

# Архитектура А/Б тестирования: сделай сам

Сергей Аверин  
Vadoo



- Социальная сеть для знакомств
- 195 миллионов зарегистрированных пользователей
- 10 миллионов пользователей в день
- 2000+ серверов
- 50 тыс. динамических запросов/сек
- Разработка в Москве и Лондоне
- MySQL, PHP, C/C++, Linux, nginx, memcached
- Много своего, в т. ч. PINBA, PHP-FPM, патчи PHP, PHPUnit, MySQL

# Для кого доклад?

Для сайтов, где

- пользователи идентифицированы
- есть возможность долгосрочно отслеживать пользователя
- пользователи нужны надолго, и заходят на сайт периодически
- нужны более сложные методики оценки успеха, чем единоразовая покупка/регистрация/заказ/звонок

# Сложность

- Выше среднего, но ничего космического тут нет
- Я уверен, что вы можете сделать то же самое за 3 недели силами 2—3 разработчиков
- Для небольших сайтов архитектура будет сильно похожа
- Результат можно измерять в реальной прибыли

# Зачем нам понадобились А/В тесты

Раньше мы запускали фичу в маленькой стране и сравнивали показатели с цифрами до запуска или в соседних странах

Но... мы растем:

- Для всех идей стран не хватит
- Нужно выбрать из нескольких похожих вариантов
- Хочется более достоверной методики оценки
- В самом худшем случае урон невелик

Почему мы сделали свое

# Мы пробовали чужое

- Смотрели несколько сторонних продуктов
- Выбрали и использовали Optimizely...

PROJECTS

▼ My Project

📊 Experiments

👤 Visitor Segments 🔒

Create New Project


No archived projects

# My Project

Project Settings

<Project Code>

 EXPERIMENTS

 VISITOR SEGMENTS 🔒





Create Experiment



## Active Experiments

There are no active experiments.

## Draft Experiments

EXPERIMENT NAME ▾	SCHEDULED	GOALS	LAST MODIFIED
badoo2		1 	Jan 24, 2013
badoo3		1 	Jan 24, 2013
badoo4		1 	Jan 28, 2013
demo		1 	Oct 24, 2013

## Archived Experiments



Did you know archiving experiments reduces the size of your snippet? [View Archived Experiments](#)



badoo4 

Set Up Goals



Start Experiment

Experiment Variations

Saved 9 months ago



Save now



Undo



Redo



Original

Variation #1

Variation #2 ▾

+ Add Variation

Editing Mode

Interactive Mode

Options



English ▾

Already a member? [Sign in](#)[Join Badoo](#)

# Meet new people

With over 195m users, Badoo is great for chatting, making friends, sharing interests, and dating!

Edit HTML

Edit Text

Edit Classes

Make Hyperlink

Edit Style

Paragraph &lt;p&gt;

Edit Element... ▸

Insert... ▸

Move and Resize

Rearrange

Remove

Track Clicks... ▸

Custom Javascript

# Проблемы Optimizely и подобных систем

- jQuery
- Внешний CDN
- JS-совместимость с нашим кодом
- «Прыгающие» страницы
- Нет локализации, поддержки разных языков
- Выбор пользователей по языку, полу — все сам, в PHP-коде
- Ограниченное применение — нет API, только JS-код

**Но главное — не это**

# А то, что анализируем только клики

Большинство докладов про А/Б тестирование — из мира e-commerce:

- анализируем посещаемость трех страниц — главной, каталога товаров, страницы товара
- анализируем воронку продаж (где пользователь уходит)
- улучшаем продажи (клики по кнопке «заказать»)
- Купил? Следующий!

# Мы — не интернет магазин

- Анализируем поведение — частота заходов, время на сайте, кол-во чатов, заливка фото...
- Нужна статистика по большому количеству параметров
- И за больший срок, чем от момента захода на сайт до клика по цели
- Продажи — не самое важное

# Чем свое лучше и хуже

## Плюсы

- Большая применимость
- Большие нагрузки
- Никакого JS с внешнего CDN
- Весь ваш функционал локализации/переменных доступен

## Минусы

- Надо кодить
- Нагрузка теперь наша
- Статистику считать самим
- Сделать свой WYSIWYG-интерфейс редактирования сложно (но и не нужно)

Как у нас устроены тесты

# Someone likes you!



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec eget fermentum diam. Ut vulputate, arcu in tincidunt volutpat, turpis lacus facilisis leo, a tincidunt est erat at velit. Vivamus dictum blandit congue. Nullam rhoncus vel nibh et posuere. Phasellus imperdiet orci non odio suscipit semper. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nulla facilisi. Morbi vestibulum vitae massa eu ultricies. Vestibulum a eros eget eros porta venenatis ut sed ligula. Nam in leo interdum, mattis dolor pellentesque, facilisis ligula. Nullam in nunc semper, vulputate nulla dignissim, placerat augue.

Go for it!

Письмо с плохим кликрейтом



# Someone likes you!



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec eget fermentum diam. Ut vulputate, arcu in tincidunt volutpat, turpis lacus facilisis leo, a tincidunt est erat at velit. Vivamus dictum blandit congue. Nullam rhoncus vel nibh et posuere. Phasellus imperdiet orci non odio suscipit semper. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nulla facilisi. Morbi vestibulum vitae massa eu ultricies. Vestibulum a eros eget eros porta venenatis ut sed ligula. Nam in leo interdum, mattis dolor pellentesque, facilisis ligula. Nullam in nunc semper, vulputate nulla dignissim, placerat augue.

[Go for it!](#)

Вариант для всех

# 1 girl voted for you!



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec eget fermentum diam. Ut vulputate, arcu in tincidunt volutpat, turpis lacus facilisis leo, a tincidunt est erat at velit. Vivamus dictum blandit congue. Nullam rhoncus vel nibh et posuere. Phasellus imperdiet orci non odio suscipit semper. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nulla facilisi. Morbi vestibulum vitae massa eu ultricies. Vestibulum a eros eget eros porta venenatis ut sed ligula. Nam in leo interdum, mattis dolor pellentesque, facilisis ligula. Nullam in nunc semper, vulputate nulla dignissim, placerat augue.

Go for it!

