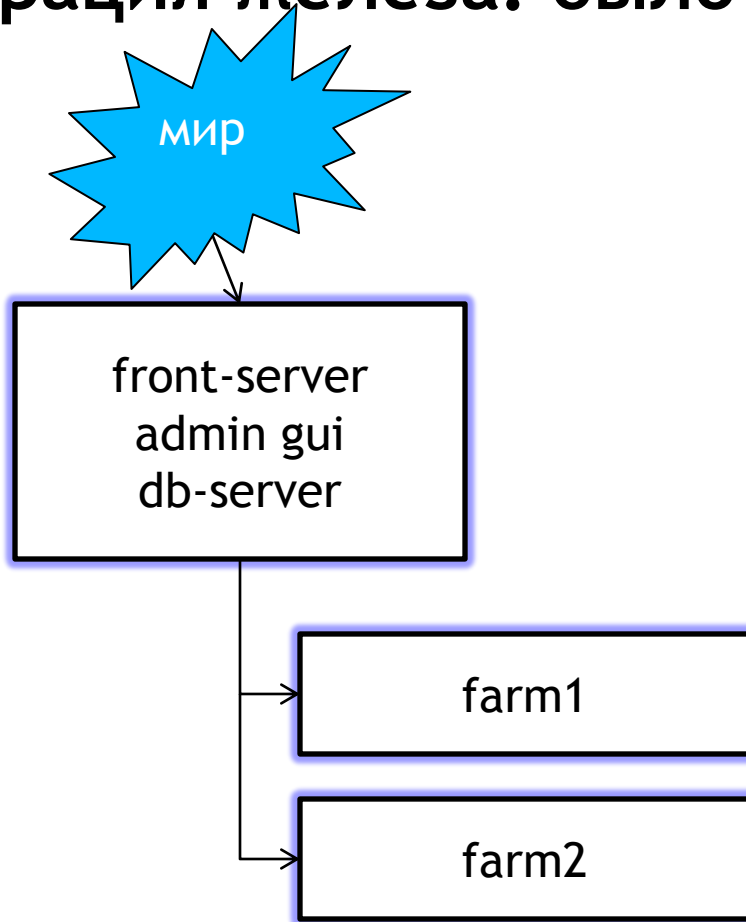
Конфигурация железа: было

Хорошо:

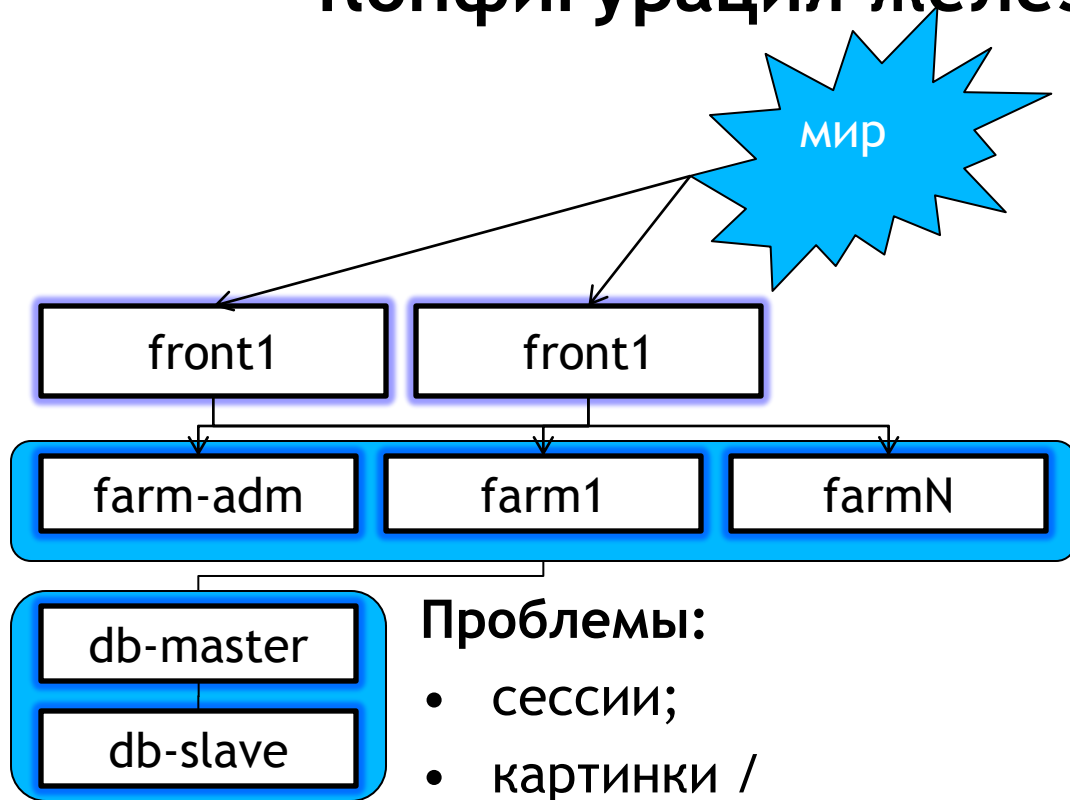
- Хорошо работает при небольшой нагрузке;
- Все очень просто.

