## SRE 4 (03.12.2021) | Теория SRE | Часть 4: SLO alerts

Для каждого из наших SLO (availability, latency) сделаем по два алерта:

- warning чувствительный к изменениям, но шумный а следовательно интересен только для ретроспективного анализа.
- раде дольше реагирует на изменения, но менее шумный, а значит подходящий для уведомлений.

## **Recording rules**

Для начала, создадим два основных recording rule. Для этого, создадим объект PrometheusRule . Можно это сделать через файлик p8s-rules.yaml положив туда следующий контент:

Замечание: Важно заменить \$city на имя вашего города

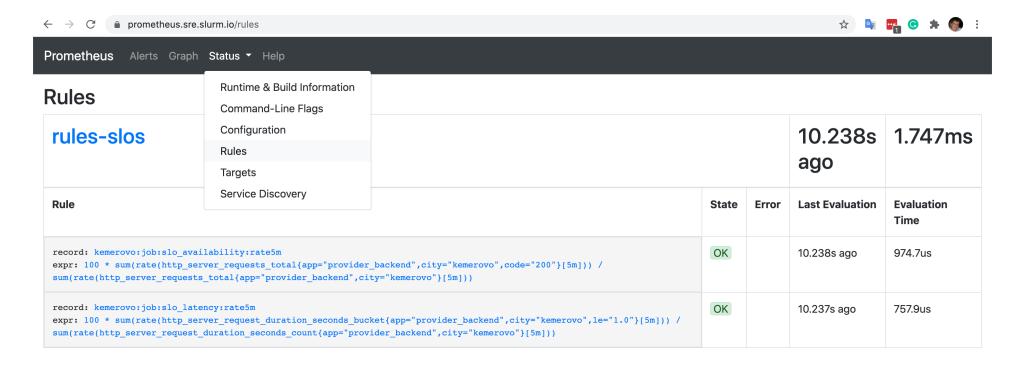
```
/ sum(rate(http_server_requests_total{app="provider_backend", city="$city"}[5m]))

- record: $city:job:slo_latency:rate5m
    expr: |
        100
        * sum(rate(http_server_request_duration_seconds_bucket{app="provider_backend", city="$city", le="1.0"}[5m]))
        / sum(rate(http_server_request_duration_seconds_count{app="provider_backend", city="$city"}[5m]))
```

Применяем конфигурацию в кластере:

```
kubectl apply -f p8s-rules.yaml
```

Убеждаемся, что prometheus подхватил изменения (процесс может занять до нескольких минут). В закладке Status -> Rules должны появится созданные recording rules.



## Алерты

После создания recording rule, можно непосредственно приступить к определению алертов. Создается подобным образом. Создаем файл p8s-alerts.yaml

## Замечание: Не забываем про \$city

```
kind: PrometheusRule
apiVersion: monitoring.coreos.com/v1
metadata:
   name: prometheus-slo-alerts
   namespace: $city
spec:
```

```
groups:
- name: alerts-slo
  rules:
    alert: $cityAvailabilitySLOPage
      expr: $city:job:slo availability:rate5m < 99</pre>
      for: 5m
      labels:
        severity: page
        team: $city
      annotations:
        summary: "Availability SLO is low"
        description: "Availability SLO is below target 99% for already 5 minutes (current value: {{ printf \"%.2f\" $vare}
    - alert: $cityAvailabilitySLOWarning
      expr: $city:job:slo availability:rate5m < 99.5</pre>
      for: 1m
      labels:
        severity: warning
        team: $city
      annotations:
        summary: "Availability SLO is low"
        description: "Availability SLO is approaching target 99% (current value: {{ printf \"%.2f\" $value }}%)"
    alert: $cityLatencySLOPage
      expr: $city:job:slo_latency:rate5m < 99</pre>
      for: 5m
      labels:
        severity: page
        team: $city
      annotations:
        summary: "Latency SLO is low"
```

```
description: "Latency SLO is below target 99% for already 5 minutes (current value: {{ printf \"%.2f\" $value }}

- alert: $cityLatencySLOWarning
    expr: $city:job:slo_latency:rate5m < 99.5
    for: 1m
    labels:
        severity: warning
        team: $city
    annotations:
        summary: "Latency SLO is low"
        description: "Latency SLO is approaching target 99% (current value: {{ printf \"%.2f\" $value }}%)"</pre>
```

Тут мы создали по два алерта для availability и latency SLO.

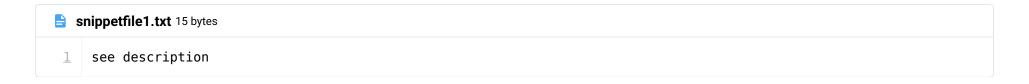
Применяем конфигурацию в кластере:

```
kubectl apply -f p8s-alerts.yaml
```

После этого, в prometheus в закладке Alerts должны появится четыре алерта.



Edited 2 hours ago



Write a comment or drag your files here...

<u>Markdown</u> is supported