

Метрики приложения



Артем Лысенко
31.10.2024

\$ whoami

Артем Лысенко



- Kotlin backend в RiskTech платформе
- ТА в академии бэкенда
- 4 курс ФКН ВГУ

План



- ➡ Что такое метрики?
- ➡ Типы метрик
- ➡ Как метрики попадают в хранилище?
- ➡ Лучшие практики

Зачем наблюдать за системой

- Поиск узких мест в приложении
- Проведение нагрузочного тестирования
- Контроль деплоя
- Алерты



Как наблюдать за системой

Логи

- Объем растет линейно
- Сложный поиск
- Подходит для разбора каждого запроса отдельно

Трейсы

- Показывает жизненный путь запроса
- Показывает критический путь запроса
- Подходит для разбора каждого запроса отдельно

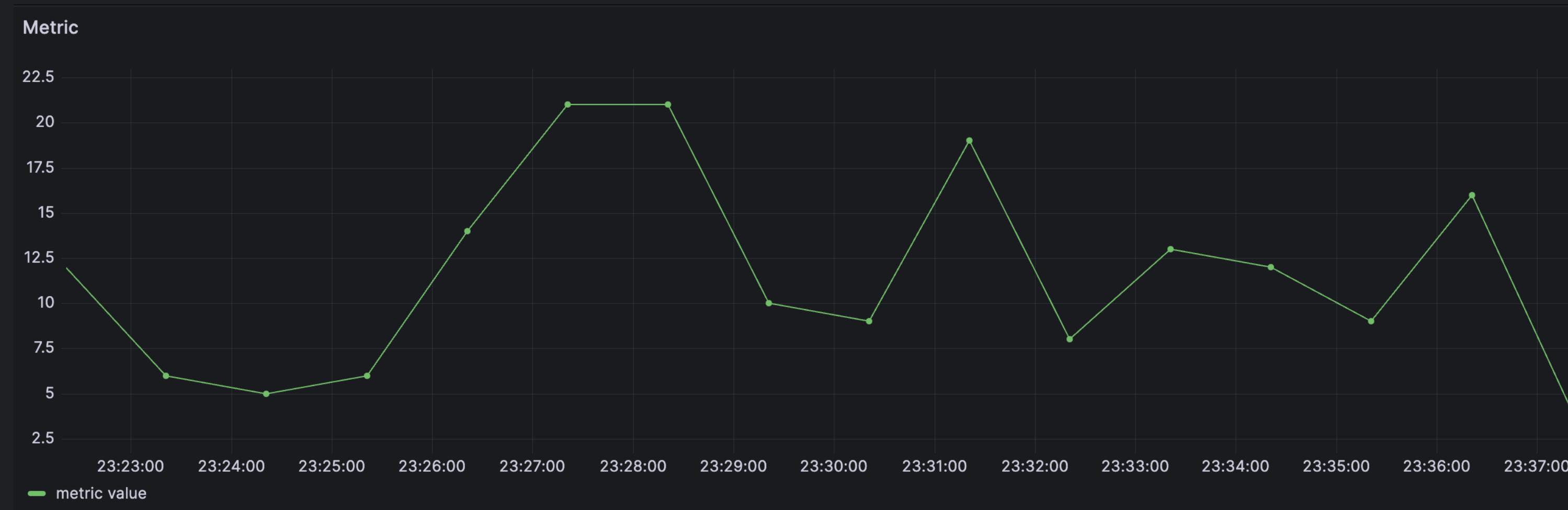
Метрики

- Агрегированная информация о системе
- в каждый момент времени
- Зачастую неточная

Что такое метрика

Метрика — показатель, которым система характеризуется в каждый момент времени

Временной ряд — дискретное представление метрики во времени



$$[(t_1, m_1), (t_2, m_2), (t_3, m_3), \dots, (t_n, m_n), \dots]$$

ЧТО ТАКОЕ МЕТРИКА

```
# HELP logback_events_total Number of log events that were enabled by the effective log level
# TYPE logback_events_total counter
```

```
logback_events_total{application="my-app",level="trace"} 0.0
logback_events_total{application="my-app",level="info"} 73.0
logback_events_total{application="my-app",level="debug"} 0.0
logback_events_total{application="my-app",level="error"} 2.0
logback_events_total{application="my-app",level="warn"} 2.0
```

- Имя метрики
- Набор тэгов
- Значение метрики

- `__name__` – имя метрики
- `@timestamp` – тэг времени

Кардинальность

Кардинальность — показатель, характеризующий уникальность значений в наборе данных

