



Программный
проект

Разработка интерфейса сравнения и анализа производительности ревизий ClickHouse

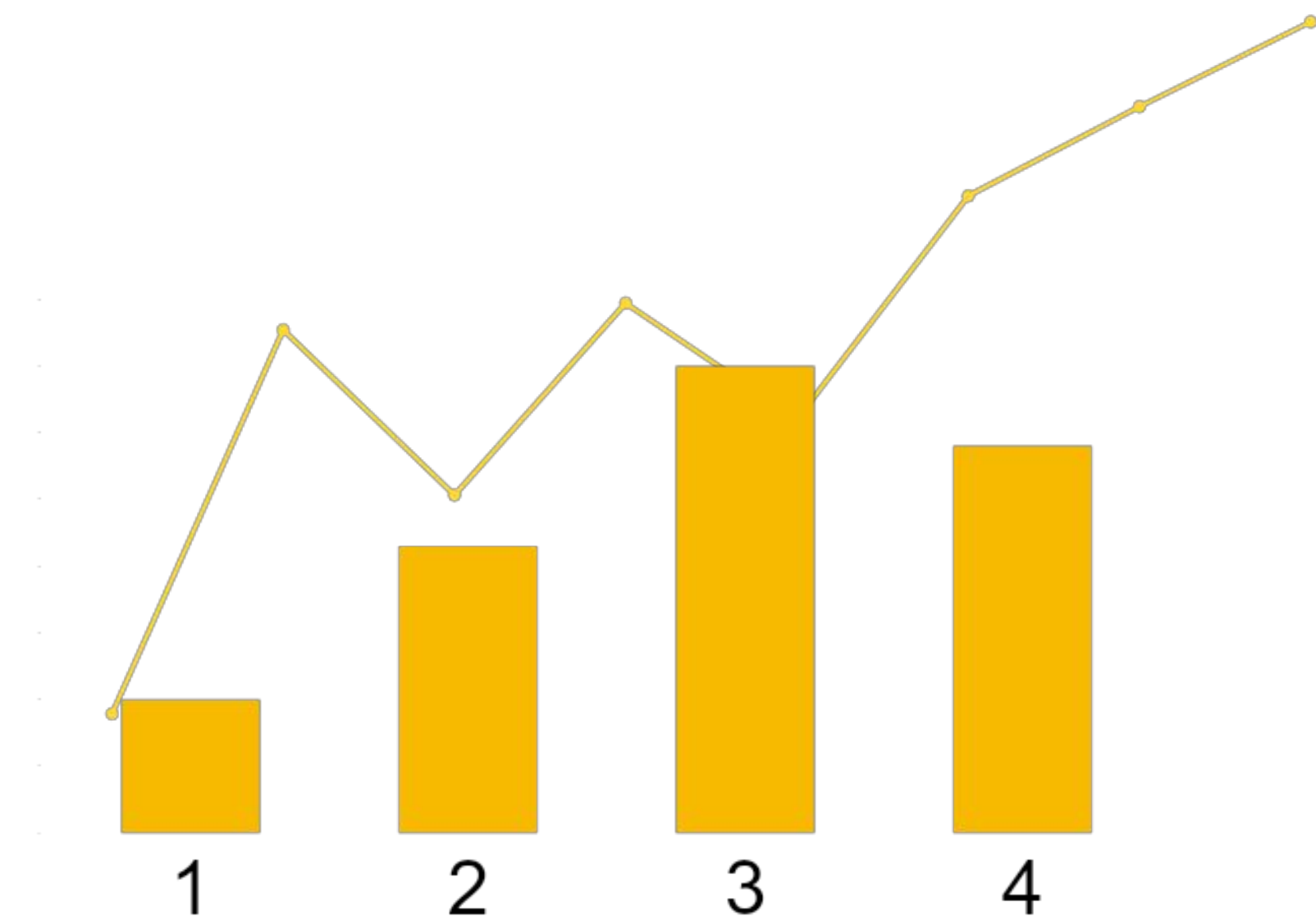
Development of the ClickHouse Revision Performance Comparison and Analysis Interface

Выполнила студентка
4 курса группы БПМИ164
Обрезкова
Дарья Валерьевна

Руководитель ВКР:
Миловидов
Алексей
Николаевич

Визуализация тестовых данных в сети Интернет

- Неизменяемые в зависимости от данных веб-компоненты
- При визуализации результатов тестирования редко используются графики и диаграммы
- Результаты тестирования представляются в виде passed/failed тестов



Ревизии

ClickHouse

- Слияние разных версий кода
- Тестирование
- Анализ результативности

новой

полученной версии

- Яндекс open-source проект
- СУБД для онлайн обработки аналитических запросов
- Визуализация результатов тестирования в виде статичного HTML без графиков и JavaScript кода

Использование

- Устранение ошибок
- Обнаружение уязвимостей
- Проверка структуры и стиля кода

Цели и задачи

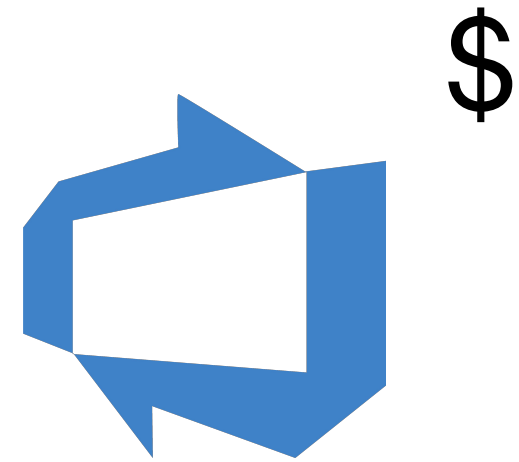
Цель Создать веб-интерфейс для быстрого понимания результатов тестирования и отслеживания качества работы над кодом

