

주제 : 웹 서비스 모니터링 아키텍처 구축

조은새 : 웹 서비스 모니터링 아키텍처 구축

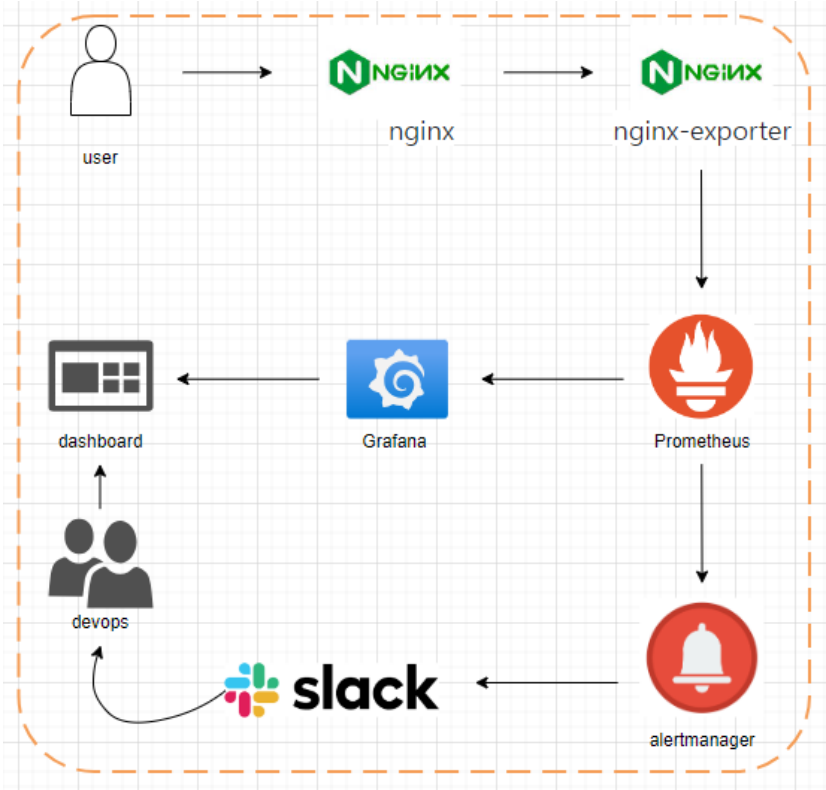
공상욱 : PPT 제작

완료 사항

- | | |
|-------------------------------|----------------------------------|
| 1. Ubuntu | => (RAM 8G, CPU 4, Memory 45G) |
| 2. Nginx, Nginx-exporter | => helm install |
| 3. Prometheus | => helm install |
| 4. Prometheus, Nginx-exporter | => IP Target 설정 |
| 5. Prometheus, Grafana | => 연결 |
| 6. AlertManager | => slack 연동 |

동작 흐름

1. Nginx에서 지표를 수집하고 노출하는 nginx-exporter가 실행됩니다.
2. Prometheus는 정의된 간격으로 nginx-exporter의 엔드포인트에서 지표를 스크랩하고 수집합니다. (1분 ~ 3분)
3. Prometheus는 수집한 데이터를 저장하고, 사용자가 쿼리할 수 있도록 합니다.
4. Grafana는 Prometheus로부터 수집한 데이터를 사용자에게 대시보드를 제공하고 시각적 모니터링 정보를 표시 합니다.
5. AlertManager는 Prometheus에서 발생한 경고를 관리하고 전달, Slack 리시버를 설정하여 알람을 Slack으로 전송
6. 알람 메시지는 명확하고 유용한 정보를 포함하여 운영팀이 신속하게 대응할 수 있습니다.



Slack 리시버는 Prometheus AlertManager에서 발생한 알람을 Slack으로 전송하는 메커니즘

```
receivers:  
- name: 'slack-notifications'  
  slack_configs:  
    - send_resolved: true  
      text: "{{ .CommonAnnotations.summary }}"  
      title: "{{ .CommonLabels.severity }}: {{ .CommonLabels.alertname }}"  
      channel: '#your-slack-channel'  
      api_url: 'https://hooks.slack.com/services/your/slack/webhook/url'
```

1.Nginx

1. **역할:** 웹 서버 및 리버스 프록시로 사용되는 Nginx는 클라이언트 요청을 받아들이고 정적 파일을 제공하며, 동적 콘텐츠를 처리하고 백엔드 서비스로의 요청을 전달하는 역할을 합니다.
2. **모니터링:** Nginx의 웹 서버 지표를 수집하기 위해 Prometheus Exporter를 사용할 수 있습니다.

