OTUS

Онлайн-образование

Не забыл включить запись

Меня хорошо видно && слышно?

Ставьте плюсы, если все хорошо

Напишите в чат, если есть проблемы

План

- Снова про логи
- Централизованные системы логирования
- ELK, EFK, Graylog, etc.
- Что там с приложениями?

Зачем нужны логи?

- Видимость и понимание того, как работают наши системы
- Поиск ошибок и их причин
- Контекстное дополнения метрик систем или процессов

Где хранить логи?

Локальное хранение

- Не понимаем, что где происходит, пока сами не зайдем в систему и не посмотрим
- При большой ферме серверов не хватит рабочего времени на обход всех машин
- Нет возможности быстро локализовать проблему - следовательно, и ее решить

Почему не просто....

- Логи смотрят только когда "припекает"
- А когда "припекает", то и смотреть их становиться неудобно
- Поиск и работа с plain форматом
- Немалое количество сервисов и серверов
- Локальное хранение на серверах приложений ограничено по времени
- А еще бывает что сами серверы ограничены жизненным циклом

Централизованная система логирования

- Центральный сервер(ы) агрегирует всю информацию по логам
- Единая точка доступа ко все информации
- Возможность проведения анализа по всем системам
- Возможность кластеризации

Проблематика

- Централизованная система логирования не отменяет локального хранение логов
- Локальное хранение логов по-прежнему является самым надежным способом хранения
- Возможна потеря логов, если центральный сервер загружен или не доступен

Возможные требования

- Длительность хранения и целостность данных
- В централизованной системе логирования не должны хранится приватные данные
- Разделять централизованное хранение и данные для анализа