Задачи

Обработать данные о результатах тестирования ревизий ClickHouse

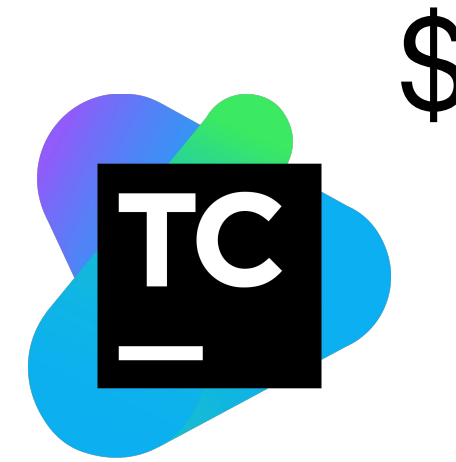
- Разработать набор визуальных элементов (графиков, таблиц и т.п.) для удобного и информативного отображения полученных данных
- Разработать статическое веб-приложение с интерфейсом сравнения и анализа производительности ревизий ClickHouse

Обзор существующих решений



\$

Azure DevOps

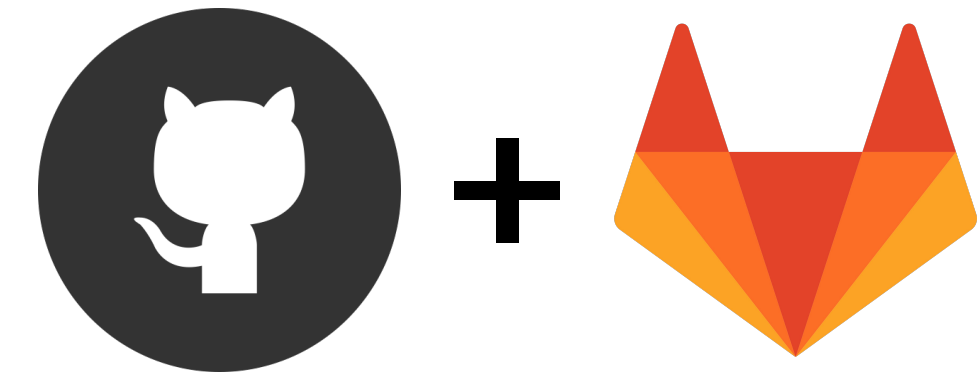


\$

TeamCity



Jenkins



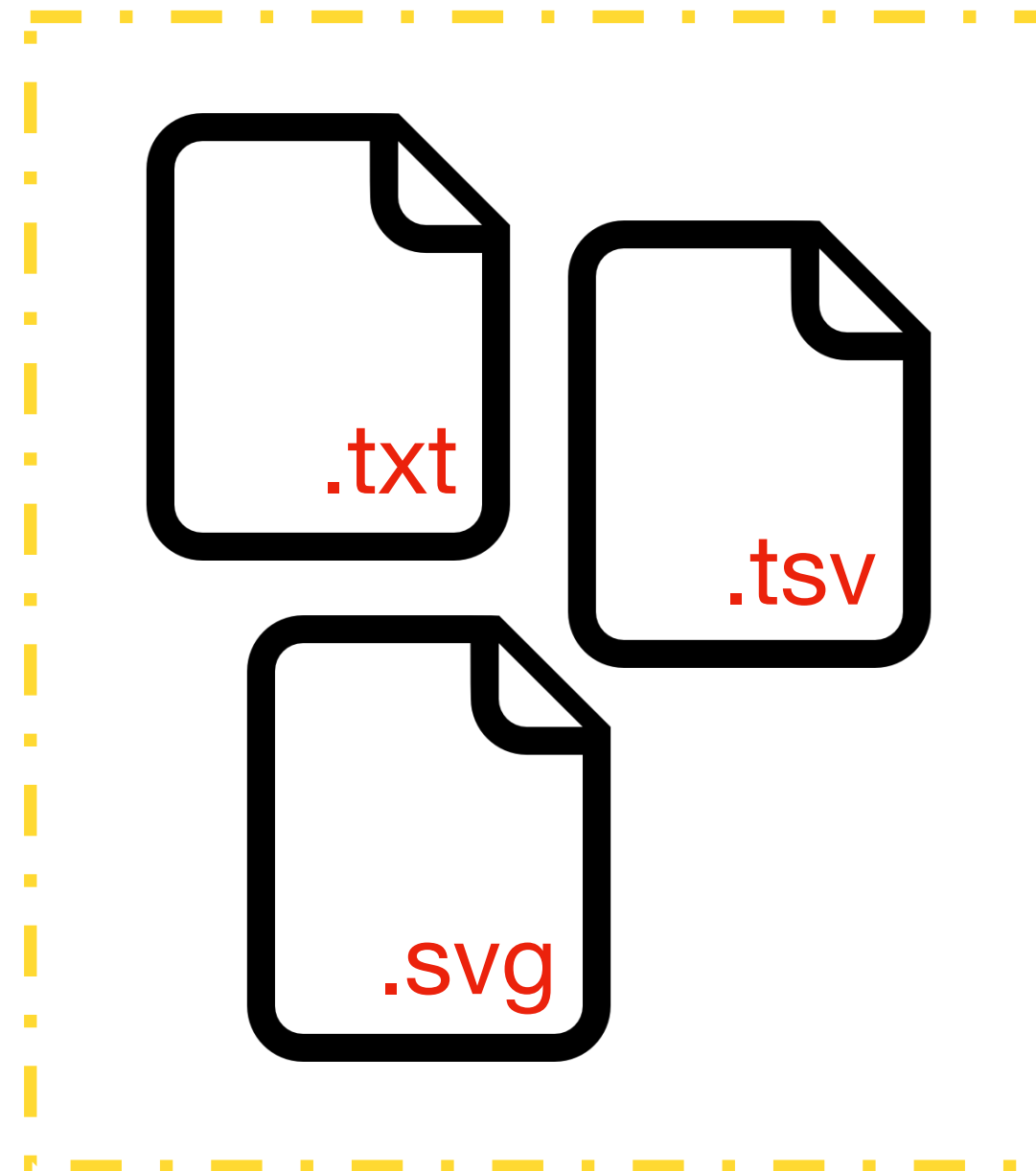
GitHub or GitLab

Критерии отбора

- Бесплатная программа
- Содержит необходимый функционал
- Содержит графики и таблицы с результатами тестирования

Работа с файлами

Работа с
файлами



make_json_to_send.py

Отбор необходимых

файлов

из общих результатов

тестирования



Считывание результатов

из

файлов и их обработка

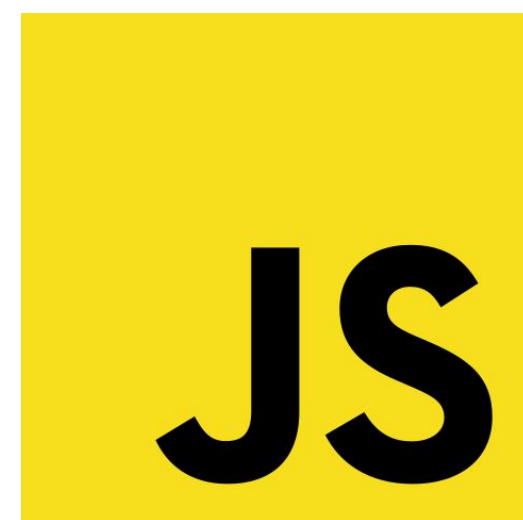


Сохранение данных в формате

.json

в доступную frontend-коду папку

Стек технологий для интерфейса



JavaScript



Vue.js

React.js



Nuxt.js



Highcharts.js

Chart.js

D3.js

+



CSS3



HTML5

Статический сайт с Nuxt.js

```
clickhouse-testing-interface /
-- | assets /
-- | components /
-- | dist /
----- | _nuxt /
----- | test_data /
----- | 200.html
----- | index.html
-- | layouts /
-- | node_modules /
-- | pages /
-- | plugins /
-- | static /
----- | test_data /
----- | commit /
----- | json-to-send.json
----- | commit_1 /
----- | ...
-- | store /
-- | nuxt.config.js
-- | README.md
```



- папка, всеми статическими
файлами (кодом js, картинками,
шрифтами и входными данными
json)