2. nginx-exporter

1. 역할

- 1-1. Nginx-exporter는 Nginx 웹 서버의 상태 및 지표를 수집합니다.
- 1-2. 이러한 정보를 Prometheus에서 소비할 수 있는 형식으로 노출합니다.

3.Prometheus

1. **역할:** 모니터링 및 경고를 위한 오픈 소스 시스템으로, 여러 서버 및 서비스에서 수집된 지표를 저장하고 조회합니다.
2. **수집:** 여러 서비스에서 지표를 주기적으로 스크랩하여 데이터베이스에 저장합니다.
3. **알림:** 지정된 규칙에 따라 경고 및 알림을 생성하고, 이러한 정보를 Alertmanager로 전달합니다.

4.Grafana

1. **역할:** 모니터링 데이터를 시각적으로 표현하고 대시보드를 생성하는 툴입니다.
2. **시각화:** Prometheus에서 수집한 지표를 기반으로 다양한 차트 및 대시보드를 생성하여 사용자에게 시각적으로 제공합니다.
3. **알림 대시보드:** Prometheus Alertmanager와 통합하여 경고 및 알림에 대한 대시보드를 생성할 수 있습니다.

동작 확인

```

goodbird@goodbird-k8s:~$ kubectl get pod -n project
NAME                                     READY   STATUS    RESTARTS   AGE
my-grafana-549f99fb67-n6fsm             1/1     Running   16 (59m ago)  25h
my-nginx-584ffcb496-z2hmb               2/2     Running   0              10s
my-prometheus-alertmanager-0            1/1     Running   16 (59m ago)  25h
my-prometheus-kube-state-metrics-7d9f488d69-57thh  1/1     Running   31 (53m ago)  25h
my-prometheus-prometheus-node-exporter-sjfs6  1/1     Running   16 (59m ago)  25h
my-prometheus-prometheus-pushgateway-86f4f879d8-mjgcj  1/1     Running   16 (59m ago)  25h
my-prometheus-server-db457977c-7wxh4      2/2     Running   32 (59m ago)  25h
goodbird@goodbird-k8s:~$ kubectl get svc -n project
NAME                                     TYPE          CLUSTER-IP      EXTERNAL-IP      PORT(S)                                     AGE
my-grafana                             LoadBalancer  10.109.25.110   192.168.49.12   80:30308/TCP                             25h
my-nginx                               LoadBalancer  10.97.208.57    192.168.49.10   80:31192/TCP,9113:31275/TCP              19s
my-prometheus-alertmanager              ClusterIP      10.97.121.230   <none>           9093/TCP                                  25h
my-prometheus-alertmanager-headless     ClusterIP      None            <none>           9093/TCP                                  25h
my-prometheus-kube-state-metrics         ClusterIP      10.102.30.19    <none>           8080/TCP                                  25h
my-prometheus-prometheus-node-exporter   ClusterIP      10.102.85.134   <none>           9100/TCP                                  25h
my-prometheus-prometheus-pushgateway     ClusterIP      10.103.75.157   <none>           9091/TCP                                  25h
my-prometheus-server                    LoadBalancer  10.106.187.79   192.168.49.11   80:30895/TCP                             25h
goodbird@goodbird-k8s:~$

```

nginx, nginx-exporter, prometheus, Grafana 의 pod, svc

Namespace => project

spec.type => LoadBalancer

- nginx, nginx-exporter 시각적인 동작 확인
- Prometheus, Grafana 모니터링을 위해



Welcome to nginx!


If you see this page, the nginx web server is successfully installed and working. Further configuration is required.

For online documentation and support please refer to nginx.org.
Commercial support is available at nginx.com.