Плохо:

- Разные требования к железу у разных ролей;
- Сложно отлавливать баги и искать узкие места из-за влияния других подсистем.



Конфигурация железа: стало



Хорошо:

- все еще просто;
- легко добавлять и убирать железо;
- железо разделено по ролям.

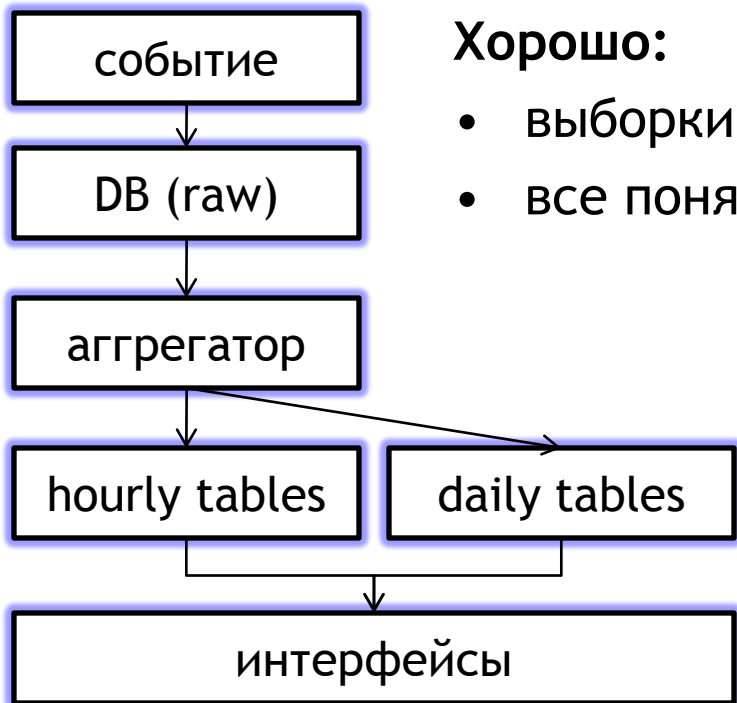
Плохо:

- Вывод железа из продакшена в случае отказа не автоматизирован.

Проблемы:

- сессии;
- картинки / пользовательские файлы;
- sql запросы, модифицирующие данные

Статистика (1)



Хорошо:

- выборки очень просты;
- все понятно;