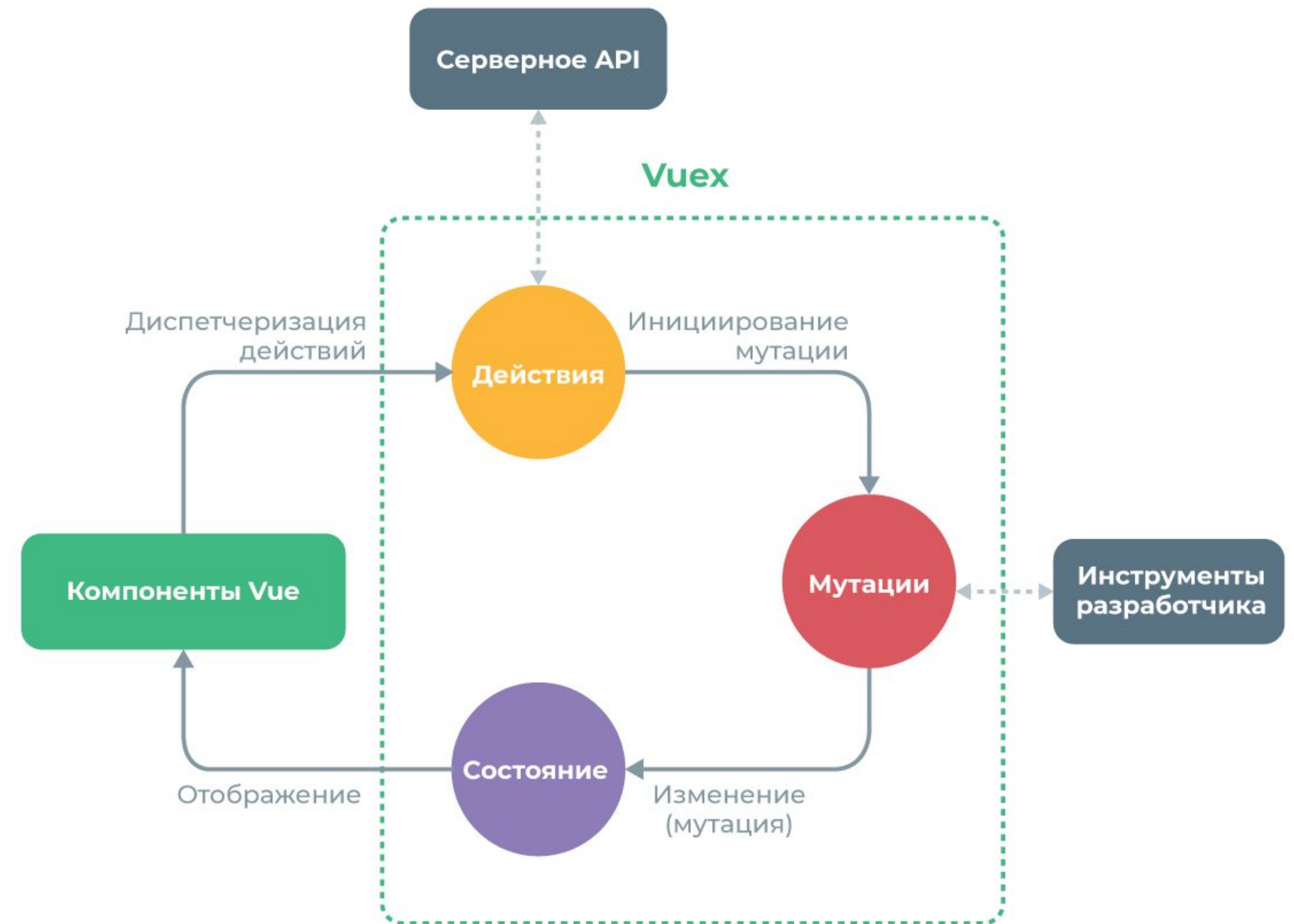


- файл с плагинами
(Highcharts.js)