Вариант для мальчиков

# Someone likes you!

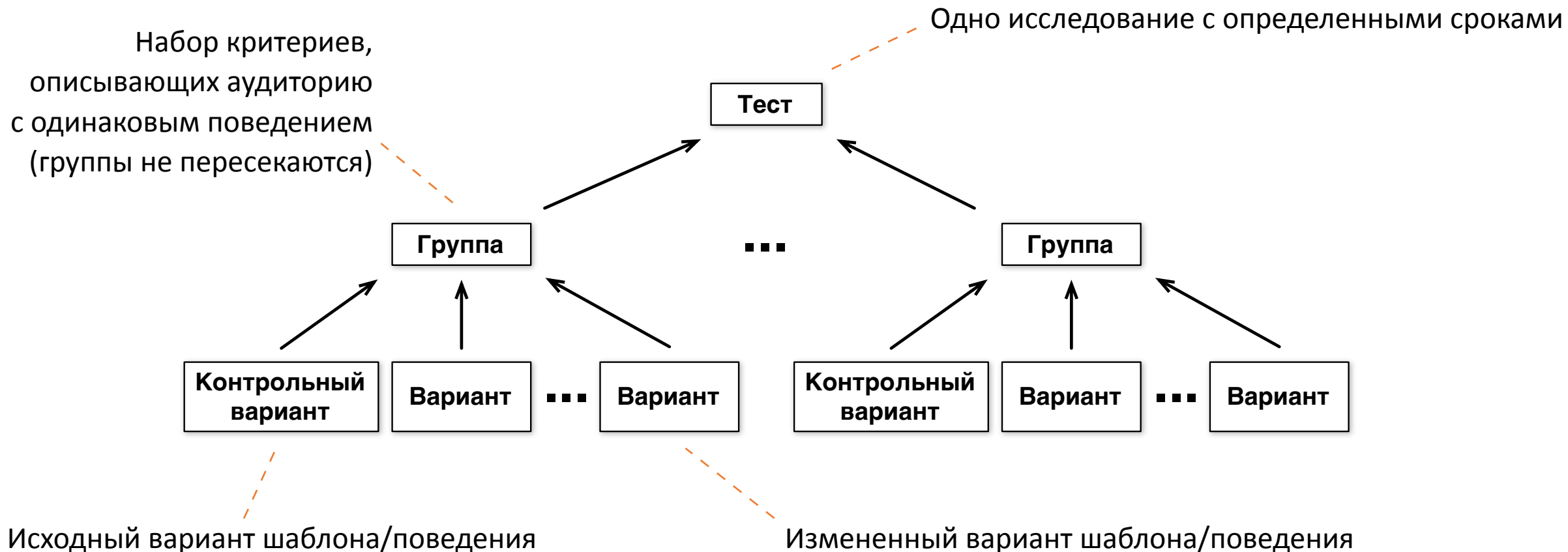


Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec eget fermentum diam. Ut vulputate, arcu in tincidunt volutpat, turpis lacus facilisis leo, a tincidunt est erat at velit. Vivamus dictum blandit congue. Nullam rhoncus vel nibh et posuere. Phasellus imperdiet orci non odio suscipit semper. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nulla facilisi. Morbi vestibulum vitae massa eu ultricies. Vestibulum a eros eget eros porta venenatis ut sed ligula. Nam in leo interdum, mattis dolor pellentesque, facilisis ligula. Nullam in nunc semper, vulputate nulla dignissim, placerat augue.

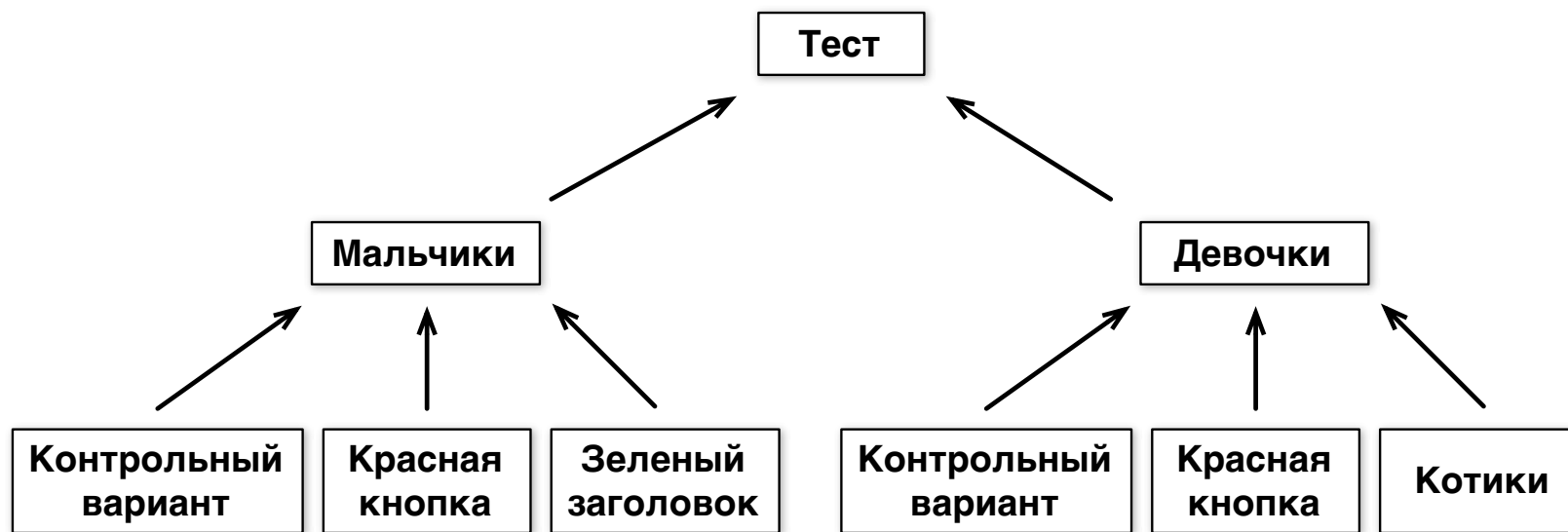
Go for it!

Вариант для девочек

# Структура теста



# Структура теста



# Результаты

Группа	Вариант	Отправлено	Кликов	Кликрейт
Мальчики	Контрольный	2000	100	5 %
	Красная кнопка	2000	110	5,5 % (+0,5 %)
	Зеленый заголовок	2000	130	<u>6,5 % (+1,5 %)</u>
Девочки	Контрольный	1000	35	3,5 %
	Красная кнопка	1000	30	3 % (-0,5 %)
	Котики	1000	50	<u>5 % (+1,5 %)</u>

**С ВЫСОТЫ ПТИЧЬЕГО ПОЛЕТА**

# Хотелось, что бы это работало так:

1) Заходим на страницу, выбираем на ней текст, создаем несколько версий текста. Или выбираем где-то в бекофисе константу, создаем тест с несколькими значениями этой константы.



