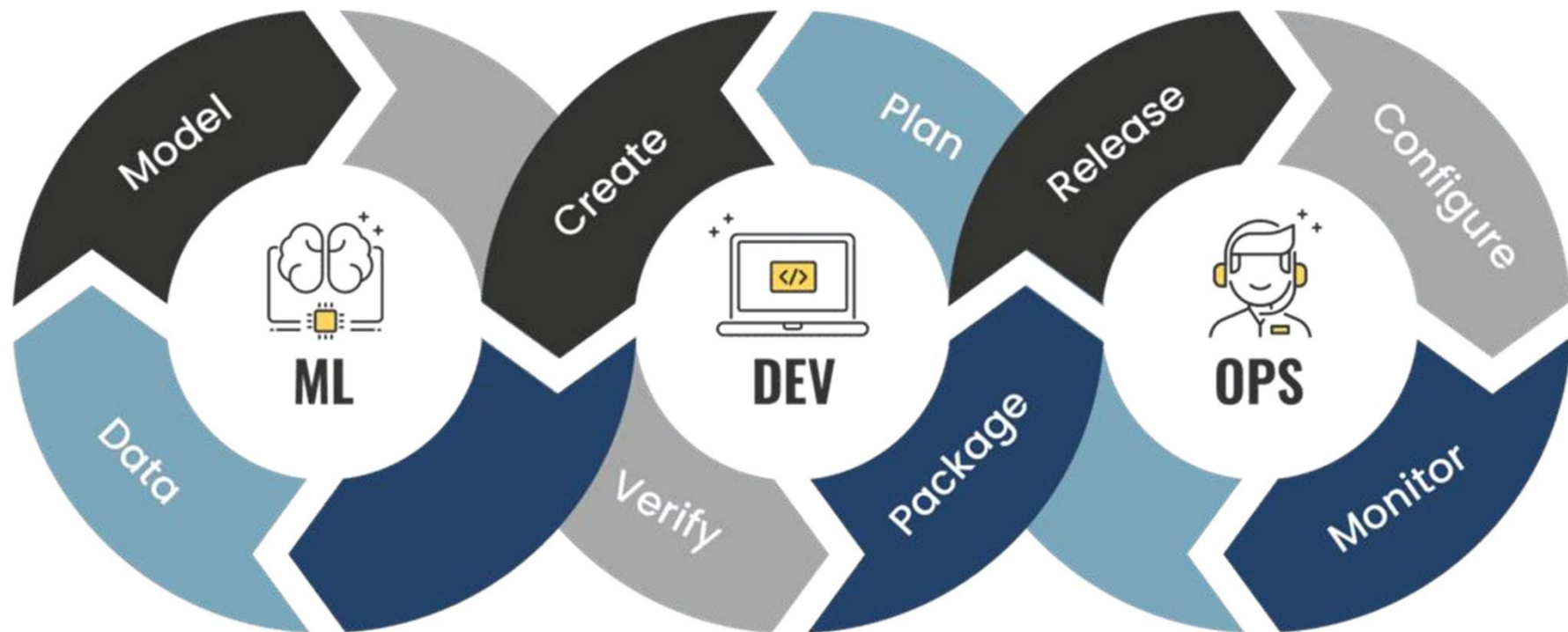