- userId
- username
- ip-адрес
- пользовательский ввод



Типы метрик

- Gauge
- Counter
- Histogram
- Summary



Gauge

Gauge — числовая метрика, которая может произвольно изменять свое значение

- `jvm_memory_used_bytes` – объем использованной памяти
- `system_cpu_usage` – потребление CPU
- `executor_queue_remaining_tasks` – число задач в очереди
- `disk_free_bytes` – место на диске



Counter

Counter — счетчик, монотонно возрастающее число

- jvm_classes_unloaded_classes_total – число выгруженных классов
- http_server_requests_seconds_count – общее число запросов
- logback_events_total – число логов



Histogram

Histogram — метрика, характеризующая распределение некоторой величины

Бакет	Интервал	Значение
le=0.2	[0, 200] ms	15
le=0.5	[0, 500] ms	15 + 25 = 40
le=1.0	[0, 1000] ms	15 + 25 + 10 = 50
le=+Inf	[0, +Inf] ms	15 + 25 + 10 + 2 = 52

- `http_server_requests_seconds` – распределение времени ответа
- Размер тела ответа внешнего сервиса



Summary

Summary — метрика, которая локально агрегирует гистограмму

Квантиль	Значение
quantile=0.5	150ms
quantile=0.75	450ms
quantile=0.95	1500ms
quantile=0.1	10000ms

- jvm_gc_pause_seconds – время GC пауз
- http_client_requests_seconds – распределение времени ответа клиентов



Push vs Pull

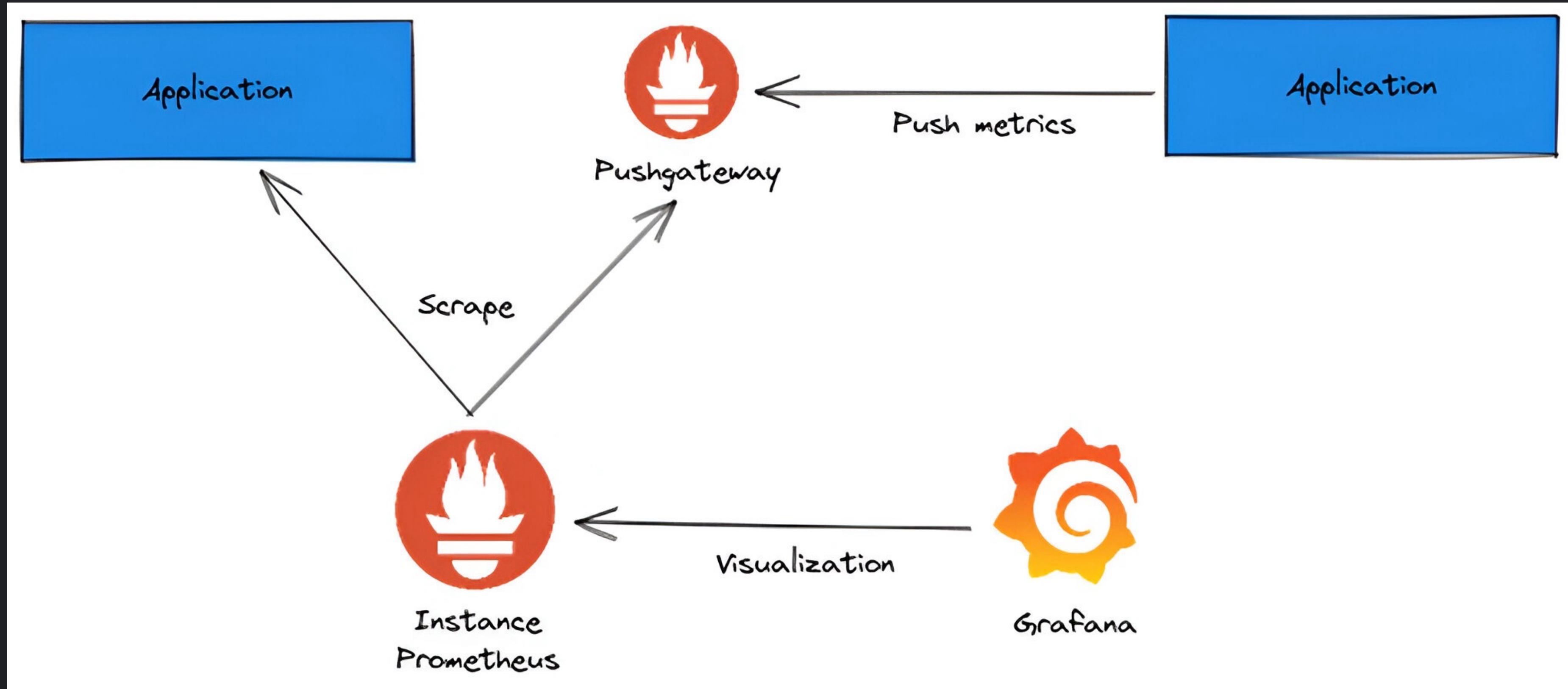
Push

- Клиентское приложение само отправляет метрики
- Необходимо если клиентское приложение имеет периодический характер или имеет большое число сэмплов
- Обманчиво просто