# Хотелось, что бы это работало так:

~~1) Заходим на страницу, выбираем на ней текст, создаем несколько версий текста. Или выбираем где то в бекофисе константу, создаем тест с несколькими значениями этой константы.~~

1) Подготавливаем несколько шаблонов/css-файлов/значений

2) Настраиваем параметры теста в бекофисе

3) Вписываем в код выбор шаблона и привязку его к пользователю

4) Выбираем метрику, по которой мы будем оценивать результат

5) Прогоняем тест, собирая эту метрику и кто какой шаблон видел

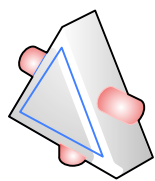
6) Берем кого-то умного и анализируем результаты



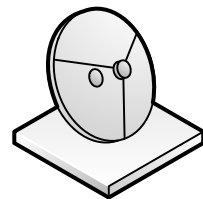
А теперь архитектура

# А теперь архитектура

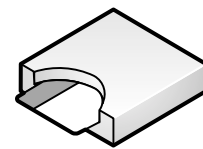
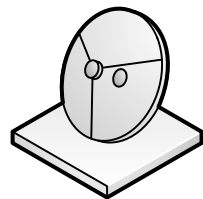
Сделаем сначала систему сбора статистики



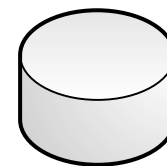
**API**



**Scribe-транспорт**

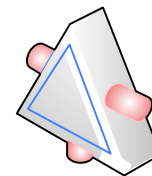


**Разборщик  
статистики**



**База данных**

# API



## Функция отправки событий

- просто пакет данные+текущую дату, и передает в транспорт

```
Stats::sendEvent($event_type, $user_id)
```

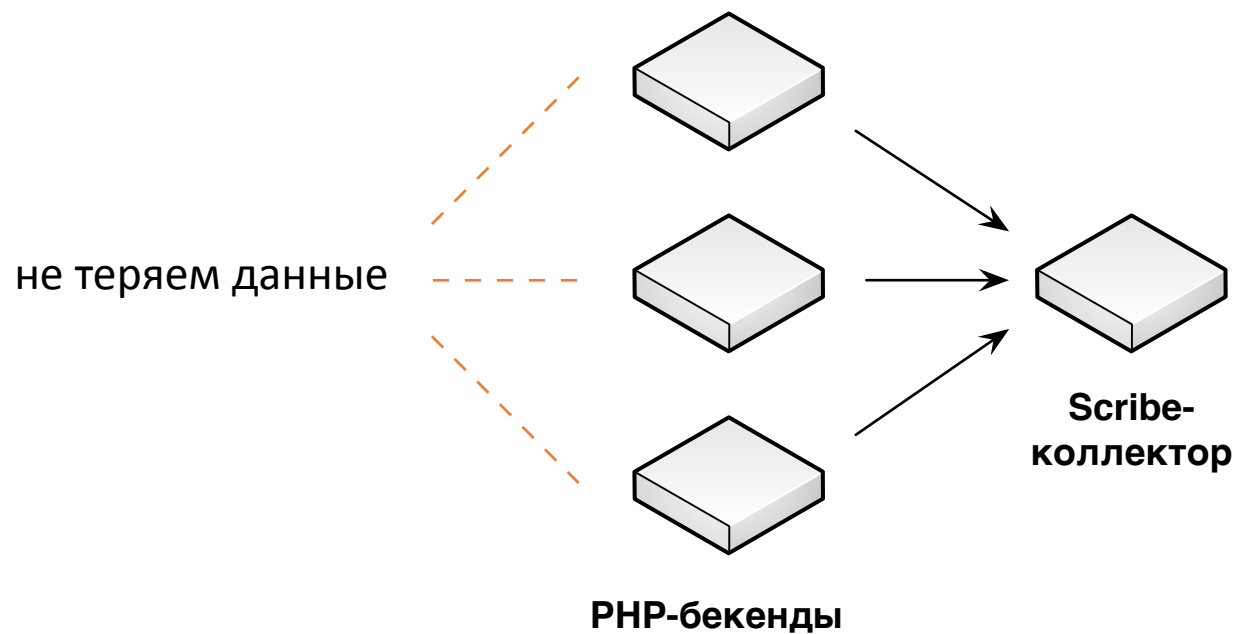
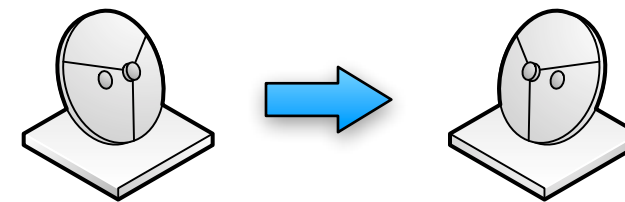
```
Stats::sendEvent($event_type, $user_id, $param)
```

тип события, строка

значение, например,  
«проголосовал за» — 1  
«проголосовал против» — 0

# Scribe-транспорт

100 000 000+ записей в день



файлы с данными,  
новый файл каждые 2 минуты,  
каждое событие в отдельной папке

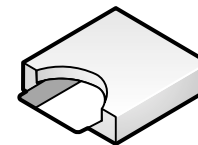
```
2. ssh
scripts40:scribe}~%(@> ls -la | grep -E 'bda_|ab_'
drwxr-xr-x  2 wwwrun www   520192 Сен 10 12:20 ab_fw_hit
drwxr-xr-x  2 wwwrun www  6615040 Сен  7  2012 bda_debug_redirect
drwxr-xr-x  2 wwwrun www  6877184 Окт 23 16:02 bda_presence_sid
drwxr-xr-x  2 wwwrun www  6934528 Май 16  2012 bda_presence_uid
drwxr-xr-x  2 wwwrun www 1560576 Окт 23 16:01 bda_push_queue_errors
drwxr-xr-x  2 wwwrun www  6299648 Май 16  2012 bda_usage_sid
drwxr-xr-x  2 wwwrun www  6144000 Май 16  2012 bda_usage_uid
drwxr-xr-x  2 wwwrun www   262144 Окт 23 16:02 local_ab_fw_hit
scripts40:scribe}~%(@> █
```

```
2. ssh  
bda_presence_sid-2013-10-23_00944  
bda_presence_sid-2013-10-23_00945  
bda_presence_sid-2013-10-23_00946  
bda_presence_sid-2013-10-23_00947  
bda_presence_sid-2013-10-23_00948  
bda_presence_sid-2013-10-23_00949  
bda_presence_sid-2013-10-23_00950  
bda_presence_sid-2013-10-23_00951  
bda_presence_sid-2013-10-23_00952  
bda_presence_sid-2013-10-23_00953  
bda_presence_sid-2013-10-23_00954  
bda_presence_sid-2013-10-23_00955  
bda_presence_sid-2013-10-23_00956  
bda_presence_sid-2013-10-23_00957  
bda_presence_sid-2013-10-23_00958  
bda_presence_sid-2013-10-23_00959  
bda_presence_sid-2013-10-23_00960  
bda_presence_sid-2013-10-23_00961  
bda_presence_sid-2013-10-23_00962  
bda_presence_sid-2013-10-23_00963  
bda_presence_sid-2013-10-23_00964  
bda_presence_sid-2013-10-23_00965  
bda_presence_sid-2013-10-23_00966  
bda_presence_sid_current  
scripts40:bda_presence_sid}<%%(@
```



