Как мы собираем логи Nginx в ClickHouse

Глеб Гончаров
Системный администратор

О компании

Инфраструктура

- Геораспределённая
- 400+ физических серверов
- 200+ виртуальных серверов

Балансировщики нагрузки

- Nginx
- 40 krps HTTP/HTTPS
- 2 Gbps
- 800M запросов
- 10 Gb сжатых логов

Причины

Проблемы

- Нет консолидации
- Нет агрегации
- Нет индекса
- Сложные выражения обработки строк

Авария

- DoS
- 3 минуты простоя
- Человеческий фактор

Задачи

Задачи

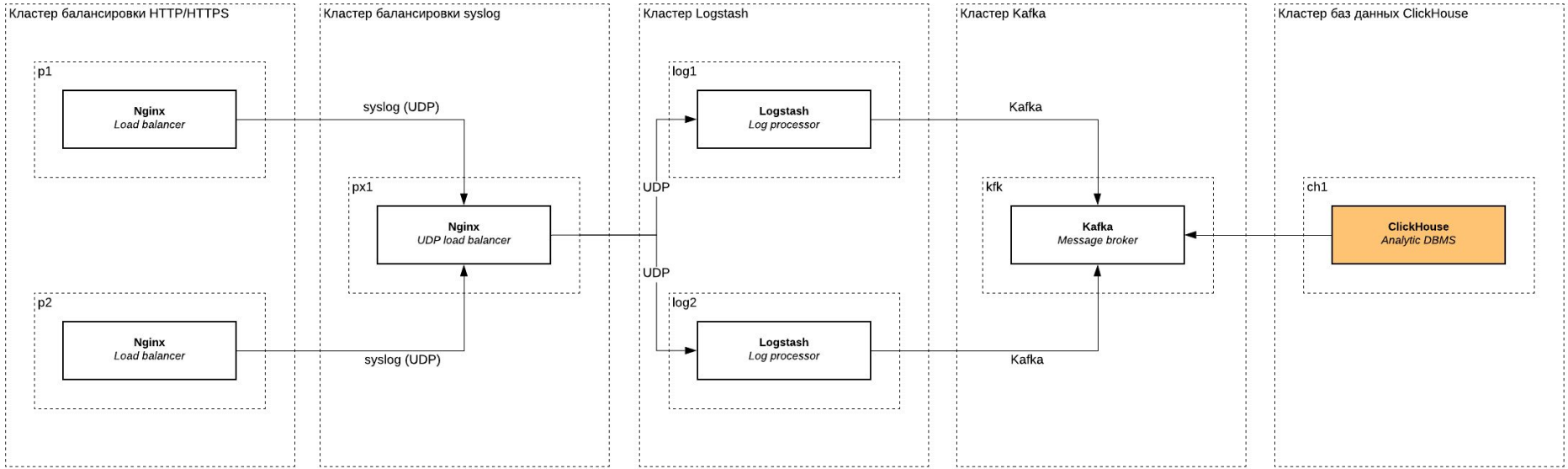
- Разработать систему доставки логов балансировщиков
- Обеспечить отказоустойчивость решения
- Унифицировать формат логирования
- Строить агрегаты метрик приложений
- Предоставить веб-интерфейс для поиска запросов

Решение

Состав кластера

● Балансировщики нагрузки	Nginx
● UDP-балансировщики	Nginx
● Обработчики логов	Logstash
● Брокеры сообщений	Kafka
● Аналитическая БД	ClickHouse

Архитектурная схема



Мониторинг

- Zabbix
- Prometheus
- Grafana

Трудности

Nginx и Syslog

Nginx не умеет пересылать в syslog по TCP.

Решения

- `socat(1)`
- Локальный rsyslog и RELP
- Отправлять по UDP

ClickHouse и retention policy

Партицирование по времени не задействует ClickHouse TTL.

Решения

- Удалять партиции внешним скриптом
- Использовать составной ключ партицирования

ClickHouse и DateTime

ClickHouse хранит данные с типом DateTime как UNIX timestamp

Решения

- Хранить миллисекунды отдельно
- Ничего не делать

ClickHouse и время в Nginx

ClickHouse не понимает формат времени из `$time_local` и `$time_iso8601`

Решения

- Приводить в Nginx с помощью Lua, Njs или Perl
- Приводить в Logstash

```
Nginx       2019-06-27T05:41:51+03:00
```

Nginx и время ответа от сервера

`$upstream_response_time` может быть пустой строкой

Решения

- `map` в Nginx
- Преобразовывать в Logstash
- Приводить тип в ClickHouse

Сообщения и HTTP-методы

HTTP-методов больше, чем кажется.

Решения

- Взять список из IANA HTTP Method Registry
- Хранить строкой

Сообщения и User-Agent

Некоторые клиенты отправляют escape-последовательности в UA

Решения

- Заменять escape-символы перед повторным парсингом JSON
- Ничего не делать

```
Mozilla/5.0 (Linux; Android 8.1; Xperia\xA0\XA2 Ultra Build/LMY47I)
```

Итоги

Плюсы

- Сократили время поиска до 1000 раз
- Организовали мониторинг приложений
- Отыскали скрытые ошибки при конфигурировании

Минусы

- Обрабатываем 99,99992% сообщений
- Нужно поддерживать

Вопросы

ФанБокс
funbox.ru
@funbox_team

Глеб Гончаров
gon.gl
@gongled