- хранилище
store

Хранилище store



* Рисунок из официальной документации Vue.js

Интерфейс - графики и диаграммы



Performance
comparison of analytical DBMS



Dashboard



Tables



All Queries



Logs



Download

RESULTS OF THE COMMIT № b8be58559598ec31450d14c8cf526ee79dc571ac

TOTAL TESTS:
1248

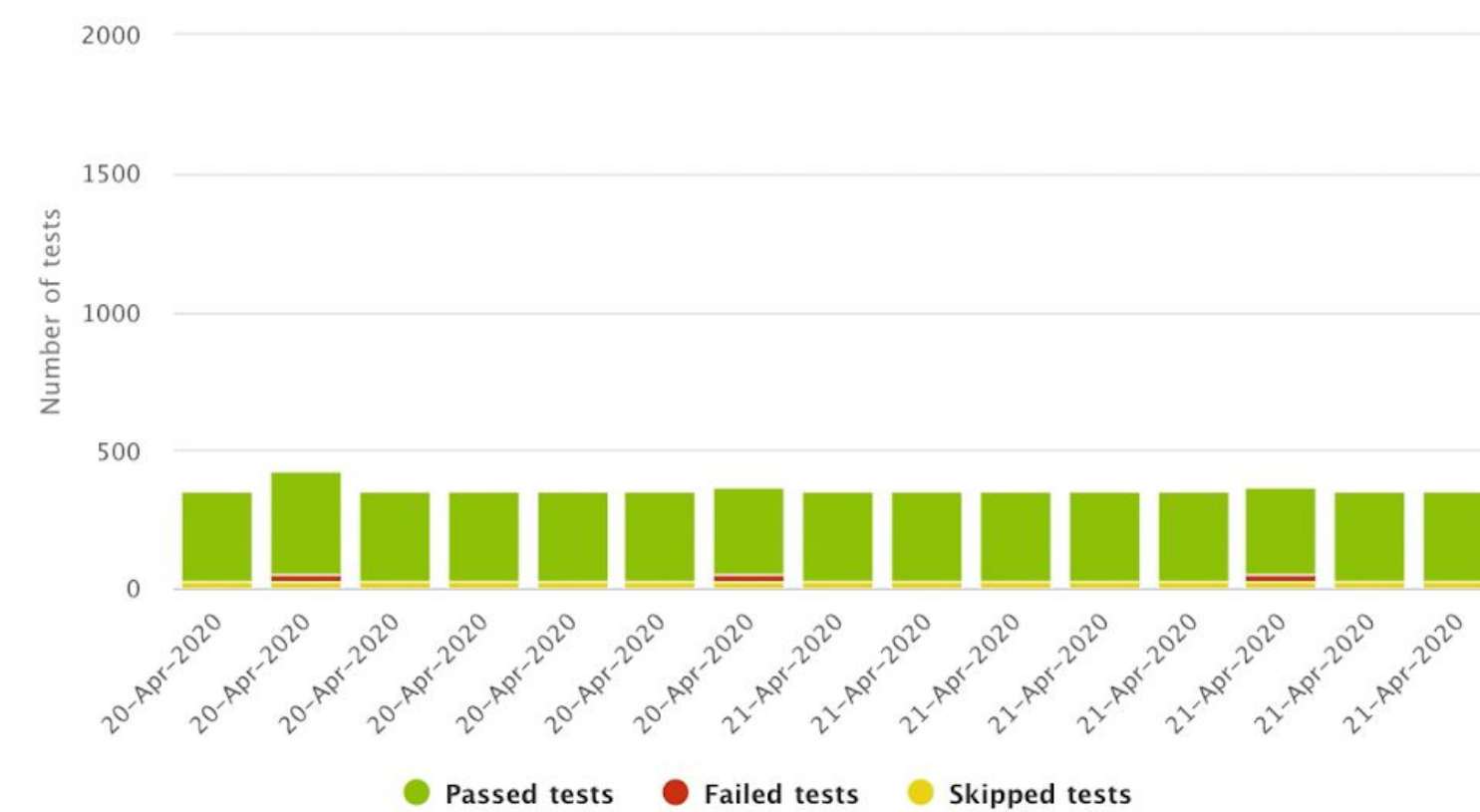
Failed tests Skipped tests Passed tests

1226
TESTS
PASSED

2
TESTS
FAILED

20
TESTS