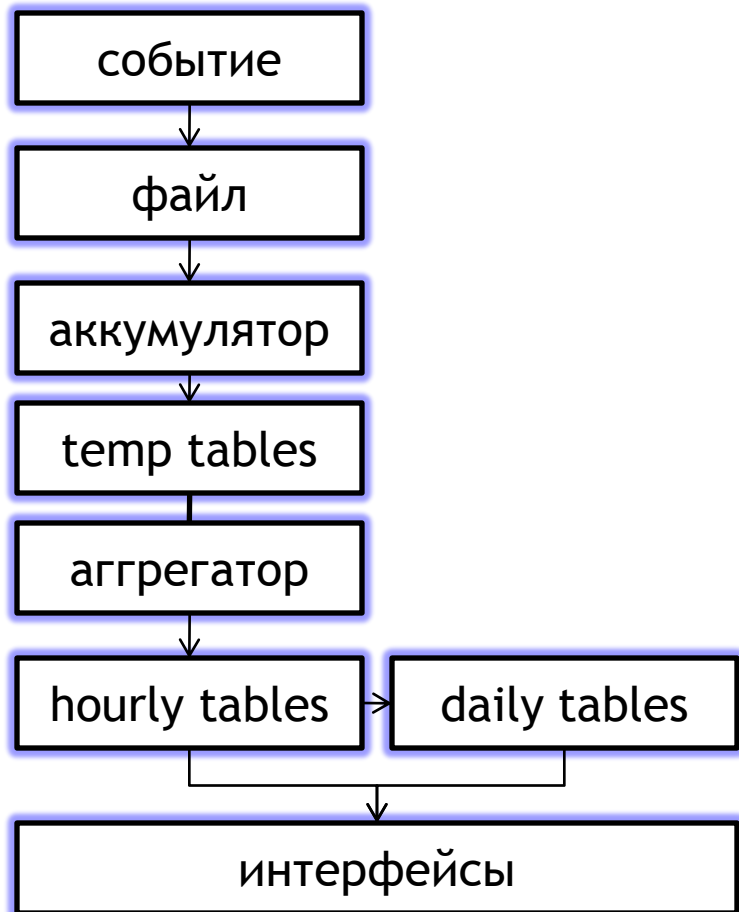
Плохо:

- медленно;
- сложно удалять старые сырые данные;
- сложно масштабировать.

Проблемы:

- Однажды мы начали писать быстрее, чем смогли обрабатывать;
- Внешние ключи не только полезны, но и вредны.

Статистика (2)



Хорошо:

- выборки очень просты;
- все понятно;

Плохо:

- Временные таблицы надо чистить;

Статистика (3)



Хорошо:

- Агрегация работает годно;

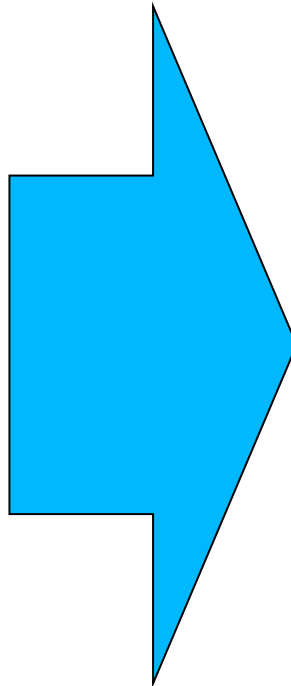
Плохо:

- Данные в дневные таблицы попадают не одновременно с часовыми (переживем).

Биллинг: было

Утро:

 Рекламодатель 100 рублей	 Владелец сайта 100 рублей
 Рекламодатель 150 рублей	 Владелец сайта 10 рублей
 Рекламодатель 180 рублей	 Владелец сайта 20 рублей




Вечер:

 Рекламодатель 50 рублей	 Владелец сайта 120 рублей
 Рекламодатель 70 рублей	 Владелец сайта 12 рублей
 Рекламодатель 100 рублей	 Владелец сайта 30 рублей

Биллинг: стало

Счет	сумма
User 1	1 000
User 2	100
User 3	100
System	150
...	...
...	...