Pull

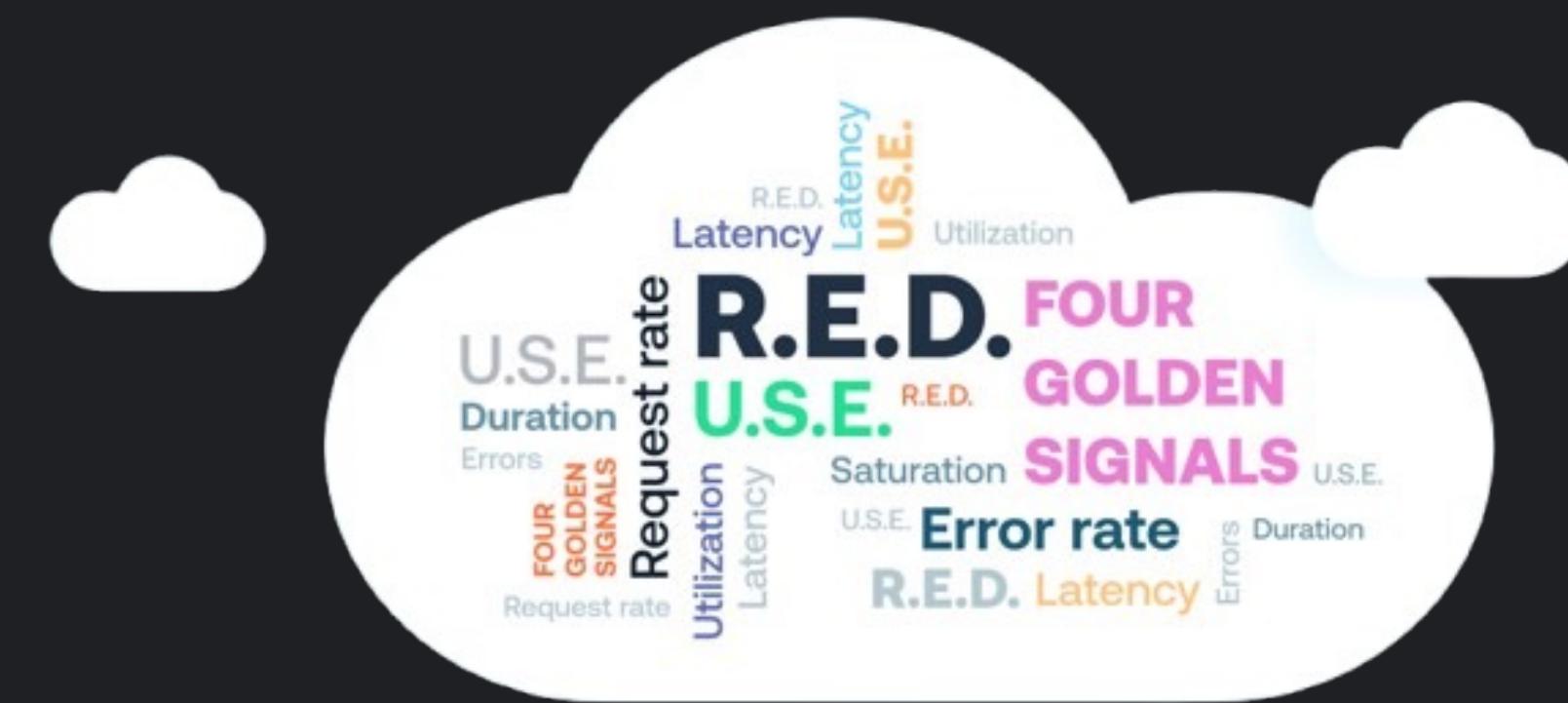
- Система мониторинга сама опрашивает клиентское приложение
- Если клиентское приложение развернуто в несколько инстансов, то система мониторинга должна сходить в каждый

Push vs Pull



Лучшие практики

- RED
- USE
- LETS (Google 4 Golden Signals)
- STELA (LETS + Availability)