SKIPPED

HISTORY OF ALL COMMITS



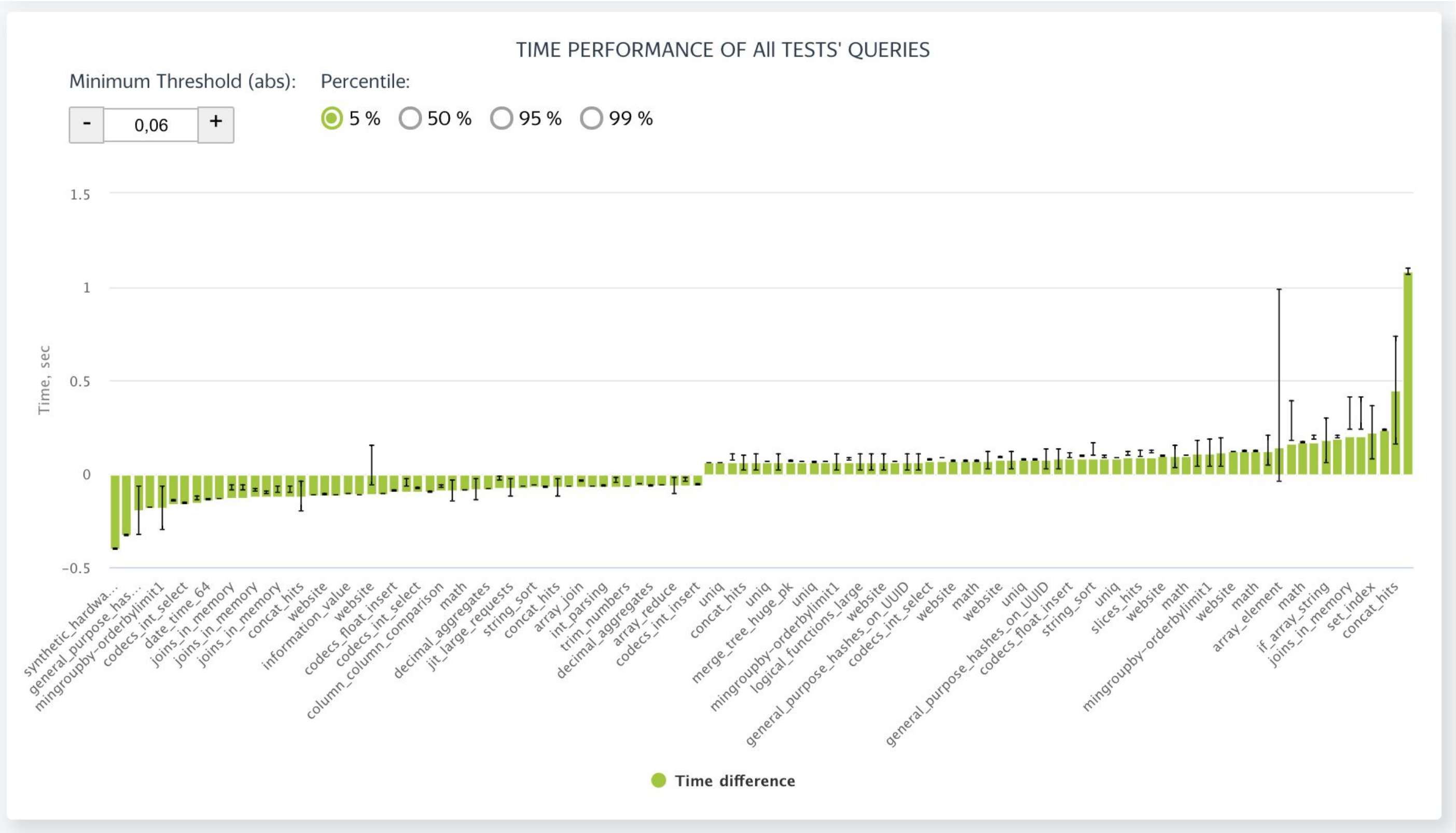
TIME PERFORMANCE OF ALL TESTS

Minimum Threshold (abs): Percentile: ☒ 5 % ☐ 50 % ☐ 95 % ☐ 99 %



Интерфейс - графики и диаграммы

Относительная разница между новым и старым временем выполнения каждого запроса



Значени
я

- Каждый запрос выполняется несколько раз
- Берется среднее геометрическое относительной разности нового и стартового времени