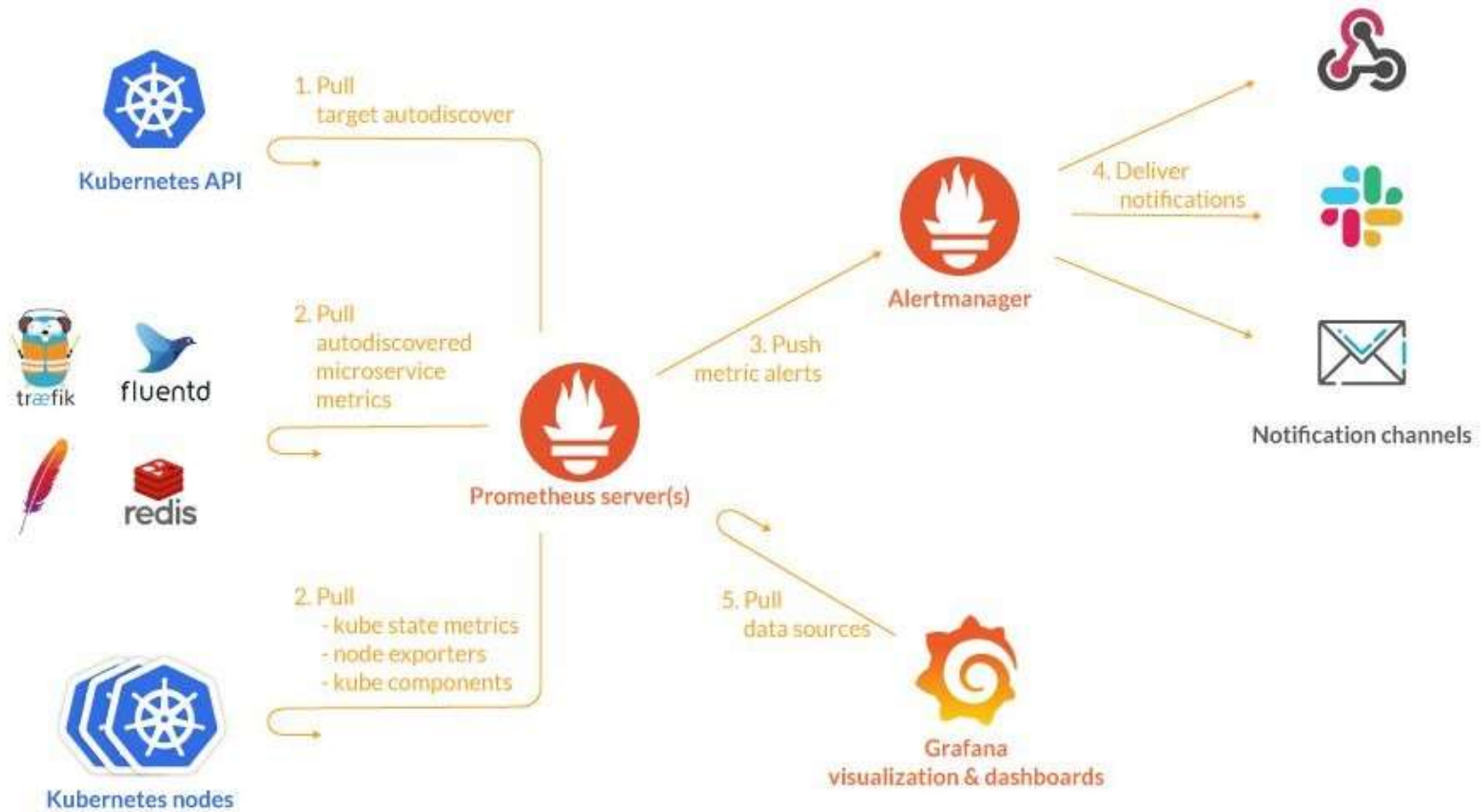


