

Skillbox

Observability

Работа с логированием

Андрей Гордиенков

Solution Architect

ABAX

На прошлом уроке

- Проблема наблюдаемости приложения
- Составляющие характеристики наблюдаемости

На этом уроке

- Цель логов
- Использование простых и структурированных логов
- Подходы к сбору логов

Размеры данных

International Data Corporation (IDC) прогнозирует, что глобальная сфера данных вырастет более чем в пять раз за семь лет, с 33 зеттабайт в 2018 году до 175 зеттабайт в 2025 году.

(Зеттабайт - это 10^{21} байт или миллион петабайт).

Телеметрия от:

- Умных фабрик, IoT
- Научные исследования (автомобили, научные установки, наблюдения)
- Поведение пользователей

Логирование

Анализ логов начинается со:

- сбора
- объединения
- подготовки логов по всему предприятию

Индексирование, очистка и нормализация наборов данных играют свою роль, и все эти задачи должны выполняться с высокой скоростью и эффективностью

Цель логирования

Аномалии

A .. B .. C
A .. B .. C
A .. B .. Z

Связи

Если A, то
скорее всего
будет B

Тренды

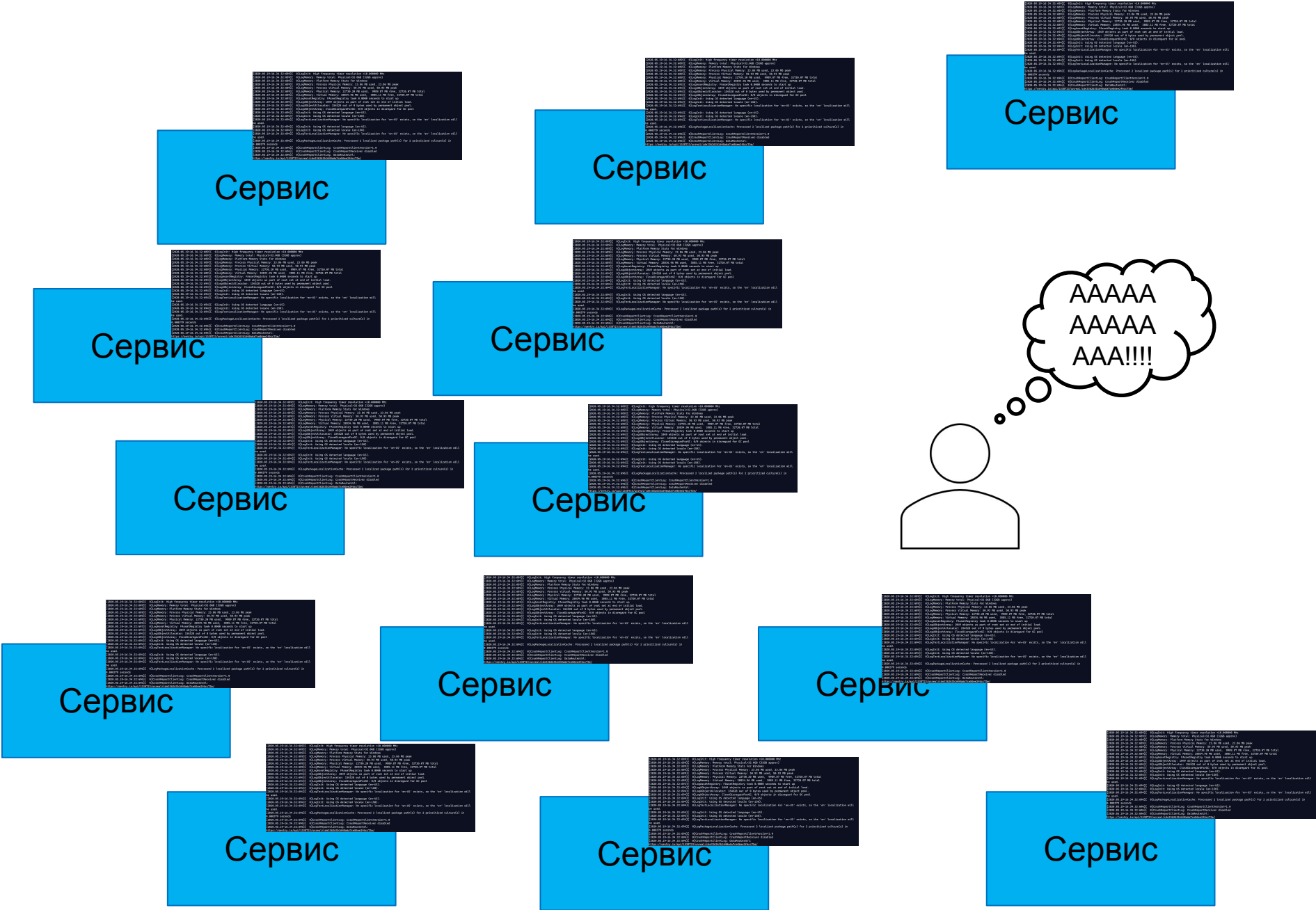
A, C, E, G, I

Единый источник

Монолит

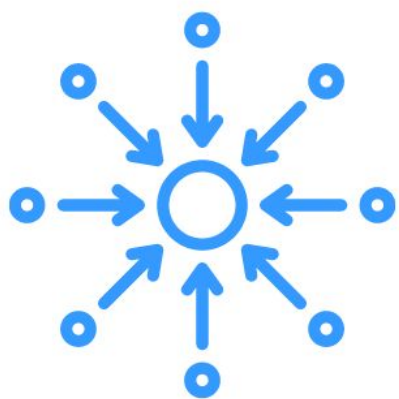
```
[2020.05.19-16.34.32:689][ 0]LogInit: High frequency timer resolution =10.000000 MHz
[2020.05.19-16.34.32:689][ 0]LogMemory: Memory total: Physical=32.0GB (32GB approx)
[2020.05.19-16.34.32:689][ 0]LogMemory: Platform Memory Stats for Windows
[2020.05.19-16.34.32:689][ 0]LogMemory: Process Physical Memory: 23.86 MB used, 23.86 MB peak
[2020.05.19-16.34.32:689][ 0]LogMemory: Process Virtual Memory: 50.93 MB used, 50.93 MB peak
[2020.05.19-16.34.32:689][ 0]LogMemory: Physical Memory: 22730.20 MB used, 9989.87 MB free, 32720.07 MB total