RED

- Rate – число запросов в единицу времени
- Errors – число запросов с ошибкой в единицу времени
- Duration – время выполнения запросов

RED



USE

- Utilisation – насколько загружен или утилизирован ресурс
- Saturation – насколько насыщен ресурс
- Errors – ошибки

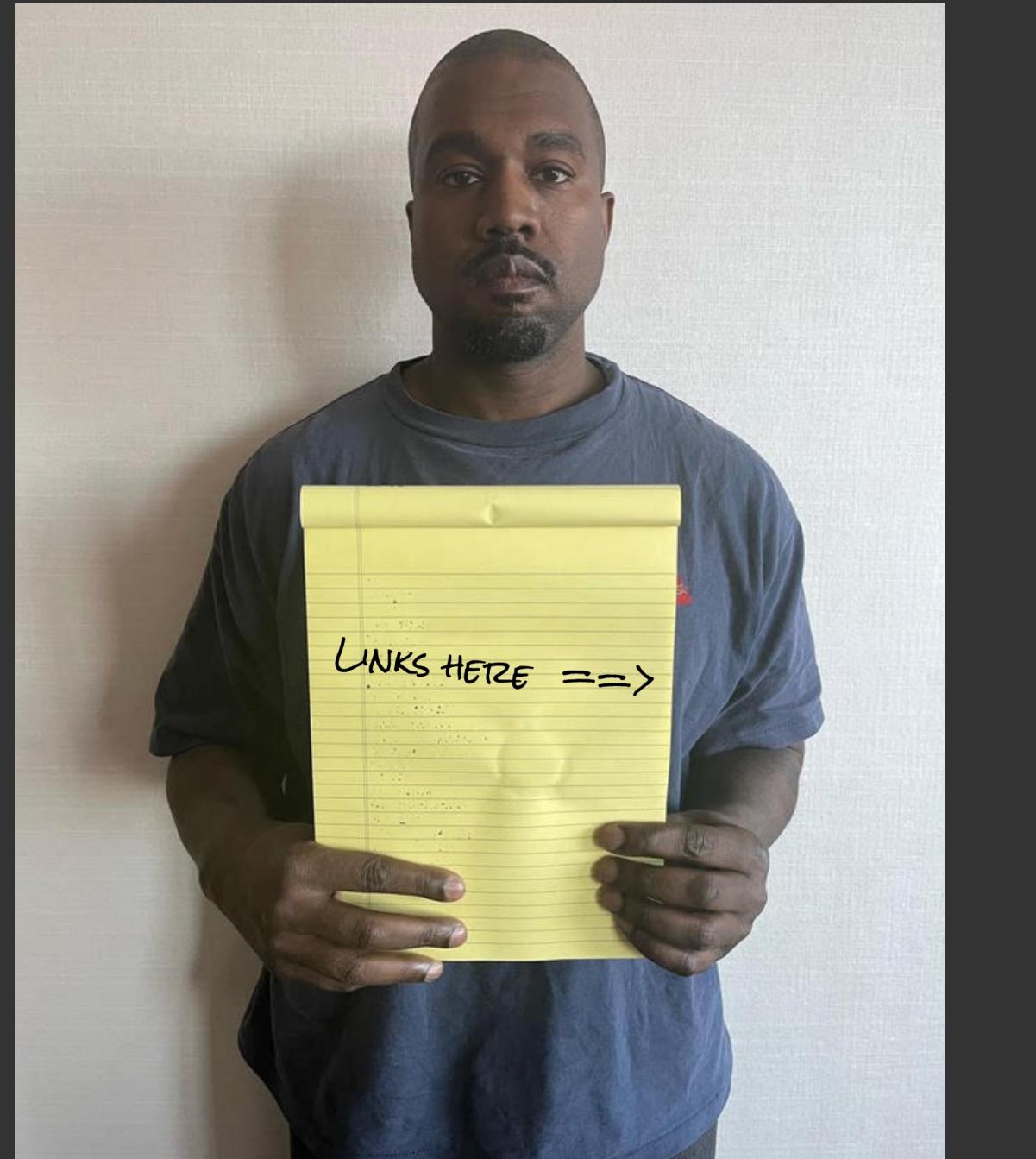
USE



LETS и STELA

- Latency – время выполнения запросов
- Errors – число запросов с ошибкой в единицу времени
- Traffic – число запросов в единицу времени
- Saturation – насыщение
- Availability – доступность

Ссылки



[Цикл статей по метрикам](#)



[Введение в метрики](#)



[Проект на Github'е](#)





Спасибо!