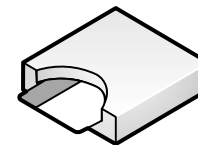
```
2. ex1 (~) - VIM (ssh)
1 20131023142047 | 248283903 | 5
2 20131023142047 | 304193811 | 5
3 20131023142047 | 343842353 | 1
4 20131023142047 | 332514236 | 5
5 20131023142047 | 281931214 | 5
6 20131023142047 | 161450473 | 5
7 20131023142047 | 336323314 | 5
8 20131023142047 | 317765286 | 5
9 20131023142047 | 344405780 | 1
10 20131023142047 | 344407284 |
~
~
~
~
~
~
~
~
~
~
~
```

# Скрипт-процессор статистики



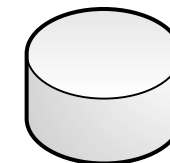
- Разбирает эти файлы, и кладет в БД
- Мультипоточный разбор с блокировками файлов (в PHP — flock())
- События каждого типа — в отдельные таблицы
- По одной таблице на день
- Выкидываем повторные события там, где это можно
- Там, где нельзя — делаем BULK INSERT'ы
- Для некоторых — агрегируем в памяти/файлах и пишем результат

# Скрипт-процессор статистики

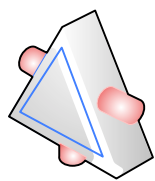


ts	user_id	value
2013-09-27 00:00:00	228041695	1
2013-09-27 00:00:00	228041695	1
2013-09-27 00:00:00	1202059665	2
2013-09-27 00:00:00	1202059665	2
2013-09-27 00:00:00	228041695	1
2013-09-27 00:00:00	228041695	1
2013-09-27 00:00:00	181248514	1
2013-09-27 00:00:00	330017771	1
2013-09-27 00:00:00	330017771	1
2013-09-27 00:00:00	330017771	1
2013-09-27 00:00:00	330017771	1
2013-09-27 00:00:00	1232338253	2
2013-09-27 00:00:00	1208234220	2
2013-09-27 00:00:00	1208234220	2

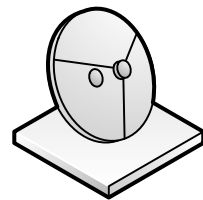
# БД



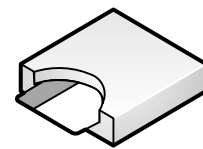
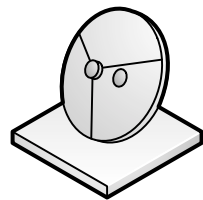
- Подойдет обычная SQL-СУБД, например, MySQL или Postgres
- Желательно отдельный сервер, с большим кол-вом ядер
- Для MySQL лучше использовать XtraDB



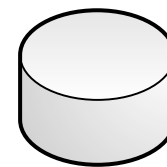
**API**



**Scribe-транспорт**

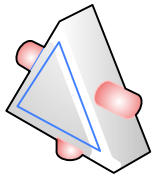


**Разборщик  
статистики**

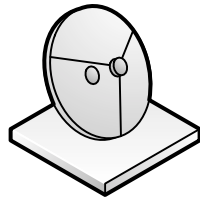


**База данных**

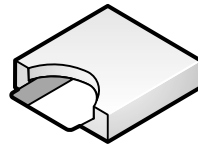
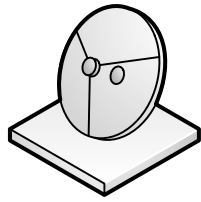
Дополним схему до  
архитектуры А/Б тестирования



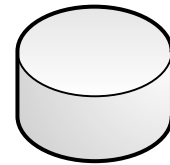
**API  
статистики**



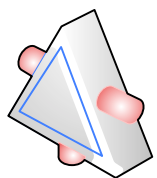
**Scribe-транспорт**



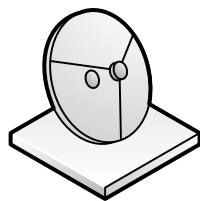
**Разборщик  
статистики**



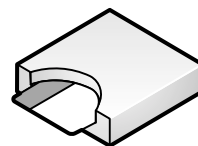
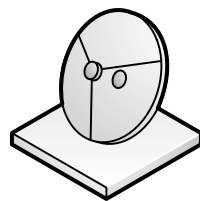
**БД  
статистики**



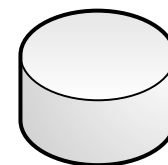
**API  
статистики**



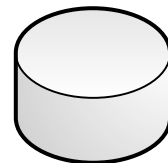
**Scribe-транспорт**



**Разборщик  
статистики**



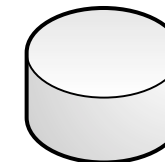
**БД  
статистики**



**БД тестов**

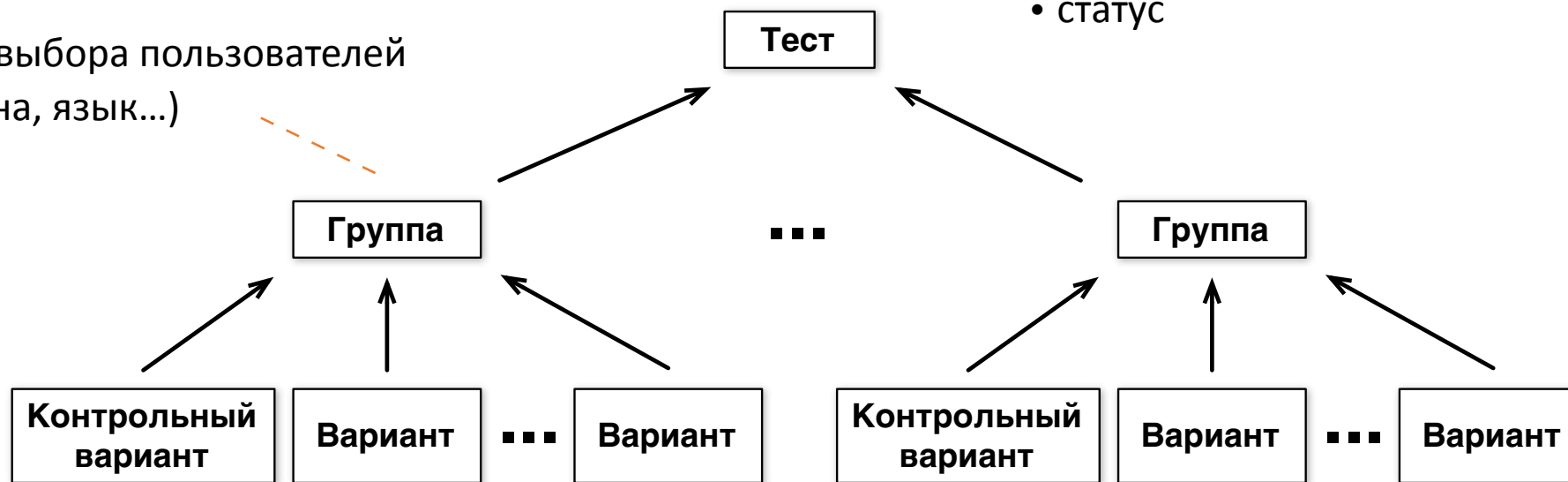


# БД тестов

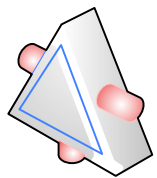


- id (сквозная нумерация)
- имя
- критерии выбора пользователей (пол, страна, язык...)