[2020.05.19-16.34.32:689][ 0]LogMemory: Virtual Memory: 28834.96 MB used, 3885.11 MB free, 32720.07 MB total
[2020.05.19-16.34.32:693][ 0]LogAssetRegistry: FAssetRegistry took 0.0000 seconds to start up
[2020.05.19-16.34.32:694][ 0]LogUObjectArray: 1049 objects as part of root set at end of initial load.
[2020.05.19-16.34.32:694][ 0]LogUObjectAllocator: 154528 out of 0 bytes used by permanent object pool.
[2020.05.19-16.34.32:694][ 0]LogUObjectArray: CloseDisregardForGC: 0/0 objects in disregard for GC pool
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected language (en-US).
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected locale (en-150).
[2020.05.19-16.34.32:694][ 0]LogTextLocalizationManager: No specific localization for 'en-US' exists, so the 'en' localization will
be used.
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected language (en-US).
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected locale (en-150).
[2020.05.19-16.34.32:694][ 0]LogTextLocalizationManager: No specific localization for 'en-US' exists, so the 'en' localization will
be used.
[2020.05.19-16.34.32:696][ 0]LogPackageLocalizationCache: Processed 2 localized package path(s) for 2 prioritized culture(s) in
0.000379 seconds
[2020.05.19-16.34.32:696][ 0]CrashReportClientLog: CrashReportClientVersion=1.0
[2020.05.19-16.34.32:696][ 0]CrashReportClientLog: CrashReportReceiver disabled
[2020.05.19-16.34.32:696][ 0]CrashReportClientLog: DataRouterUrl:
https://sentry.io/api/1538723/unreal/cde336263b1640a6a7ce86ee2f6cc75e/
```

Проблема



Логирование

Как получить пользу от распределенного логирования?