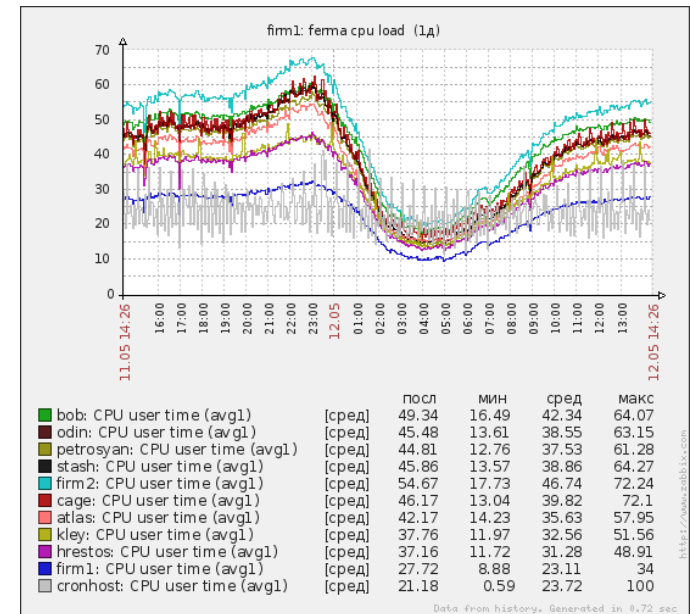
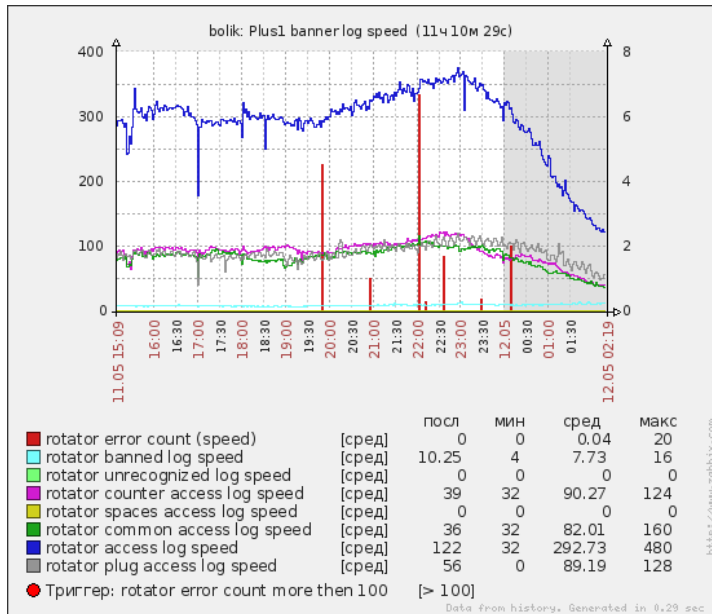
Счет	сумма
User 1	1 100
User 2	200
User 3	500
System	3 000
...	...
...	...



time	Credit_bill_id	Debet_bill_id	Amount	comment
...	User_1	User_2	100	за клики
...