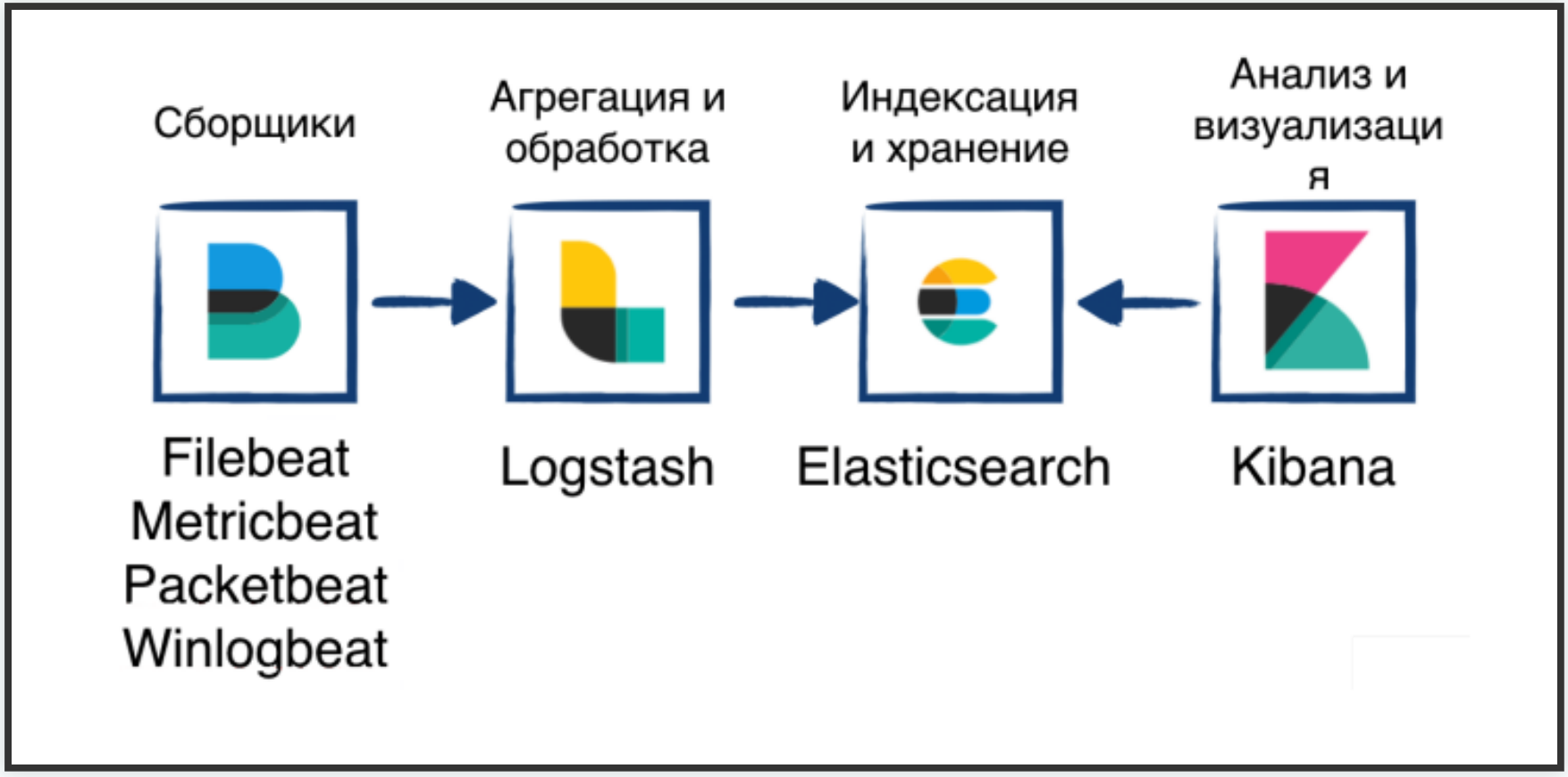
Требования к системам логирования

- Горизонтальная масштабируемость
- Надежность (отсутствие потери логов)
- Близость к real-time
- Должна быть недорогой

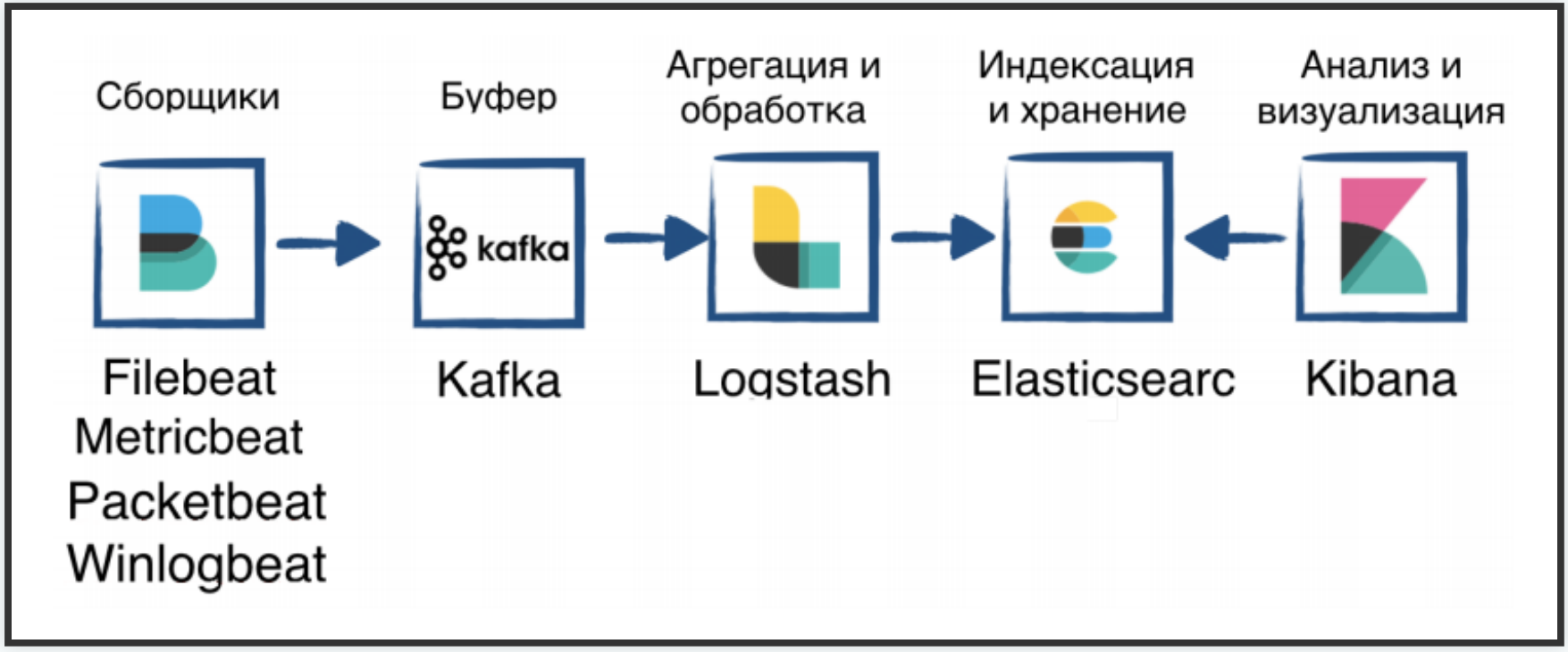
Текущий инструментарий

- Open source: Elastic Stack, Graylog2
- SaaS: Splunk(+ on-premise), Loggly, Papertrail, Datadog, logz.io
- Cloud Platform Service: Stackdriver Logging (GCP), CloudWatch (AWS)