Централизованное
логирование



Возможность
поиска



Корреляция

Логирование

Простой склейкой данных не обойтись

```
[2020.05.19-16.34.32:689][ 0]LogInit: High frequency timer resolution =10.000000 MHz
[2020.05.19-16.34.32:689][ 0]LogMemory: Memory total: Physical=32.0GB (32GB approx)
[2020.05.19-16.34.32:689][ 0]LogMemory: Platform Memory Stats for Windows
[2020.05.19-16.34.32:689][ 0]LogMemory: Process Physical Memory: 23.86 MB used, 23.86 MB peak
[2020.05.19-16.34.32:689][ 0]LogMemory: Process Virtual Memory: 50.93 MB used, 50.93 MB peak
[2020.05.19-16.34.32:689][ 0]LogMemory: Physical Memory: 22730.20 MB used, 9989.87 MB free, 32720.07 MB total
[2020.05.19-16.34.32:689][ 0]LogMemory: Virtual Memory: 28834.96 MB used, 3885.11 MB free, 32720.07 MB total
[2020.05.19-16.34.32:693][ 0]LogAssetRegistry: FAssetRegistry took 0.0000 seconds to start up
[2020.05.19-16.34.32:694][ 0]LogUObjectArray: 1049 objects as part of root set at end of initial load.
[2020.05.19-16.34.32:694][ 0]LogUObjectAllocator: 154528 out of 0 bytes used by permanent object pool.
[2020.05.19-16.34.32:694][ 0]LogUObjectArray: CloseDisregardForGC: 0/0 objects in disregard for GC pool
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected language (en-US).
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected locale (en-150).
[2020.05.19-16.34.32:694][ 0]LogTextLocalizationManager: No specific localization for 'en-US' exists, so the 'en' localization will be used.
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected language (en-US).
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected locale (en-150).
[2020.05.19-16.34.32:694][ 0]LogTextLocalizationManager: No specific localization for 'en-US' exists, so the 'en' localization will be used.
[2020.05.19-16.34.32:696][ 0]LogPackageLocalizationCache: Processed 2 localized package path(s) for 2 prioritized culture(s) in 0.000379 seconds
[2020.05.19-16.34.32:696][ 0]CrashReportClientLog: CrashReportClientVersion=1.0
[2020.05.19-16.34.32:696][ 0]CrashReportClientLog: CrashReportReceiver disabled
[2020.05.19-16.34.32:696][ 0]CrashReportClientLog: DataRouterUrl:
https://sentry.io/api/1538723/unreal/cde336263b1640a6a7ce86ee2f6cc75e/
[2020.05.19-16.34.32:689][ 0]LogInit: High frequency timer resolution =10.000000 MHz
[2020.05.19-16.34.32:689][ 0]LogMemory: Memory total: Physical=32.0GB (32GB approx)
[2020.05.19-16.34.32:689][ 0]LogMemory: Platform Memory Stats for Windows
[2020.05.19-16.34.32:689][ 0]LogMemory: Process Physical Memory: 23.86 MB used, 23.86 MB peak
[2020.05.19-16.34.32:689][ 0]LogMemory: Process Virtual Memory: 50.93 MB used, 50.93 MB peak
[2020.05.19-16.34.32:689][ 0]LogMemory: Physical Memory: 22730.20 MB used, 9989.87 MB free, 32720.07 MB total
[2020.05.19-16.34.32:689][ 0]LogMemory: Virtual Memory: 28834.96 MB used, 3885.11 MB free, 32720.07 MB total
[2020.05.19-16.34.32:693][ 0]LogAssetRegistry: FAssetRegistry took 0.0000 seconds to start up
[2020.05.19-16.34.32:694][ 0]LogUObjectArray: 1049 objects as part of root set at end of initial load.
[2020.05.19-16.34.32:694][ 0]LogUObjectAllocator: 154528 out of 0 bytes used by permanent object pool.
[2020.05.19-16.34.32:694][ 0]LogUObjectArray: CloseDisregardForGC: 0/0 objects in disregard for GC pool
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected language (en-US).
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected locale (en-150).
[2020.05.19-16.34.32:694][ 0]LogTextLocalizationManager: No specific localization for 'en-US' exists, so the 'en' localization will be used.
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected language (en-US).
[2020.05.19-16.34.32:694][ 0]LogInit: Using OS detected locale (en-150).
[2020.05.19-16.34.32:694][ 0]LogTextLocalizationManager: No specific localization for 'en-US' exists, so the 'en' localization will be used.
[2020.05.19-16.34.32:696][ 0]LogPackageLocalizationCache: Processed 2 localized package path(s) for 2 prioritized culture(s) in 0.000379 seconds
[2020.05.19-16.34.32:696][ 0]CrashReportClientLog: CrashReportClientVersion=1.0
[2020.05.19-16.34.32:696][ 0]CrashReportClientLog: CrashReportReceiver disabled
[2020.05.19-16.34.32:696][ 0]CrashReportClientLog: DataRouterUrl:
https://sentry.io/api/1538723/unreal/cde336263b1640a6a7ce86ee2f6cc75e/
```

Проблемы

- Произвольное составление форматов логов.
- Простой текстовый файл
- Составление парсинга
- Изменение формата нарушает системы обработки логов

Абсолютное зло:

- Запись логирования в несколько строк

Структурные логи

2021-06-12T16:13:22+00:00 ERROR Read timed out

Когда произошло?

timestamp 2021-06-12T16:13:22+00:00

Сообщение?

level ERROR log Read timed out

Что за компонент?

service Service A team events

Что исполняется?

commit 543af4 build a.34+122 runtime Java-1.8.0

Где?

region Europe-west2 node node_e839ae12

Кто вызвал?

customerID 33875 userID 328

Трассировка?