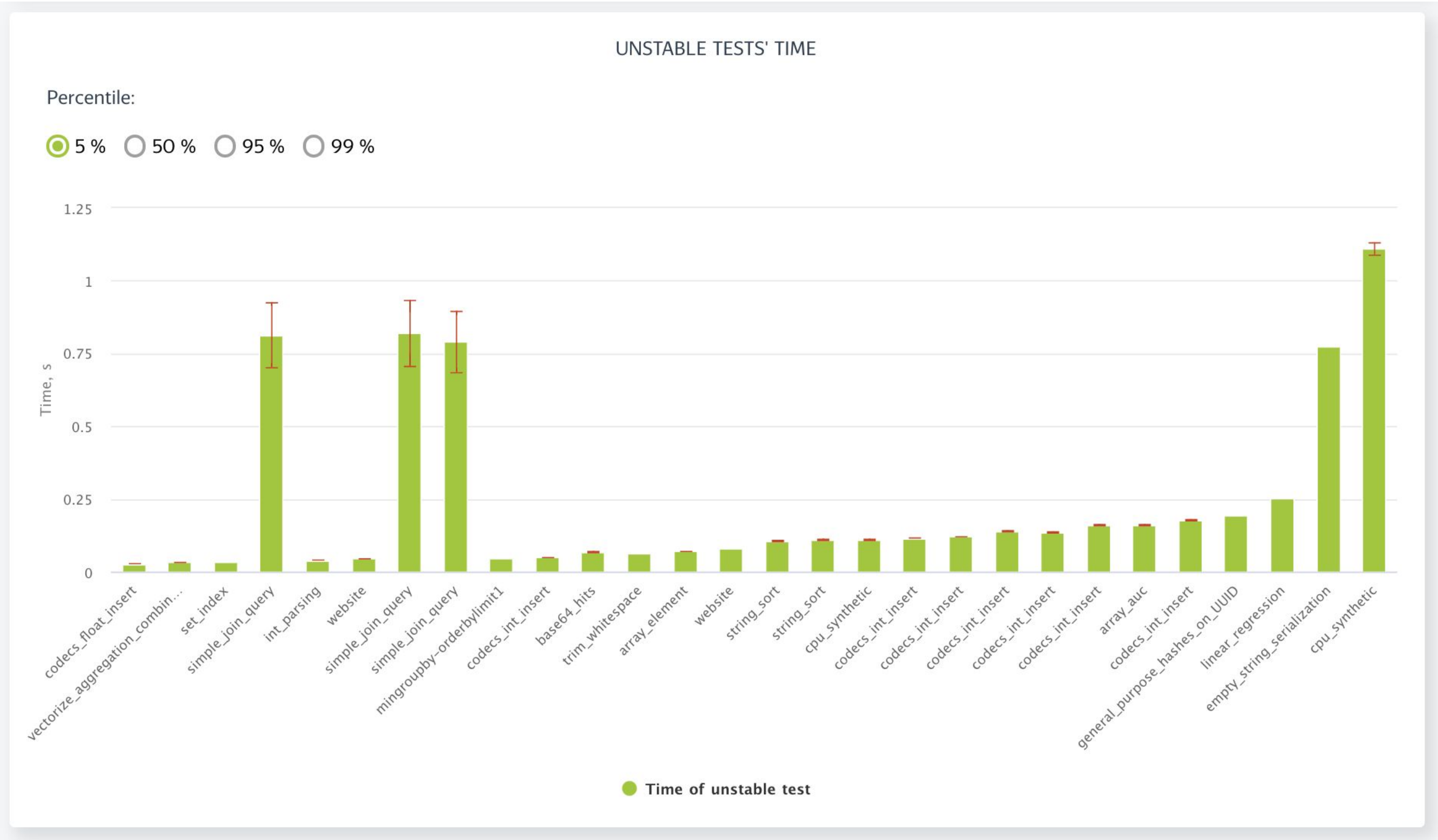
Параметр
ы

- Нижний порог разницы времен
- Доверительный интервал для каждого запроса

$$\bar{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

Интерфейс - графики и диаграммы

Относительная разница между новым и старым временем выполнения нестабильных запросов



Значени я

- Каждый запрос выполняется несколько раз
- Запрос нестабилен, если относительная разница нового и старого времени выше определенного порога