Thank you for using nginx.

```
# HELP go_gc_duration_seconds A summary of the pause duration of garbage collection cycles.
# TYPE go_gc_duration_seconds summary
go_gc_duration_seconds{quantile="0"} 4.1621e-05
go_gc_duration_seconds{quantile="0.25"} 4.1621e-05
go_gc_duration_seconds{quantile="0.5"} 0.000150098
go_gc_duration_seconds{quantile="0.75"} 0.000150098
go_gc_duration_seconds{quantile="1"} 0.000150098
go_gc_duration_seconds_sum 0.000191719
go_gc_duration_seconds_count 2
# HELP go_goroutines Number of goroutines that currently exist.
# TYPE go_goroutines gauge
go_goroutines 15
# HELP go_info Information about the Go environment.
# TYPE go_info gauge
go_info{version="go1.21.6"} 1
# HELP go_memstats_alloc_bytes Number of bytes allocated and still in use.
# TYPE go_memstats_alloc_bytes gauge
go_memstats_alloc_bytes 1.586272e+06
# HELP go_memstats_alloc_bytes_total Total number of bytes allocated, even if freed.
# TYPE go_memstats_alloc_bytes_total counter
go_memstats_alloc_bytes_total 4.77072e+06
# HELP go_memstats_buck_hash_sys_bytes Number of bytes used by the profiling bucket hash table.
# TYPE go_memstats_buck_hash_sys_bytes gauge
go_memstats_buck_hash_sys_bytes 4626
# HELP go_memstats_frees_total Total number of frees.
# TYPE go_memstats_frees_total counter
go_memstats_frees_total 14261
# HELP go_memstats_gc_sys_bytes Number of bytes used for garbage collection system metadata.
# TYPE go_memstats_gc_sys_bytes gauge
go_memstats_gc_sys_bytes 3.64684e+06
# HELP go_memstats_heap_alloc_bytes Number of heap bytes allocated and still in use.
# TYPE go_memstats_heap_alloc_bytes gauge
go_memstats_heap_alloc_bytes 1.586272e+06
# HELP go_memstats_heap_idle_bytes Number of heap bytes waiting to be used.
# TYPE go_memstats_heap_idle_bytes gauge
go_memstats_heap_idle_bytes 3.997696e+06
# HELP go_memstats_heap_inuse_bytes Number of heap bytes that are in use.
# TYPE go_memstats_heap_inuse_bytes gauge
go_memstats_heap_inuse_bytes 3.702784e+06
# HELP go_memstats_heap_objects Number of allocated objects.
# TYPE go_memstats_heap_objects gauge
go_memstats_heap_objects 2304
# HELP go_memstats_heap_released_bytes Number of heap bytes released to OS.
# TYPE go_memstats_heap_released_bytes gauge
go_memstats_heap_released_bytes 2.547712e+06
```

☐ Use local time ☐ Enable query history ☒ Enable autocomplete ☒ Enable highlighting ☒ Enable linter