requestID ec7764as32

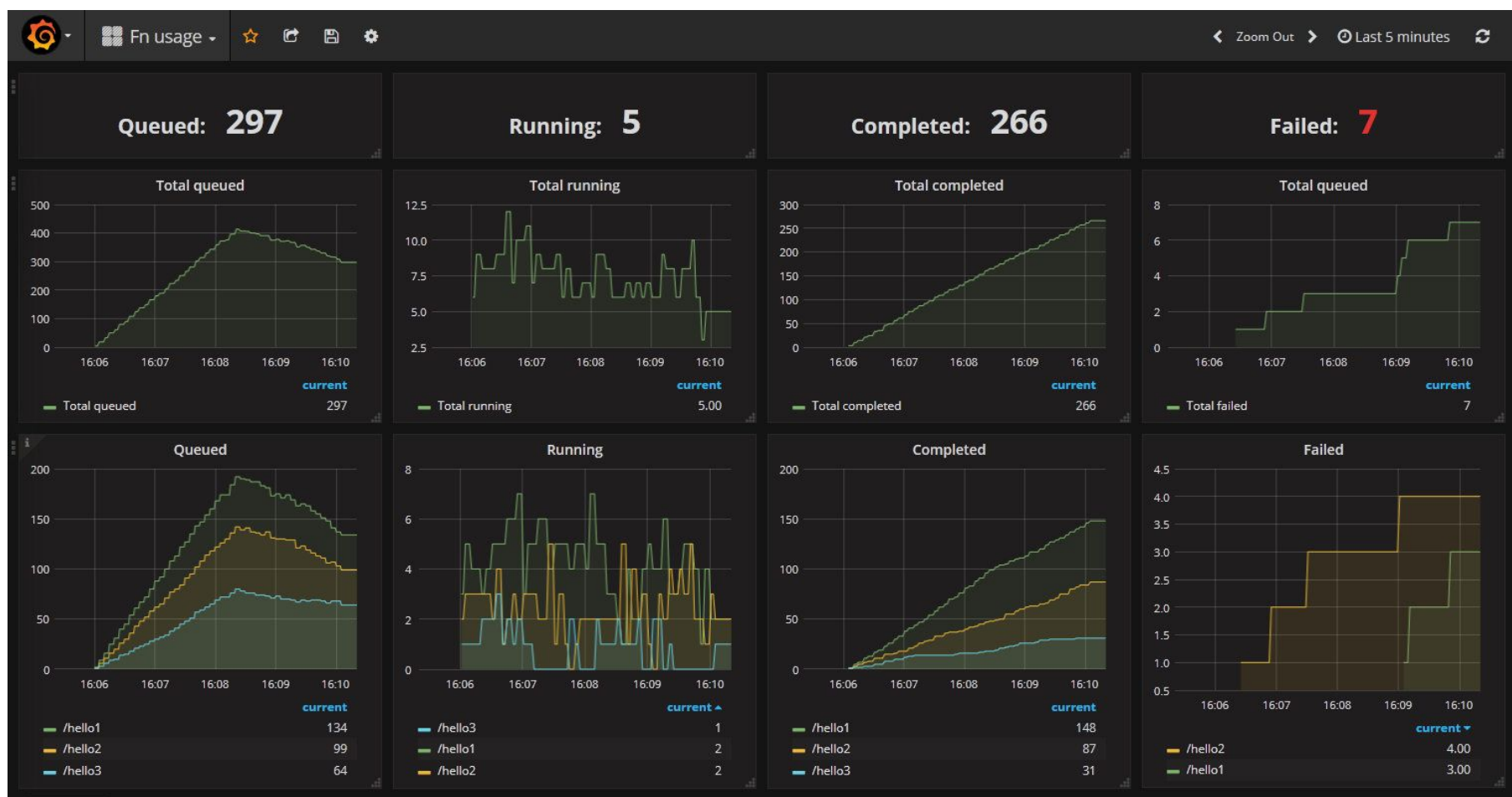
Что-то еще?

... ..