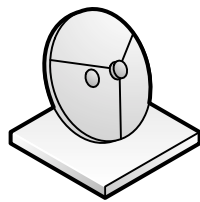
- id (сквозная нумерация)
- имя
- даты начала-конца
- статус



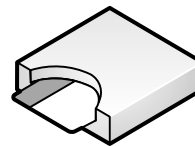
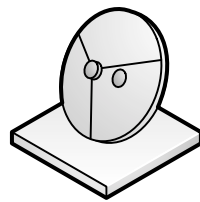
- id (сквозная нумерация)
- имя
- значение переменной/имя шаблона



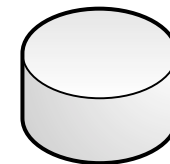
**API  
статистики**



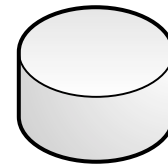
**Scribe-транспорт**



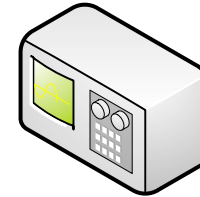
**Разборщик  
статистики**



**БД  
статистики**

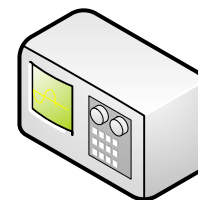


**БД тестов**



**Бекотис-  
интерфейс**

# Бек-офис-интерфейс



## Создание/редактирование тестов

- проверяет непересекаемость групп
- требует указать дату начала/конца

## Управление тестами

### Дебаг:

- покажи мне как выглядит вариант X?
- какой вариант был для юзера Z?
- попадет ли в тест юзер Y?



All

New

Running

Finished

▶ Demo test for Dave

Here we are trying to understand the interface



NEW (EMPTY VARIATION) 16.10.13

▶ Тестовый тест

Не запускать!!!



NEW (EMPTY VARIATION) 03.09.13

▶ Тест 1

Не запускать!!!



NEW (EMPTY VARIATION) 15.08.13

Here we are trying to understand the interface

NEW EXPERIMENT



10% Ireland



Ages: Any Users: 10% Gender: M & F

Languages: English (UK)

Countries: Ireland

10% users in Ireland

Variations: 1 Translations: 0 Status: Ready



All

New

Running

Finished

▶ Demo test for Dave

Here we are trying to understand the interface



NEW (EMPTY VARIATION) 16.10.13

▶ Тестовый тест

Не запускать!!!



NEW (EMPTY VARIATION) 03.09.13

▶ Тест 1

Не запускать!!!



NEW (EMPTY VARIATION) 15.08.13

Name:

Demo test for Dave

Description:

Here we are trying to understand the interface

Dates [Test duration]

17.10.2013

to

23.10.2013

CANCEL

SAVE



All

New

Running

Finished

▶ Demo test for Dave

Here we are trying to understand the interface



NEW (EMPTY VARIATION) 16.10.13

▶ Тестовый тест

Не запускать!!!



NEW (EMPTY VARIATION) 03.09.13

▶ Тест 1

Не запускать!!!



NEW (EMPTY VARIATION) 15.08.13

EDIT EXPERIMENT

Name:

Russian, English; Russia, UK; 5%

Description:

Languages & Countries:

Languages

Russian



English (UK)



Countries

Russia



United Kingdom



Ages

 to 

Gender

M

F

Users %

10

CANCEL

NEXT



All

New

Running

Finished

▶ Demo test for Dave

Here we are trying to understand the interface



NEW (EMPTY VARIATION) 16.10.13

▶ Тестовый тест

Не запускать!!!



NEW (EMPTY VARIATION) 03.09.13

▶ Тест 1

Не запускать!!!



NEW (EMPTY VARIATION) 15.08.13

NEW EXPERIMENT

**Variation 1**   
behavior 1

**Variation 2**   
behavior 2

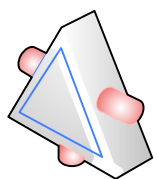
**Add Variation**  
+

Value:

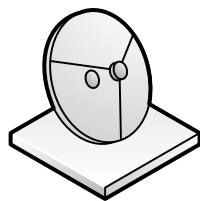
behavior\_1

CANCEL

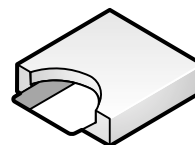
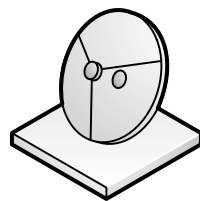
SAVE



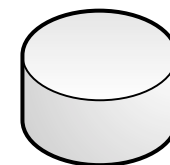
**API  
статистики**



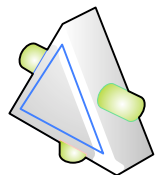
**Scribe-транспорт**



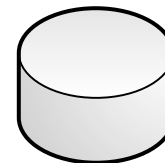
**Разборщик  
статистики**



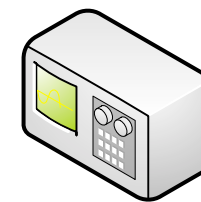
**БД  
статистики**



**API  
тестов**

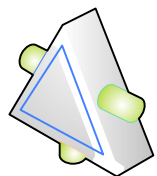


**БД тестов**

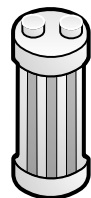


**Бекофис-  
интерфейс**

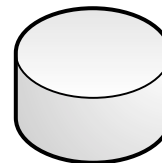




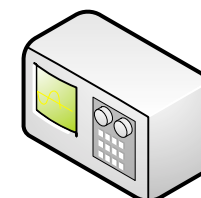
**API  
тестов**



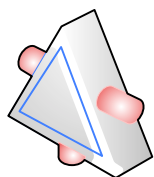
**Memcached**



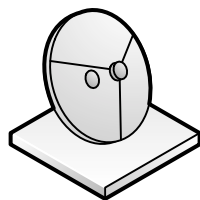
**БД тестов**



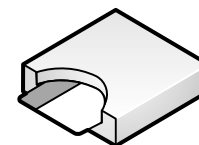
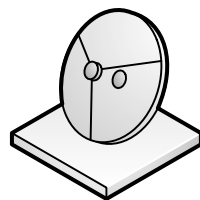
**Бекофис-  
интерфейс**