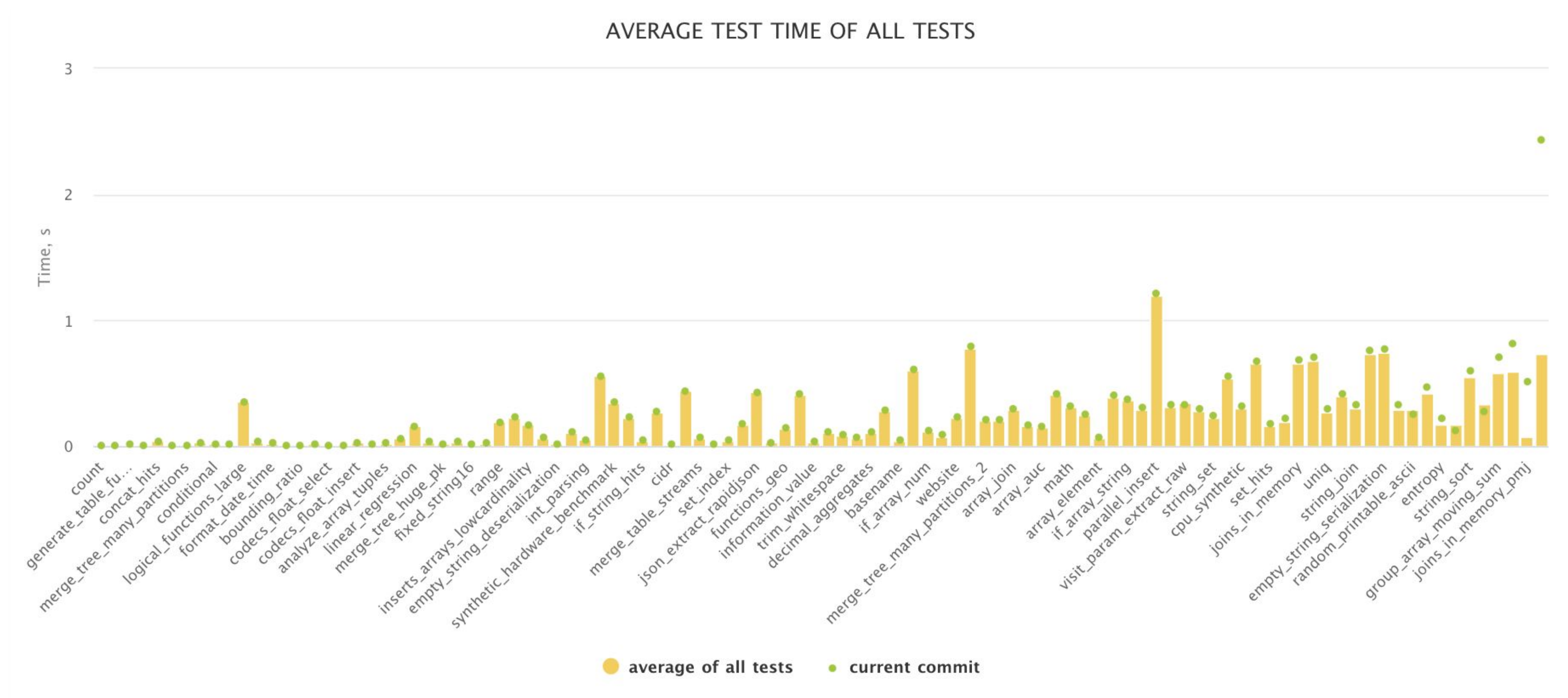
Параметр ы

- Доверительный интервал для каждого запроса

$$\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Интерфейс - графики и диаграммы

Среднее время прохождения каждого теста (набора запросов)
и среднее время предыдущих версий кода





Интерфейс - таблицы

Таблицы с результатами тестирования



Performance comparison of analytical DBMS

- Dashboard
- Tables
- All Queries
- Logs
- Download

Tested commits

	Old	New
Commit	29bb9f666565129587846f1507c9a4a5dad8a24e	b8be58559598ec31450d14c8cf526ee79dc571ac
Author	Alexander Kuzmenkov <akuzm@yandex-team.ru>	alexey-milovidov <milovidov@yandex-team.ru>
Date	Tue Apr 14 2020 00:15:58 GMT+0300 (Москва, стандартное время)	Tue Apr 14 2020 01:32:38 GMT+0300 (Москва, стандартное время)
Info	simple backport script	Merge pull request #10237 from ClickHouse/akz/mutations-to-correctly-handle-lambdas ALTER UPDATE/DELETE on Replicated* storages: Fixed "Unknown function lambda." error

Run errors ▼

Test	Error
./distributed_aggregation-err.log	clickhouse_driver.errors.ServerException: Code: 279.
./distributed_aggregation-err.log	DB::Exception: All connection tries failed. Log:

Unstable queries ▼

Old, s	New, s	Relative difference (new - old) %	Randomization distribution quantiles [5%, 50%, 95%, 99%]	Test	Query
--------	--------	-----------------------------------	--	------	-------

Интерфейс - таблицы

Таблица со всеми запросами

Performance comparison of analytical DBMS

Dashboard

Tables

All Queries

Logs

Download

All Queries ▾

Old, s	New, s	Relative difference (new - old)/old	Times speedup/ slowdown	Randomization distribution quantiles [5%, 50%, 95%, 99%]	Test	Query
0.0127	0.0104	-0.182	1.221	[0.011, 0.182, 0.244, 0.278]	mingroupby-orderbylimit1	\n SELECT key, min(value)\n FROM mingroupby_orderbylimit1_11111_key_value\n group by key format Null;\n
0.0259	0.0288	0.111	1.111	[0.095, 0.105, 0.138, 0.144]	mingroupby-orderbylimit1	\n SELECT key, min(value)\n FROM mingroupby_orderbylimit1_11111_tuple\n group by key format Null;\n
0.0209	0.0224	0.071	1.071	[0.065, 0.103, 0.249, 0.267]	mingroupby-orderbylimit1	\n SELECT key, min(value)\n FROM mingroupby_orderbylimit1_111111_key_value\n group by key format Null;\n
0.0449	0.0478	0.064	1.064	[0.064, 0.091, 0.349, 0.404]	mingroupby-orderbylimit1	\n SELECT key, min(value)\n FROM mingroupby_orderbylimit1_111111_tuple\n group by key format Null;\n
0.0849	0.0802	-0.056	1.058	[0, 0.025, 0.06, 0.069]	mingroupby-orderbylimit1	\n SELECT key, value\n FROM mingroupby_orderbylimit1_11111_key_value\n order by key, value limit 1 by key format Null;\n
0.4362	0.4327	-0.009	1.008	[0.008, 0.011, 0.018, 0.018]	mingroupby-orderbylimit1	\n SELECT key, value\n FROM mingroupby_orderbylimit1_11111_tuple\n order by key, value limit 1 by key format Null;\n
0.0919	0.1027	0.117	1.117	[0.099, 0.129, 0.173, 0.183]	mingroupby-orderbylimit1	\n SELECT key, value\n FROM mingroupby_orderbylimit1_111111_key_value\n order by key, value limit 1 by key format Null;\n

Интерфейс - логи и скачивание данных

Таблица с логами работы
кода



Performance
comparison of analytical DBMS



Dashboard



Tables



All Queries



Logs



Download

So many rows...
Please wait some seconds after clicking on column title for sorting

Logs ▼

Date	Time	Info
2020-04-14	04:09:48	user_files/test_all_expr_matches.values
2020-04-14	04:09:48	user_files/
2020-04-14	04:09:48	rm: cannot remove 'right': No such file or directory
2020-04-14	04:09:48	rm: cannot remove 'left': No such file or directory
2020-04-14	04:09:48	output/scripts/report.py
2020-04-14	04:09:48	output/scripts/report.py