Мониторим всё

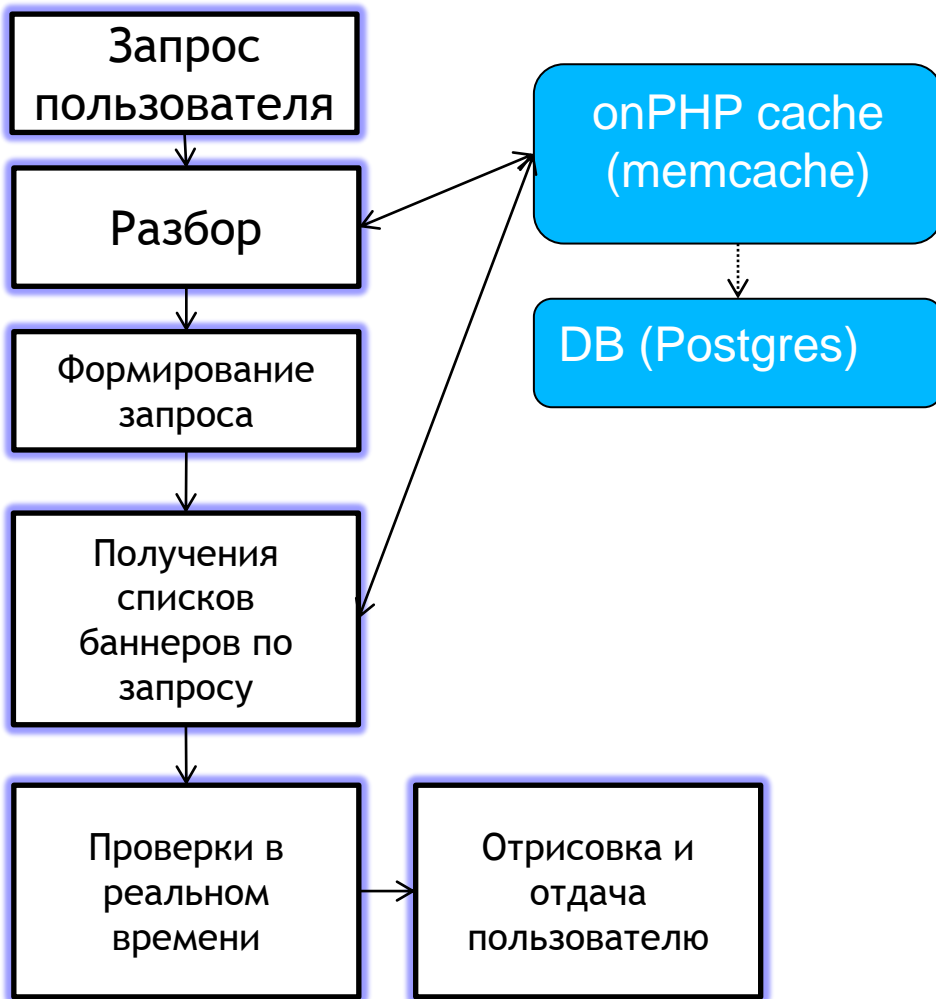
1. Состояние логов
2. Агрегация ошибок и отправка их почтой
3. Цветные графики
4. Особо критичные триггеры по sms.



Борьба за быстродействие:

- клиенты хотят получать баннер быстро;
- клиенты получают баннеры синхронно;
- аренда оборудования - это дорого.

Борьба за быстродействие: было



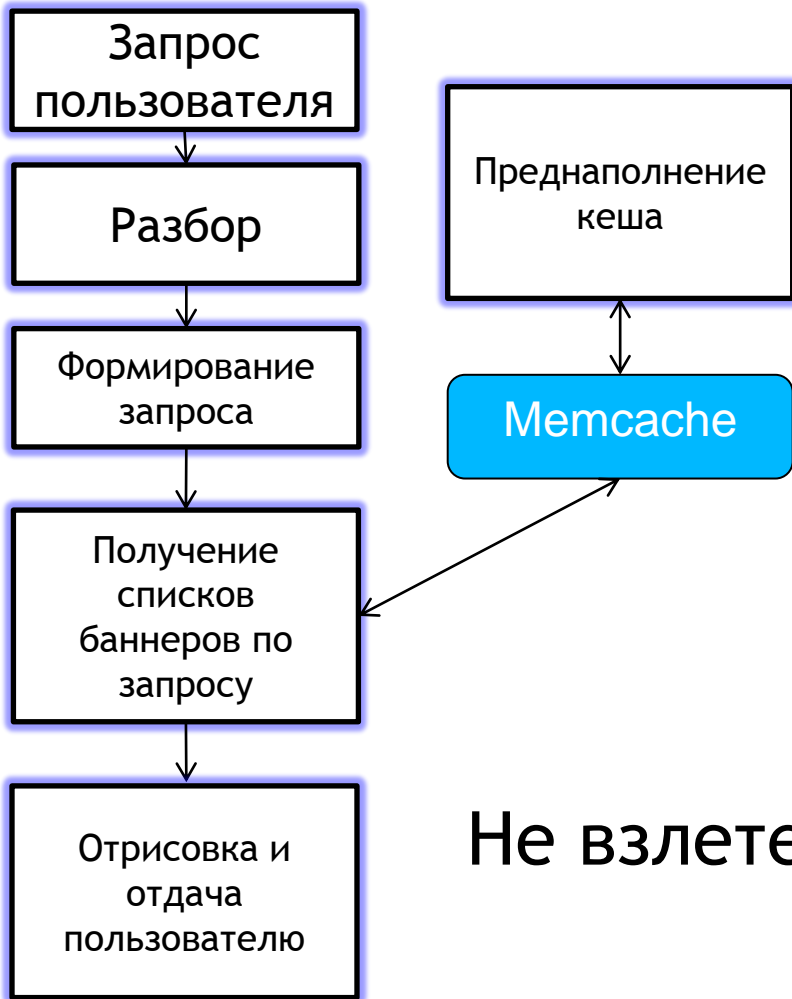
Плохо:

- Иногда ходим в базу;
- Используем родные объекты onPHP - медленная сборка.

Проблема:

- Однажды оно перестало помещаться в кеш.

Борьба за быстродействие: как хотели.

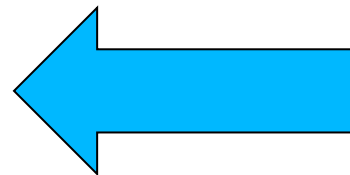
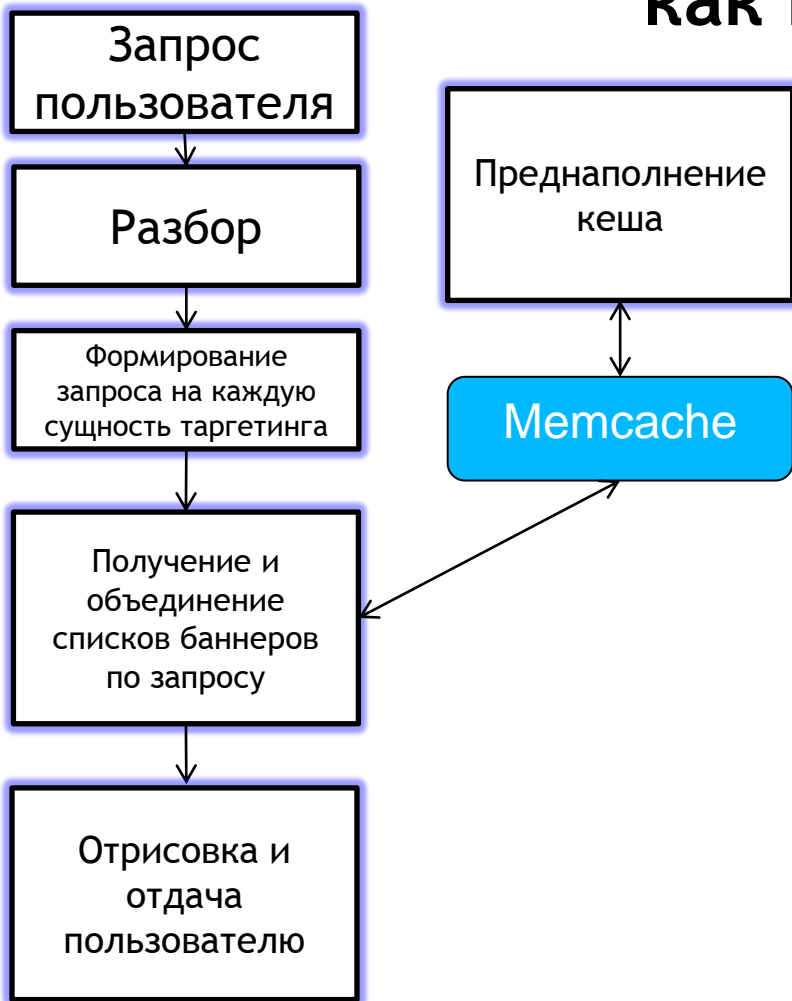


```

get model_id=42&gateway_id=1&region_id=199
VALUE model_id=42&gateway_id=1&region_id=199 0 256
a:2:{i:0;0:11:"PlainBanner":13:{s:15:"PlainBannerid";i:31
;s:17:"PlainBannername";N;s:18:"PlainBannertitle";N;s:20:
"PlainBannercontent";N;s:30:"PlainBannersingleLineContent
";N;s:17:"PlainBannerlink";N;s:20:"PlainBannerstateId";N;
s:23:"PlainBannercampaignId";N;s:19:"PlainBannertypeId";N
;s:21:"PlainBannerpictures";a:0:{}s:25:"PlainBannerfakePi
ctures";a:0:{}s:17:"PlainBannerfake";b:0;s:32:"PlainBanne
rshowOnlyWithGraphic";N;}i:1;0:11:"PlainBanner":13:{s:15:
"PlainBannerid";i:42;s:17:"PlainBannername";N;s:18:"Plain
Bannertitle";N;s:20:"PlainBannercontent";N;s:30:"PlainBan
nersingleLineContent";N;s:17:"PlainBannerlink";N;s:20:"Pl
ainBannerstateId";N;s:23:"PlainBannercampaignId";N;s:19:"
PlainBannertypeId";N;s:21:"PlainBannerpictures";a:0:{}s:2
5:"PlainBannerfakePictures";a:0:{}s:17:"PlainBannerfake";
b:0;s:32:"PlainBannershowOnlyWithGraphic";N;}}
END
  