 nginx_http_requests_total

  Execute

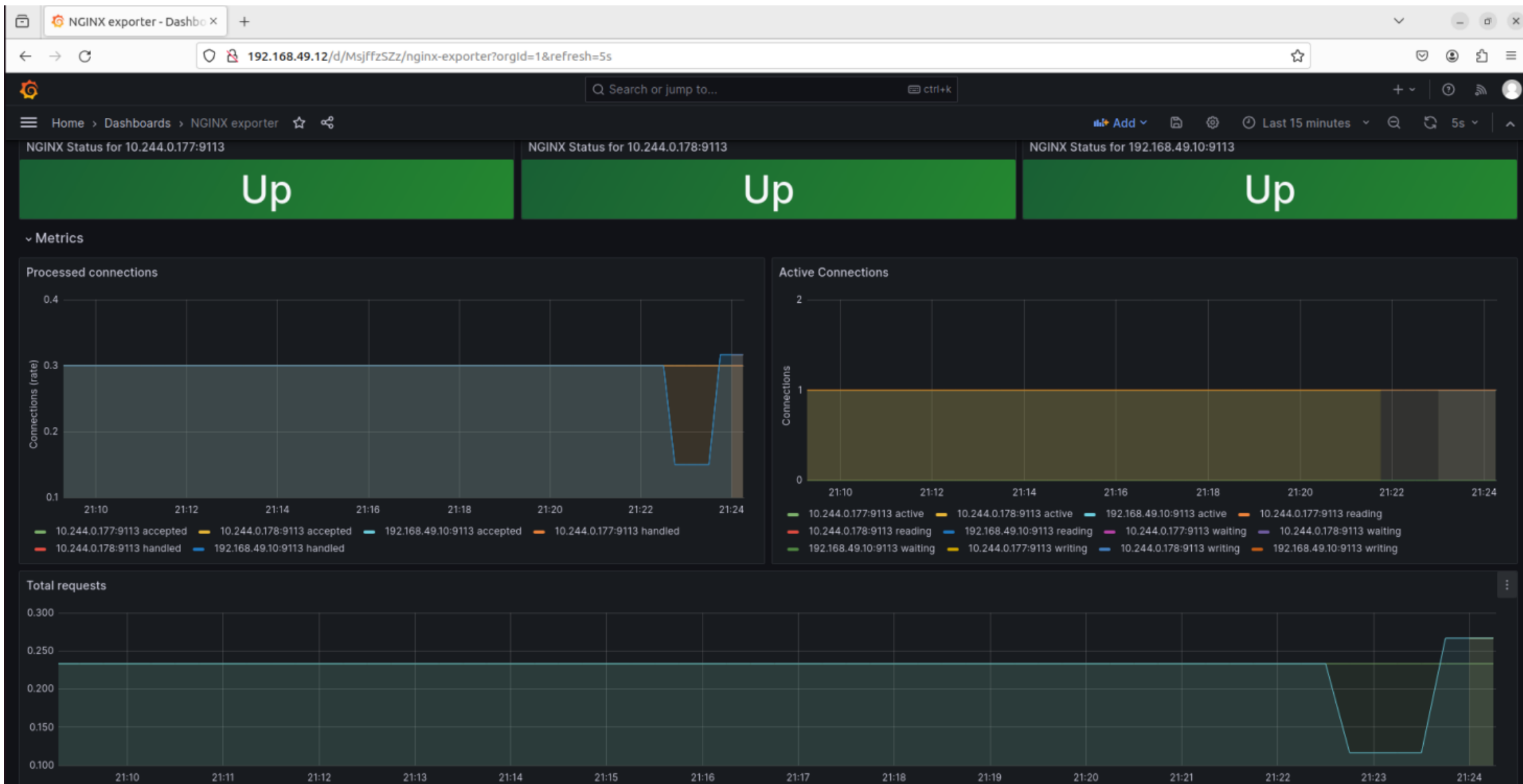
Table [Graph](#)

Load time: 7ms Resolution: 14s Result series: 2

	Evaluation time	
nginx_http_requests_total{app_kubernetes_io_instance="my-nginx", app_kubernetes_io_managed_by="Helm", app_kubernetes_io_name="nginx", app_kubernetes_io_version="1.25.3", helm_sh_chart="nginx-15.6.2", instance="10.244.0.178:9113", job="kubernetes-service-endpoints", namespace="project", node="minikube", service="my-nginx"}		10054
nginx_http_requests_total{instance="192.168.49.10:9113", job="nginx"}		10051

[Remove Panel](#)

Add Panel



Targets

All scrape pools ▾

All

Unhealthy

Expand All

🔍

Filter by endpoint or labels

✓ Unknown

✓ Unhealthy

✓ Healthy

alertmanager (1/1 up)

show less

Endpoint	State	Labels	Last Scrape	Scrape Duration	Error
http://my-prometheus-alertmanager:9093/metrics	UP	<div>Instance="my-prometheus-alertmanager:9093"</div> <div>job="alertmanager"</div> <div>▼</div>	7.778s ago	2.333ms	

kubernetes-apiservers (1/1 up)

show more

kubernetes-nodes (1/1 up)

show more

kubernetes-nodes-cadvisor (1/1 up)

show more

kubernetes-pods (2/2 up)

show more

kubernetes-service-endpoints (4/4 up)

show more

nginx (1/1 up)