Выгрузка архива с всеми
результатами
тестирования



Performance
comparison of analytical DBMS



Dashboard



Tables



All Queries



Logs

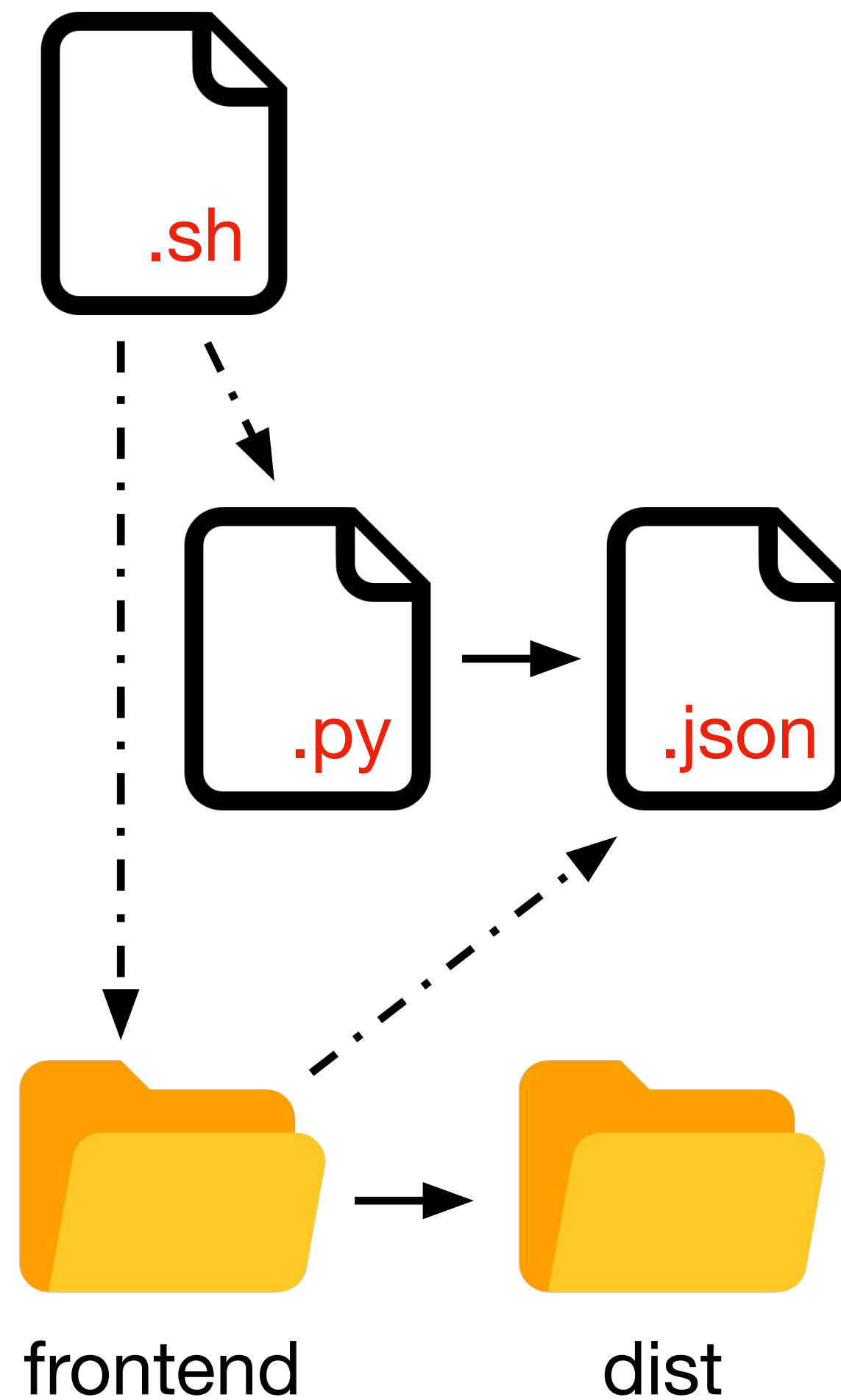


Download

TEST OUTPUT

DOWNLOAD

Сборка и запуск



Завершение

тестирования
и формирование файлов

с результатами
Запуск .sh файла,

который

запускает

файл Python

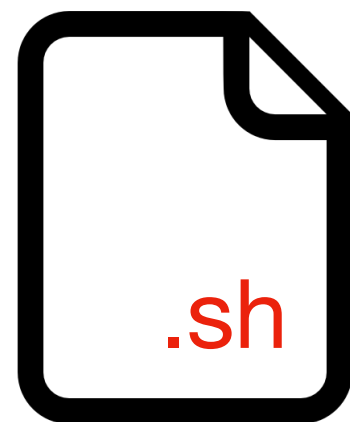
Обработка данных,
формирование файла

.json

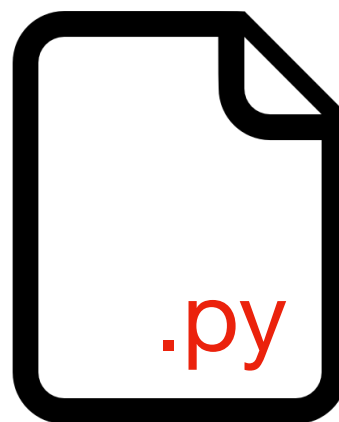
Запуск сборки статичной
страницы и формирование
папки dist

3 заключение

Итоговые
файлы



Файл
сборки



Файл обработки
данных



Файлы
интерфейса

Дальнейшие направления работы

- Добавление новых графиков
- Обновление кода при добавлении новых файлов о результатах тестирования или изменения структуры текущих



Спасибо за внимание!