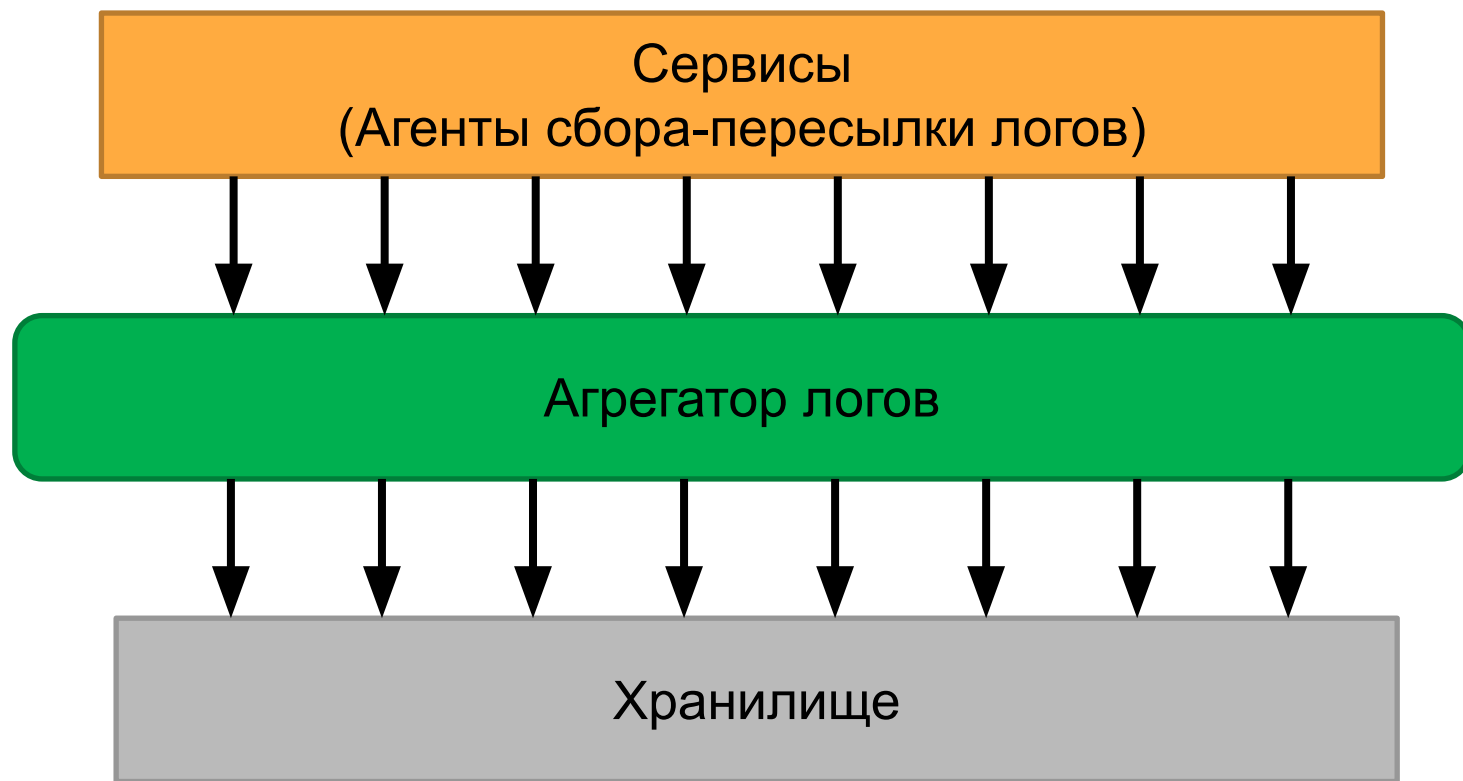
JSON

Структурные логи

Позволяют исследовать проблемы с более высокой точностью, выделяя наиболее важные моменты