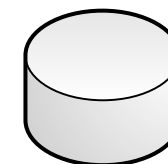
**API  
статистики**



**Scribe-транспорт**

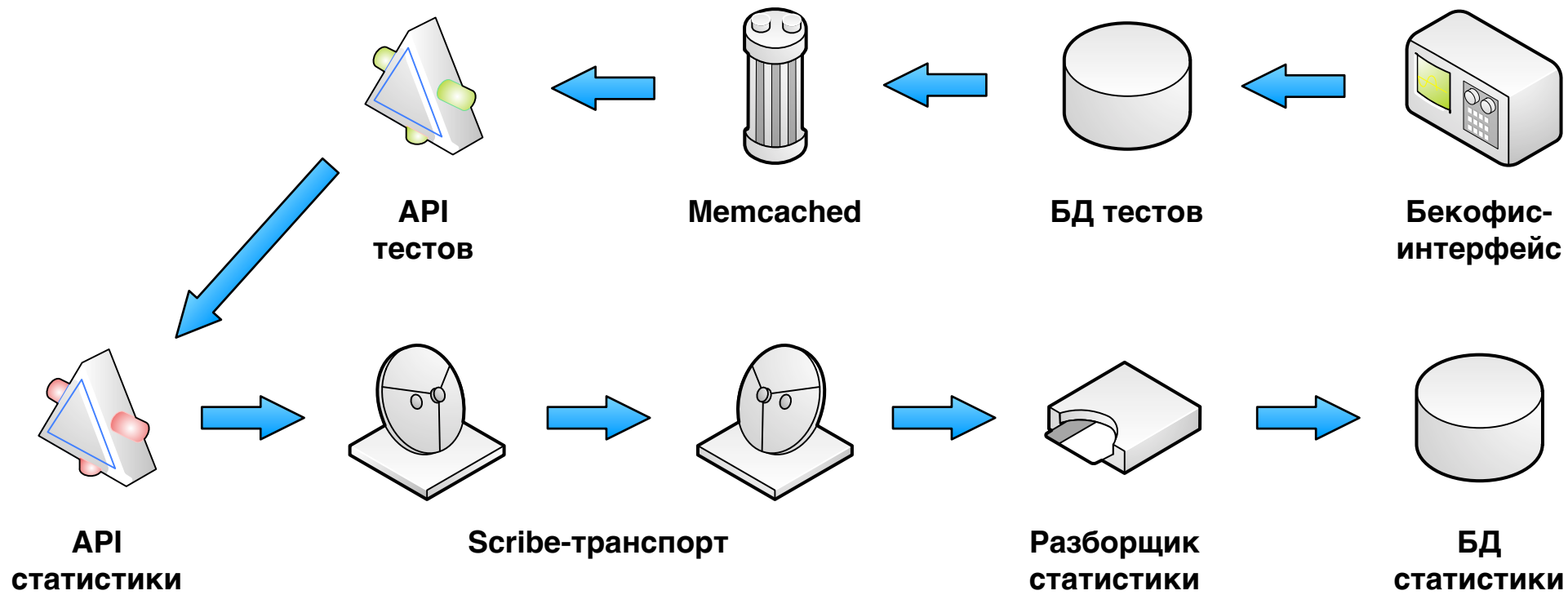


**Разборщик  
статистики**

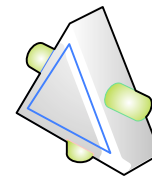


**БД  
статистики**





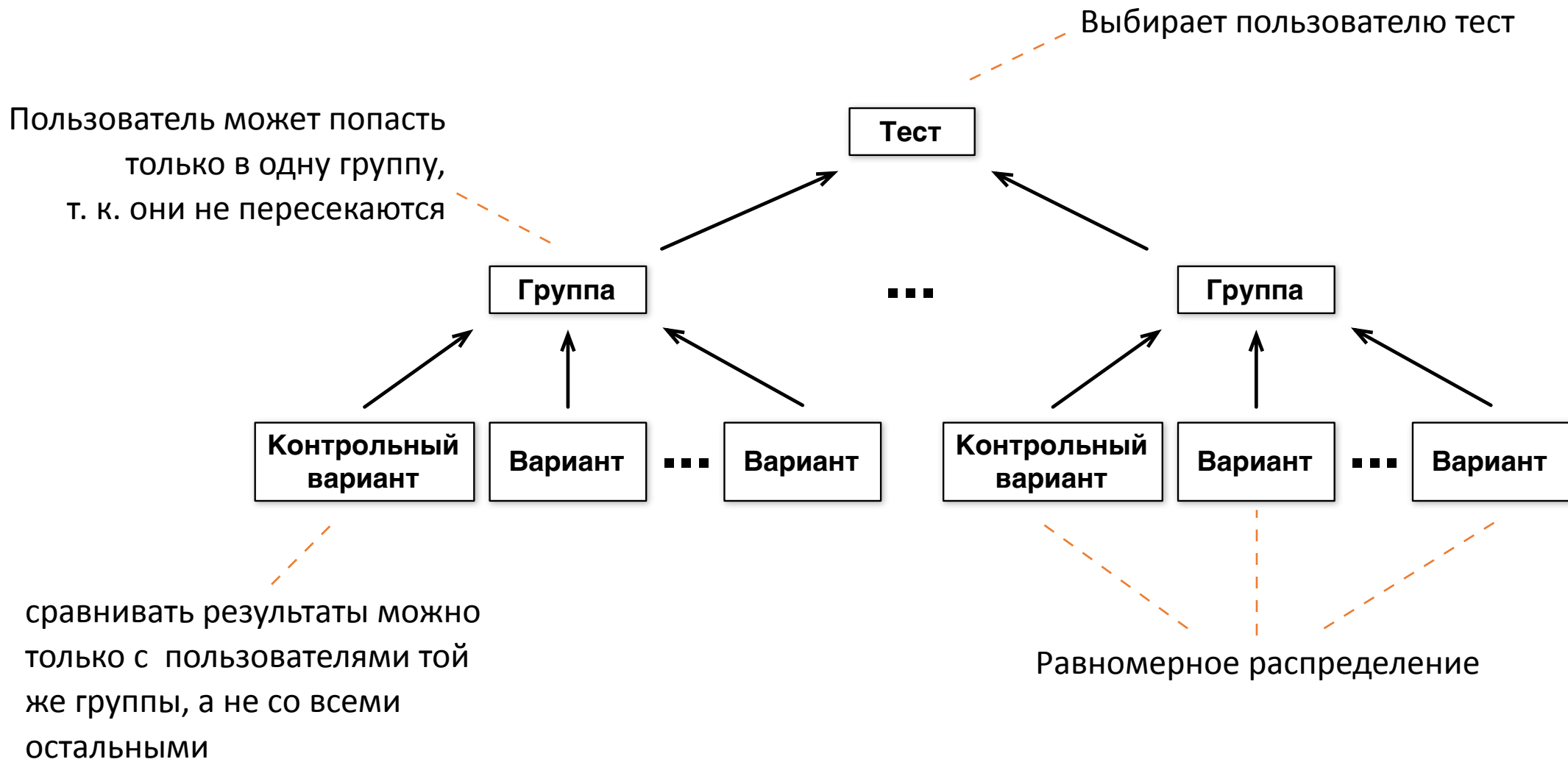
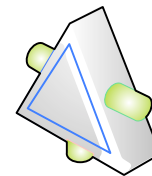
# API тестов