ELK


Сборщики
Агрегация и обработка
Индексация и хранение
Анализ и визуализация
Filebeat
Metricbeat
Packetbeat
Winlogbeat
Logstash
Elasticsearch
Kibana

ELK


Сборщики
Буфер
Агрегация и обработка
Индексация и хранение
Анализ и визуализация
kafka
Filebeat
Metricbeat
Packetbeat
Winlogbeat
Kafka
Logstash
Elasticsearc
Kibana

Beats

Сборщики:

- Из лог файлов - filebeat
- Системных событий - winlogbeat
- Метрики приложений или сервисов (при помощи модулей) - metricbeat

Filebeat

- Один из основных способов сбора данных из лог файлов
- Легковесный
- Поддерживает работу с multiline
- Начиная с 6 версии умеет также обрабатывать логи (при помощи модулей)

Fluent bit

Все тот же лог-форвардер

- Event driven (использует асинхронную модель работы)
- I/O Handler
- Routing
- Upstream Manager
- TLSv1.2 / Security

Elasticsearch

- Движок (как сервис) для хранения и поиска данных
- Поддержка партиционирования/шардирования
- Возможность выполнения запросов для огрегации данных

Logstash

- Позволяет выстроить pipeline для приема, обработки (структурирования) и отправки обработанных логов
- Можно поставить свой после filebeat, в случае большого количества преобразований

Kibana

МНОГОФУНКЦИОНАЛЬНЫЙ UI ДЛЯ ELASTICSEARCH

- Поддержка визуализации и анализа
- Инструменты для обнаружения аномалий
- Функции мониторинга
- Функции отслеживания и уведомления (Алертинг)

ELK x-pack

Платные расширения к ELK*, предоставляют доступ к:

- Мониторингу
- Машинному обучению (детекор аномалий)
- Управление правами доступа
- Алертинг