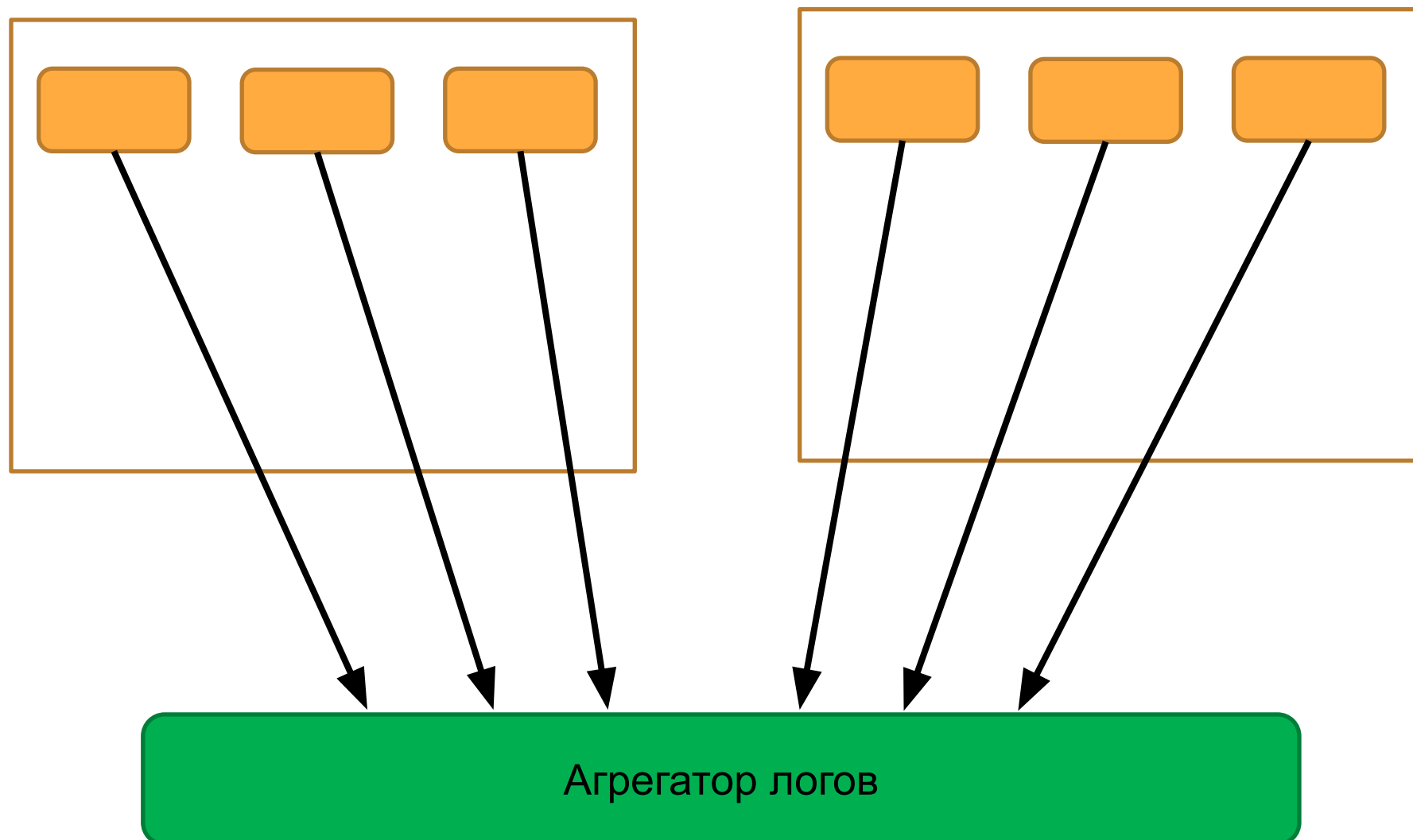


Общий подход



- Как можно быстрее доставить логи
- Структурировать для хранилища, если надо

Прямая запись



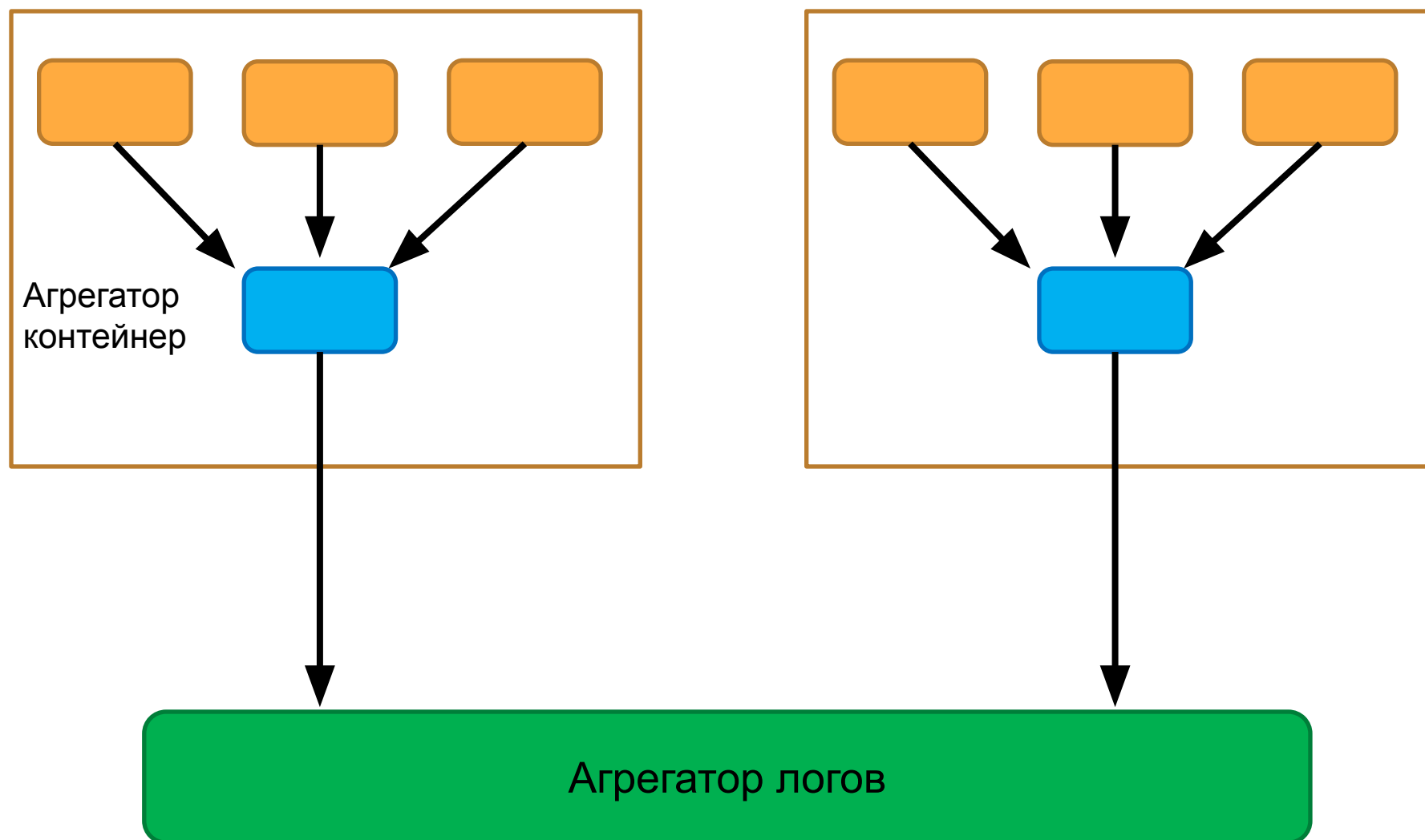
Прямая запись

Самое простое решение, но...