```

Не взлетело :(

Борьба за быстродействие: как получилось



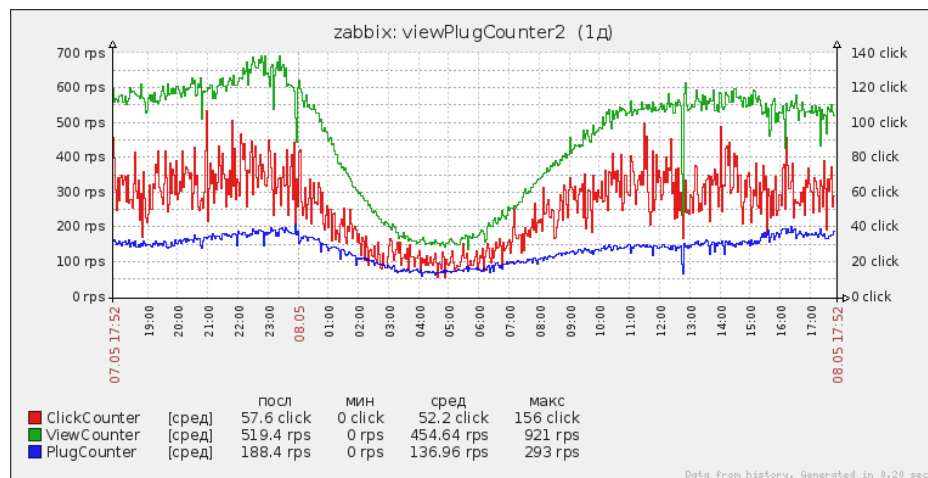
Вот так работает!

Проблемы:

- Разбиратель запроса все еще может попасть в базу;
- Редкий трафик идет в базу целиком;
- За преднаполненным кешем надо следить.

Мониторинг (2) и как мы раньше жили без pinba?

- Меряем именно то, что происходит на проде;
- Меряем все, что происходит на проде;
- Можем измерять бизнес-метрики;
- Красивые графики.



Резюме:

- Разработка - процесс итеративный.
- Ошибки бывают, не бойтесь ошибаться.
- Не надо думать на 3 года вперед, все равно все изменится.
- Мониторинг - наше всё.
- Сила в простых вещах. Когда все просто - это легко поддерживать и модифицировать.

Wapstart:

- <http://wapstart.ru/> <https://plus1.wapstart.ru/>
- Продажи: sales@co.wapstart.ru
- Разработка: dev@co.wapstart.ru
- github: <https://github.com/Wapstart>
- Habr: <http://habrahabr.ru/company/wapstart/>

Я:

- Mail/jabber: dovg@dovg.ru
- Github: <https://github.com/dovg>
- Habr: <http://habrahabr.ru/users/dovg/>
- Vk: <http://vk.com/id2082364>

Вопросы?