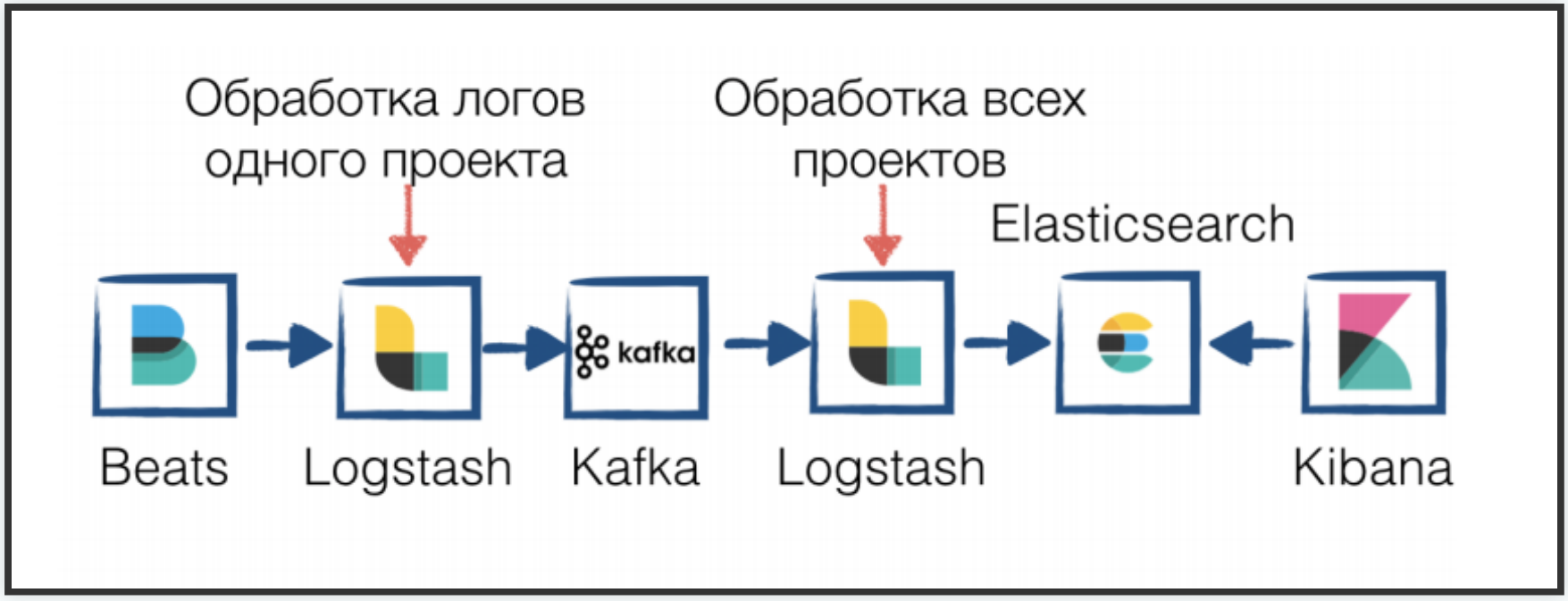
Graylog

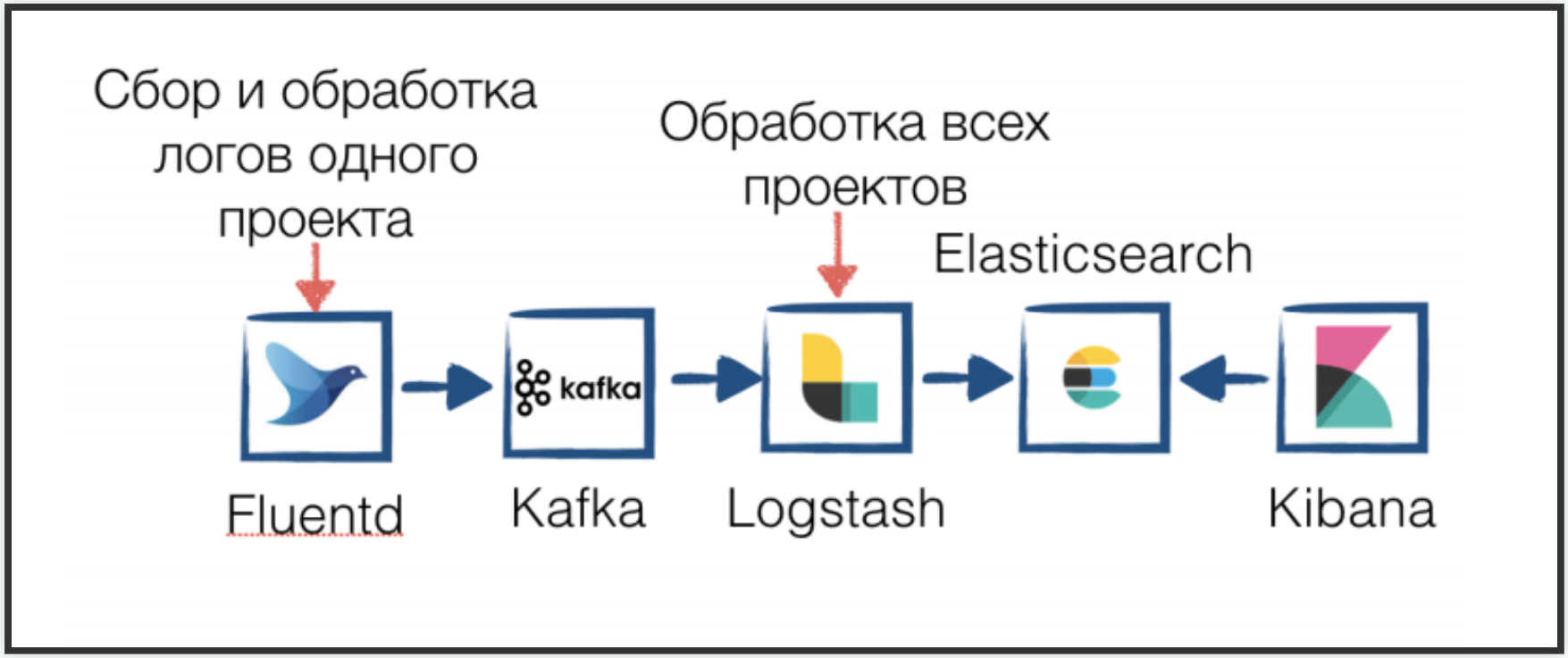
- В качестве бекэнда для хранения используется Elasticsearch
- Собственный web-интерфейс
- Дополнительные модули
  - ADVANCED SEARCH
  - FAULT TOLERANCE
  - CONTENT PACKS
  - GRAYLOG SIDECAR
- Enterprice version

Opendistro

- Лицензия Apache 2 (Opensource)
- Управление досутом (Standalone, LDAP, etc.)
- Alerting
- Поддержка SQL ^_^
- Performance Analyzer
- Разрабатывает AWS

Варианты построения систем логирования


Обработка логов
одного проекта
Обработка всех
проектов
Elasticsearch
kafka
Beats
Logstash
Kafka
Logstash
Kibana


Сбор и обработка логов одного проекта
Обработка всех проектов
Elasticsearch
kafka
Fluentd
Kafka
Logstash
Kibana

Какую активность приложения логировать?

- Запросы и ответы
- Ошибки
- Вызовы ко всем внешним сервисам и API
- Бизнес события: создание пользователя, платеж
- Время чтения/записи к БД
- etc

Возможные варианты формирования\отправки логов

- Log4j
- Structlog
- Lograge
- etc

Как структурировать логи

- Заранее договориться о формате
- Парсить существующий формат