- Фиксированный адрес агрегатора логов
- Множество соединений с агрегатором логов
- Высокая нагрузка на агрегатор

Можно вручную увеличивать количество агрегаторов и управлять адресами, на ранних стадиях жизни приложения

Агрегатор рядом с сервисами



Агрегатор рядом с сервисами

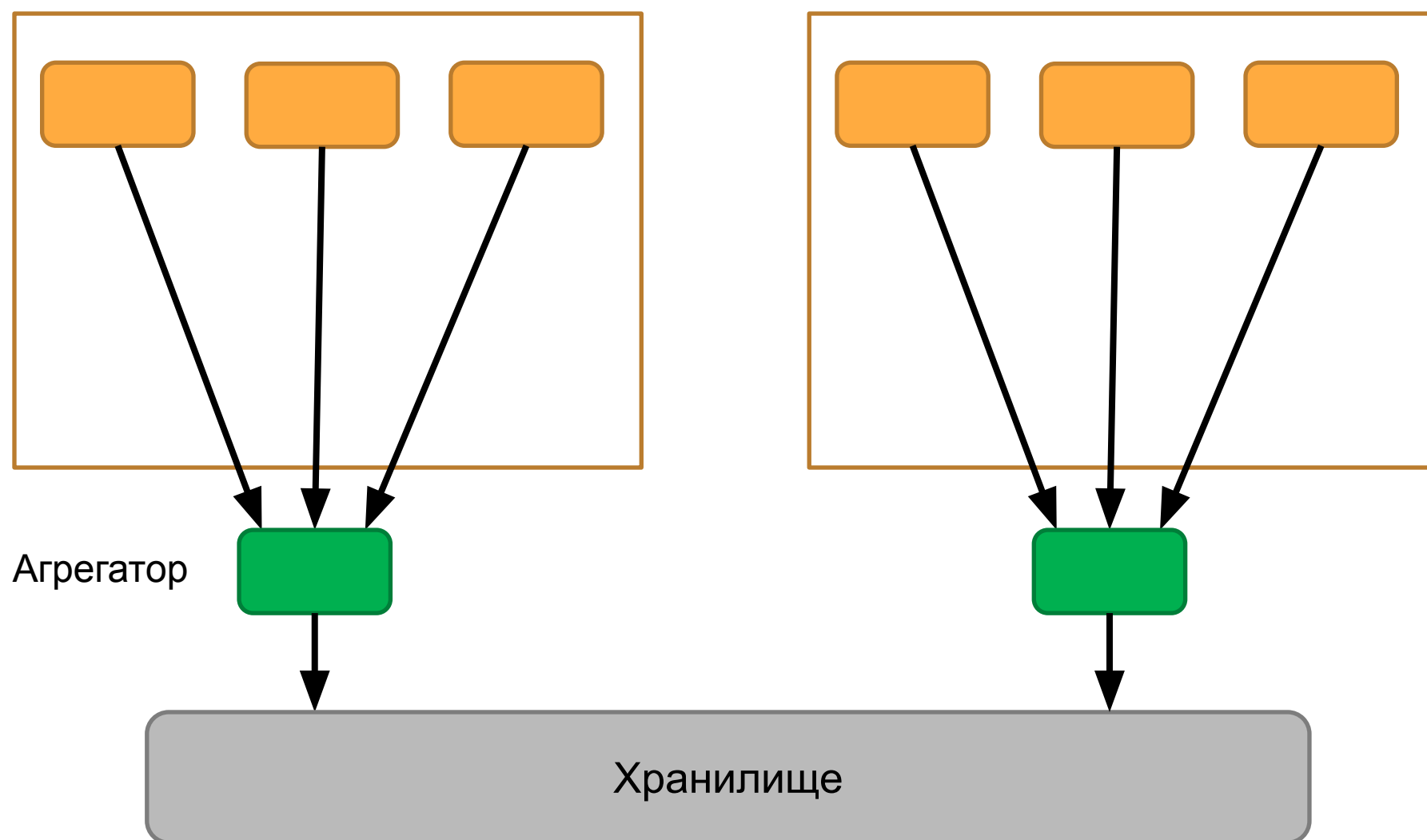
Основной минус – большее потребление ресурсов, дополнительный компонент.

Плюсы использования:

- Меньше соединений с агрегатором
- Уменьшение нагрузки на агрегатор
- Меньше настроек для сервисов
- Более гибкая конфигурация

Используется, когда логи передаются через общие сети интернет.