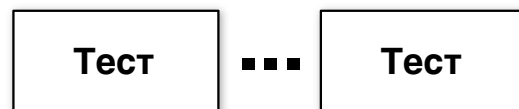
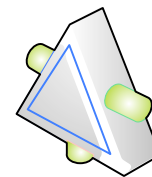
```
ABTests::hit($User, $page)
```

- Выбирает тест/группу/вариант для пользователя и страницы, где он находится (если есть подходящие)
- Отправляет эти данные в статистику как факт просмотра определенного варианта
- Возвращает либо null (не в тесте), либо заданное в варианте значение — имя шаблона/тип поведения/значение переменной
- У нас доходит до 3 млн вызовов в день, где возвращается не null

# API тестов



# API тестов: выбираем тест



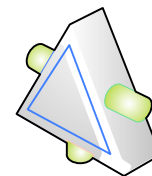
Тестировать надо так, чтобы не было взаимовлияния тестов

- $\$test\_id = \$user\_id \% \$num\_running\_tests$  — плохо

При изменении кол-ва запущенных тестов пользователи должны как можно меньше перемещаться между тестами

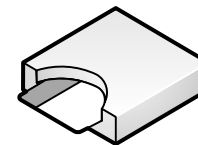
- consistent hashing
- заранее сделать несколько наборов пользователей
- выбирать и помечать тех, кто еще ни в каком тесте

# API тестов: выбираем вариант



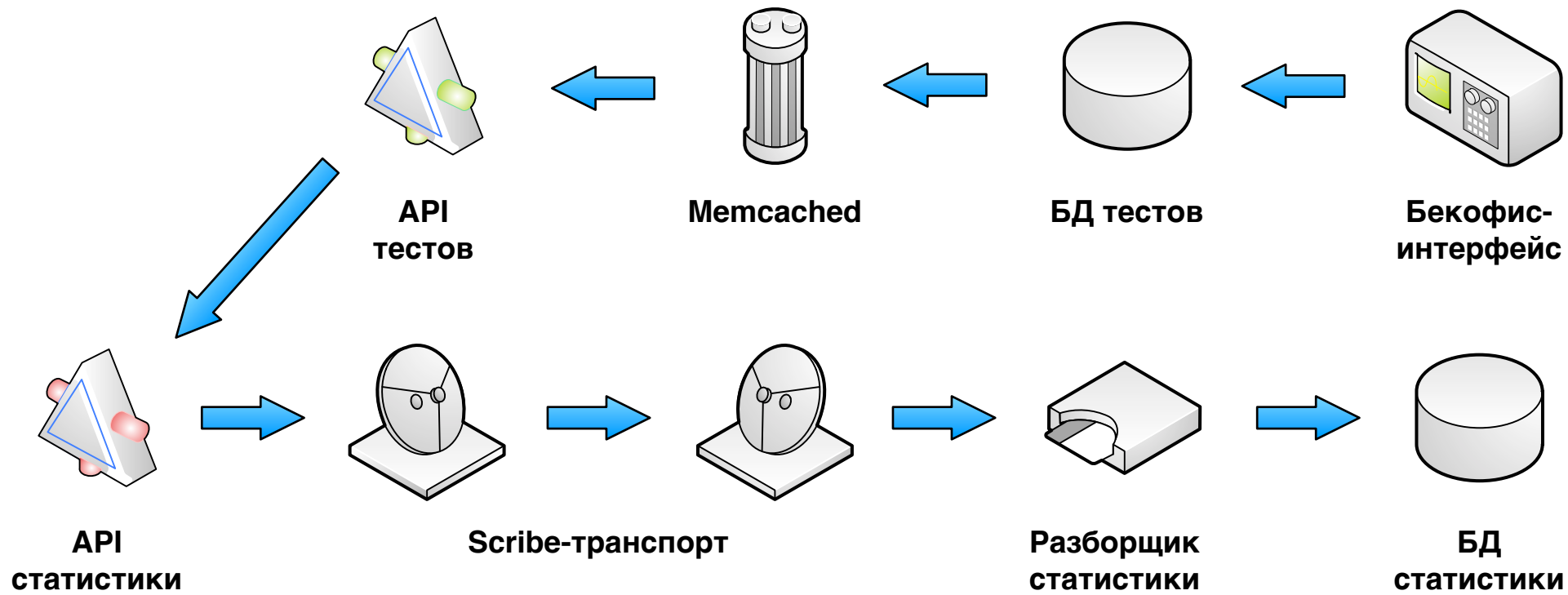
- по остатку от деления  $\$user\_id \% \$num\_variants$  — не хорошо
- по  $crc32(\$user\_id) \% \$num\_variants$  — вполне хорошее