Мониторинг сервиса и модели

MLOps = ML + Dev + Ops



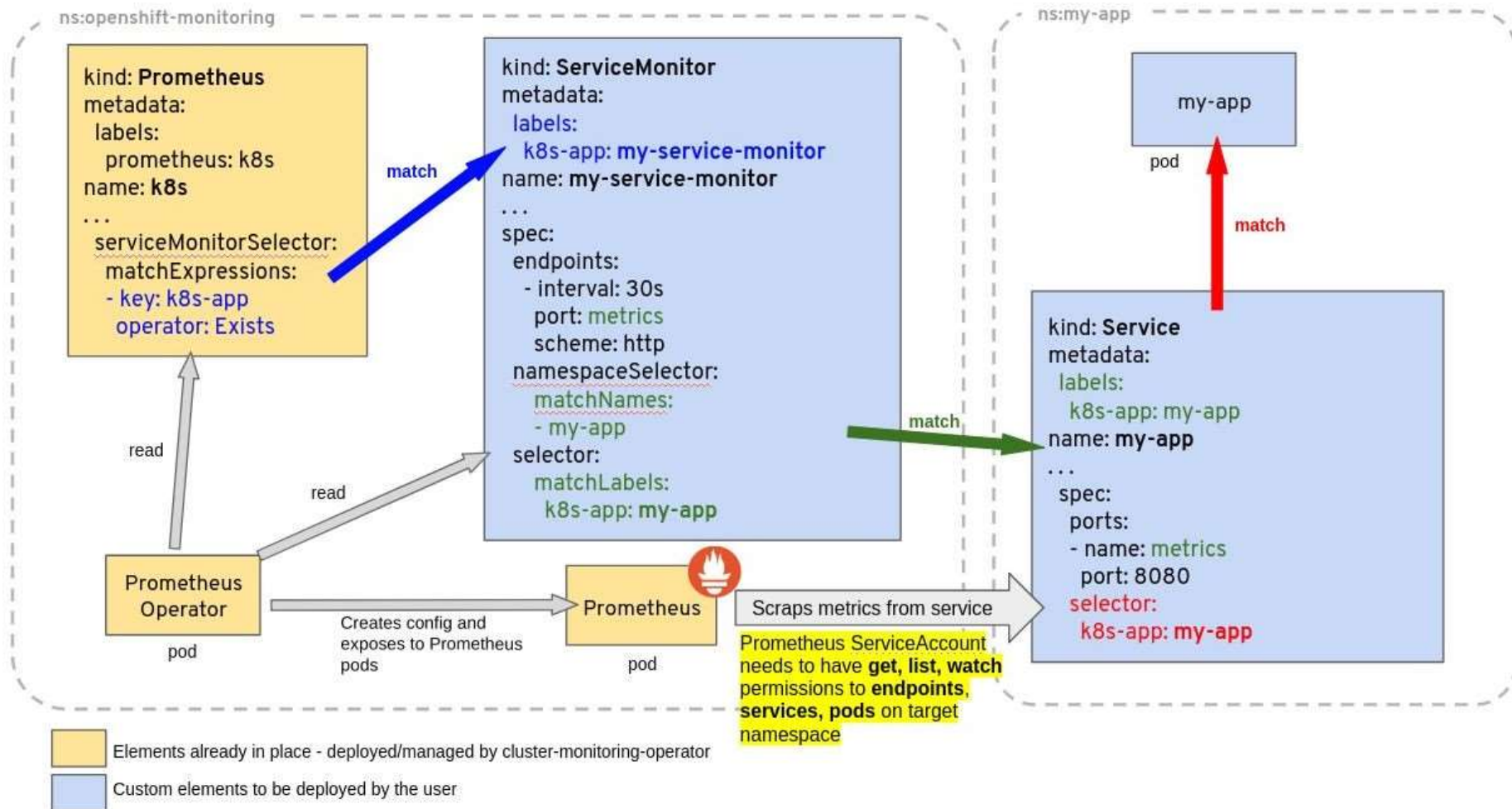
Prometheus



Подключение метрик FastAPI

1. Используем `starlette_exporter.PrometheusMiddleware`
2. Экспортируем метрики через `/metrics`
3. Добавляем метки в k8s манифест для обнаружения сервиса
4. Создаем `ServiceMonitor`
5. Создаем дашборд в Grafana

ServiceMonitor



Библиотека prometheus_client

Типы метрик:

- Counter - только увеличение
- Gauge - увеличение и уменьшение
- Summary - агрегация
- Histogram
- Info - любая нечисловая информация

Каждая метрика может иметь произвольные метки (labels)

Метрики модели

- распределение предсказаний (для обнаружения сдвигов)
- кол-во аномалий
- кол-во неизвестных категорий
- распределение входных признаков
- распределение времени предсказания

Добавление метрик

1. Создать необходимый счетчик
2. Обновлять данные счетчика при наступлении события
3. Настроить дашборд в Grafana