Агрегатор на стороне доставки



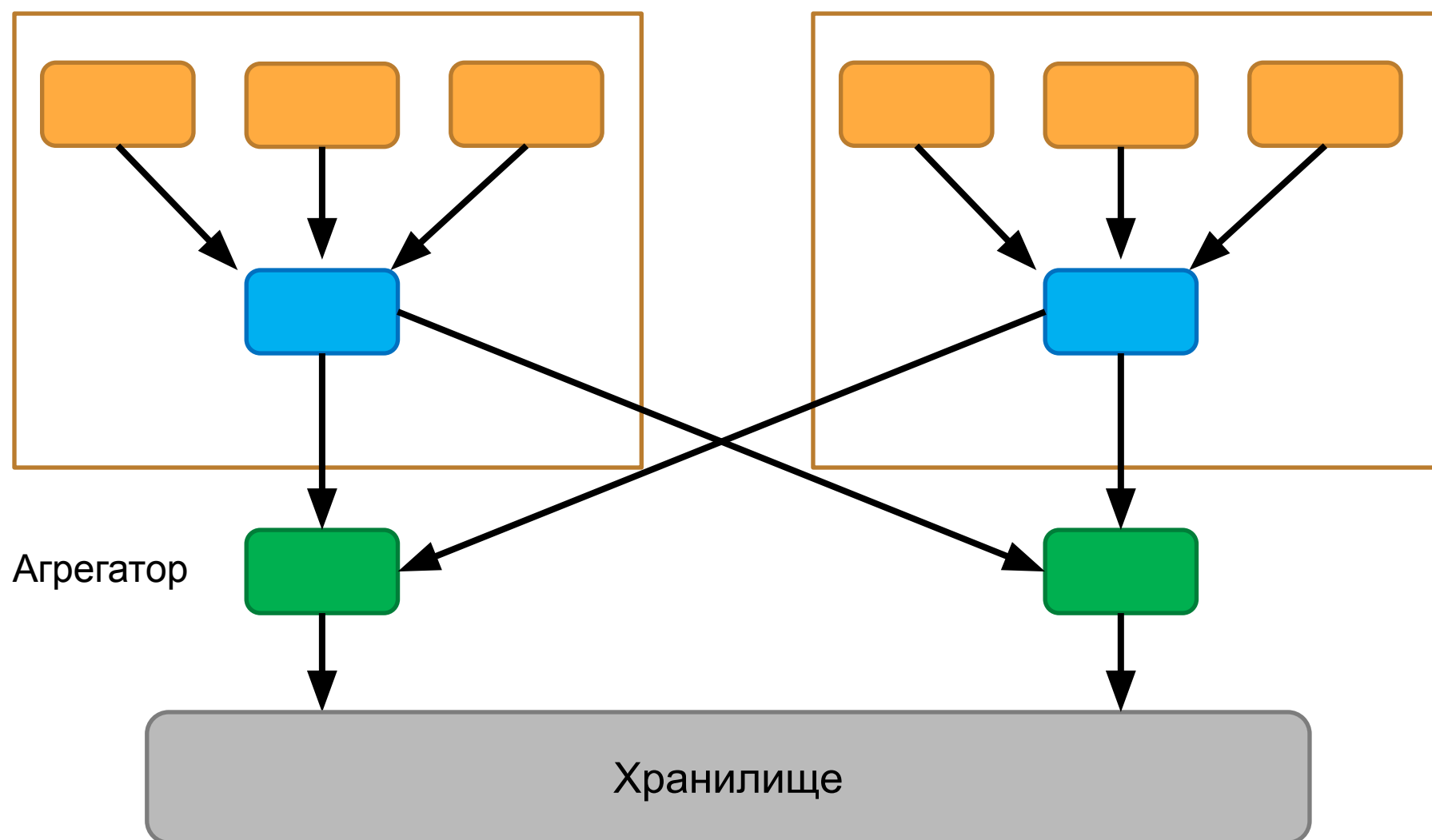
Агрегатор на стороне доставки

Избегание агрегации логов на стороне сервисов приводит к следующим издержкам:

- Изменения на стороне принятия логов вынуждают изменять конфигурацию поставщиков логов.
- Меньшая производительность из-за высокого количества соединений и данных

При таком подходе масштабирование агрегаторов без дополнительных трюков невозможно.

Агрегатор на двух сторонах



Агрегатор на двух сторонах

Наиболее оптимальная конфигурация в случае больших систем.

Минусы:

- Необходимо больше ресурсов для работы
- Сложность конфигурации

Плюсы:

- Возможность независимых изменений на обеих сторонах
- Улучшенная производительность и отказоустойчивость

Инструменты

- Elastic stack
- Grafana stack
- Облачные решения APM - Azure, AWS

Важно помнить

- Централизованное хранилище
- Возможность поиска
- Построение корреляции с другими логами и метриками системы

Выводы

- Используйте структурные логи, если есть возможность
- Используйте сайдкары для парсинга простых логов перед высылкой
- Выбирайте шаблон доставки логов в зависимости от сложности вашей системы
- Планируйте структуру логов заранее, создайте свой шаблон. Используйте служебную информацию о сервисах и запросах.

Что дальше?

- Назначение и использование метрик
- Инструментарий