# Скрипт-процессор статистики

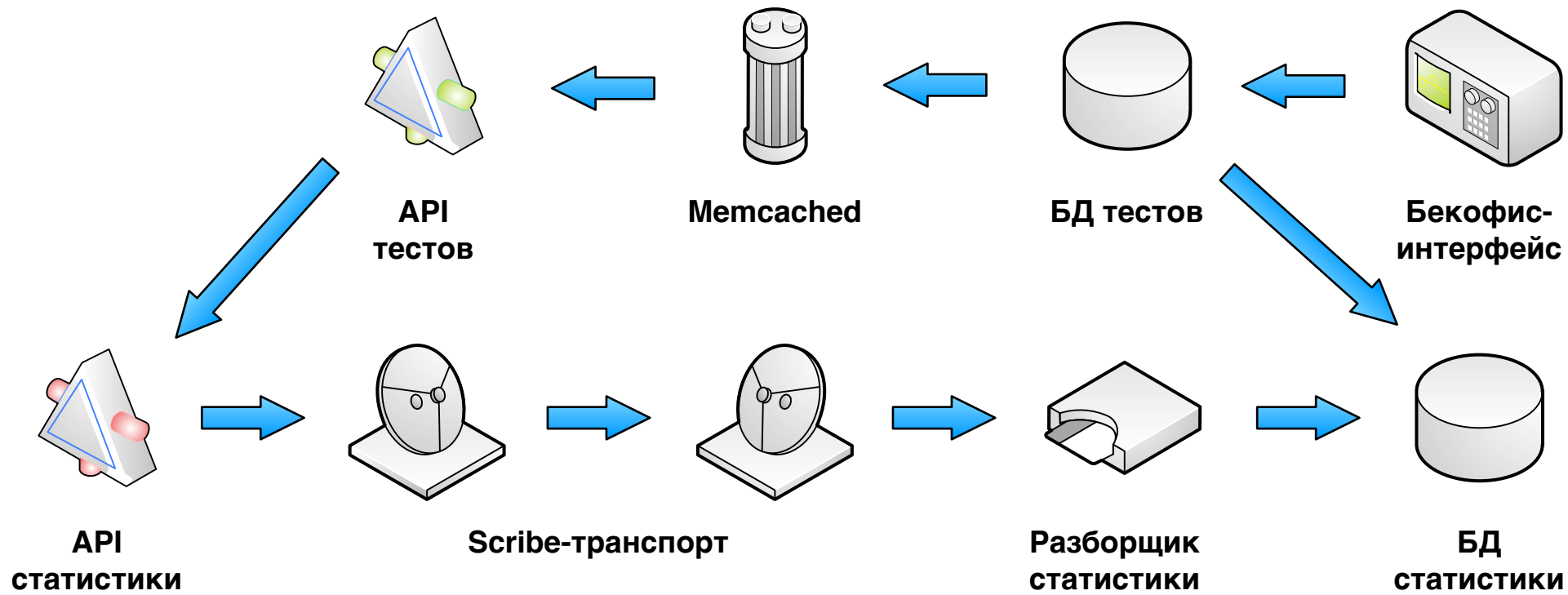


- Дорабатываем скрипт для создания таблицы фактов участия в тесте в БД статистики
- Можно сделать подсчет хитов

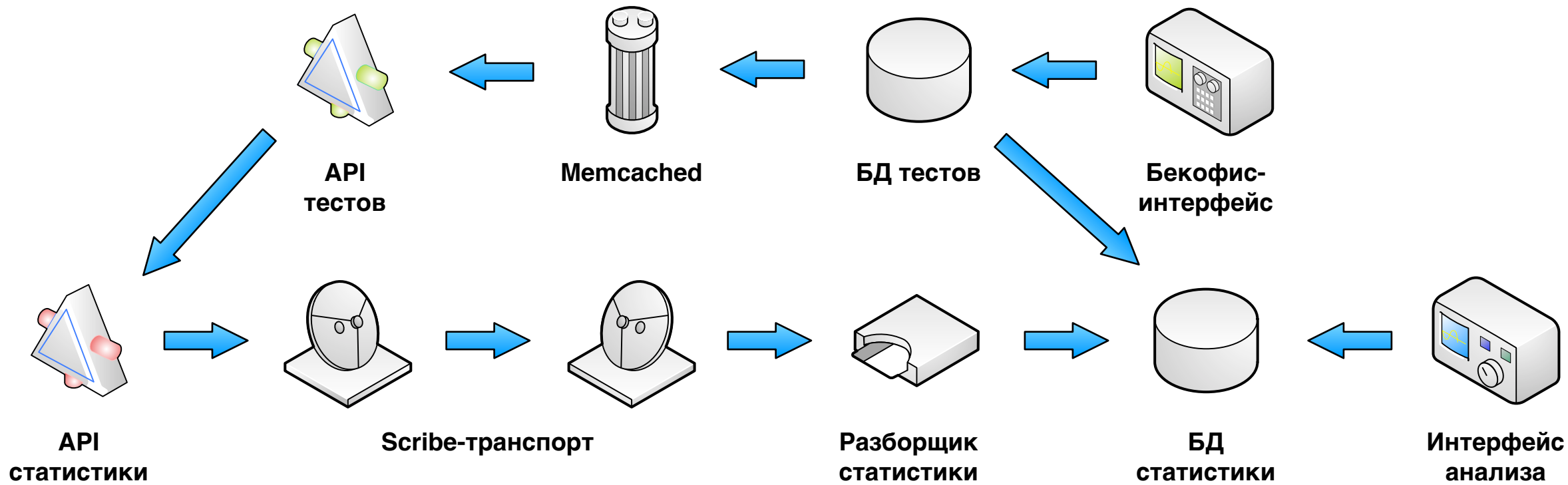
id	user_id	test_id	expt_id	var_id	hits	created	updated
1	336441661	1	2	3	381	2013-08-29 13:53:42	2013-08-30 12:17:48
185	336557489	1	2	3	377	2013-08-30 09:53:32	2013-09-06 08:47:18
198	314292808	2	3	4	48	2013-08-30 09:37:09	2013-09-09 13:27:17
225	334191012	2	3	4	28	2013-08-30 09:12:57	2013-09-04 07:02:20
242	333965942	2	3	4	5	2013-08-30 09:25:15	2013-08-30 10:35:15
258	329558096	2	3	4	2	2013-08-30 09:17:03	2013-08-31 18:43:08
263	141306373	2	3	4	1	2013-08-30 10:00:32	2013-08-30 10:00:32
309	330803344	2	3	4	2	2013-08-30 09:10:15	2013-08-30 09:10:15
336	332630150	2	3	4	10	2013-08-30 08:52:52	2013-09-04 20:49:25
357	331395121	2	3	4	4	2013-08-30 19:22:26	2013-09-09 12:04:48
361	334576284	2	3	4	12	2013-08-30 18:43:34	2013-09-05 15:37:58



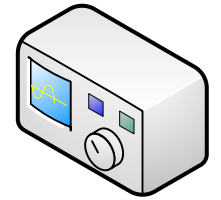




Делаете скрипт копирования тестов в базу со статистикой



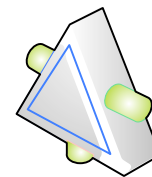
# Интерфейс для аналитика



- Программа-клиент для MySQL
  - JOIN статистики с таблицей фактов WHERE test\_id = 2 GROUP BY variant\_id
- Скрипт-генератор отчетов
- Бекофис-страница с графиками
- Система бизнес-аналитики
- ...

Как это работает

# Как это работает



- Распихиваете по коду сбор статистики
- В коде обработки запроса:
  - дергаете функцию выбора варианта
  - null и контрольный вариант — показываете то же, что и обычно
  - другое — показываете новый шаблон/вариант поведения
- Настраиваете тест в интерфейсе, прогоняем тест
- В БД получаем таблицы со статистикой, фактами просмотра пользователями вариантов и таблицы с описаниями тестов
- Анализируем результаты

# Оценка результатов

- Условия теста не должны меняться от начала до конца
- Проверяйте достоверность результатов (G-test confidence)
- Переводы часто портят результаты — делайте группы по языкам
- Проверяйте гомогенность групп по 2—3 параметрам
- Иногда для разных аудиторий побеждают разные варианты — не стесняйтесь слать мальчикам одни письма, а девочкам другие (только без фанатизма)
- Идея смотреть только кликрейт — так себе, выбирайте несколько параметров

**Много интересного:**

<http://habrahabr.ru/company/badoo/blog/>

# Спасибо!

# Вопросы?



@ryba\_xek



ryba.xek



s@averin.ru

Слайды, mind map,  
предыдущие доклады:  
<http://averin.ru/slides/>