show less

Endpoint	State	Labels	Last Scrape	Scrape Duration	Error
http://192.168.49.10:9113/metrics	UP	<div>Instance="192.168.49.10:9113"</div> <div>job="nginx"</div> <div>▼</div>	54.488s ago	3.116ms	

prometheus (1/1 up)

show more

prometheus-pushgateway (1/1 up)

show more

```
goodbird@goodbird-k8s:~$ ab -n 10000 -c 10 http://192.168.49.10/
This is ApacheBench, Version 2.3 <$Revision: 1879490 $>
Copyright 1996 Adam Twiss, Zeus Technology Ltd, http://www.zeustech.net/
Licensed to The Apache Software Foundation, http://www.apache.org/
```

```
Benchmarking 192.168.49.10 (be patient)
```

```
Completed 1000 requests
Completed 2000 requests
Completed 3000 requests
Completed 4000 requests
Completed 5000 requests
Completed 6000 requests
Completed 7000 requests
Completed 8000 requests
Completed 9000 requests
Completed 10000 requests
Finished 10000 requests
```

```
Server Software:      nginx
Server Hostname:      192.168.49.10
Server Port:          80
```

```
Document Path:        /
Document Length:       615 bytes
```

```
Concurrency Level:     10
Time taken for tests:   1.035 seconds
Complete requests:      10000
Failed requests:         0
Total transferred:      8700000 bytes
HTML transferred:       6150000 bytes
Requests per second:    9665.73 [#/sec] (mean)
Time per request:       1.035 [ms] (mean)
Time per request:       0.103 [ms] (mean, across all concurrent requests)
Transfer rate:          8212.09 [Kbytes/sec] received
```

Connection Times (ms)

	min	mean	mean[+/-sd]	median	max
Connect:	0	0	0.1	0	1
Processing:	0	1	0.3	1	5
Waiting:	0	1	0.2	0	5
Total:	0	1	0.3	1	6

ERROR: The median and mean for the waiting time are more than twice the standard deviation apart. These results are NOT reliable.

Percentage of the requests served within a certain time (ms)

50%	1
66%	1
75%	1

☒ Inactive (1) ☒ Pending (0) ☒ Firing (0)

Filter by name or labels

☒ Show annotations

/etc/config/alerting_rules.yml > example

inactive

▼ WebServerConnectivity (0 active)

name: WebServerConnectivity

expr: nginx_http_requests_total{instance="192.168.49.10:9113",job="nginx"} > 100

for: 1m

labels:

severity: critical

annotations:

description: The rate of HTTP requests is higher than 100 per minute.

summary: High request rate on nginx

✓ Inactive (0)

✓ Pending (1)

✓ Firing (0)

Q

Filter by name or labels

✓

Show annotations

/etc/config/alerting_rules.yml > example

pending (1)

▼ WebServerConnectivity (1 active)

name: WebServerConnectivity
expr: nginx_http_requests_total{instance="192.168.49.10:9113",job="nginx"} > 100
for: 1m
labels:
 severity: critical
annotations:
 description: The rate of HTTP requests is higher than 100 per minute.
 summary: High request rate on nginx

Labels	State	Active Since	Value
alertname=WebServerConnectivity instance=192.168.49.10:9113 job=nginx severity=critical	PENDING	2024-01-16T12:26:20.526679846Z	10051

Annotations

description
The rate of HTTP requests is higher than 100 per minute.
summary
High request rate on nginx

👍 Inactive (0)

👍 Pending (0)

👍 Firing (1)

🔍

Filter by name or labels

👤

Show annotations

/etc/config/alerting_rules.yml > example

firing (1)

▼ WebServerConnectivity (1 active)

name: WebServerConnectivity

expr: nginx_http_requests_total{instance="192.168.49.10:9113",job="nginx"} > 100

for: 1m

labels:

severity: critical

annotations:

description: The rate of HTTP requests is higher than 100 per minute.

summary: High request rate on nginx

Labels	State	Active Since	Value
alertname=WebServerConnectivity instance=192.168.49.10:9113 job=nginx severity=critical	FIRING	2024-01-16T12:26:20.526679846Z	10065

Annotations

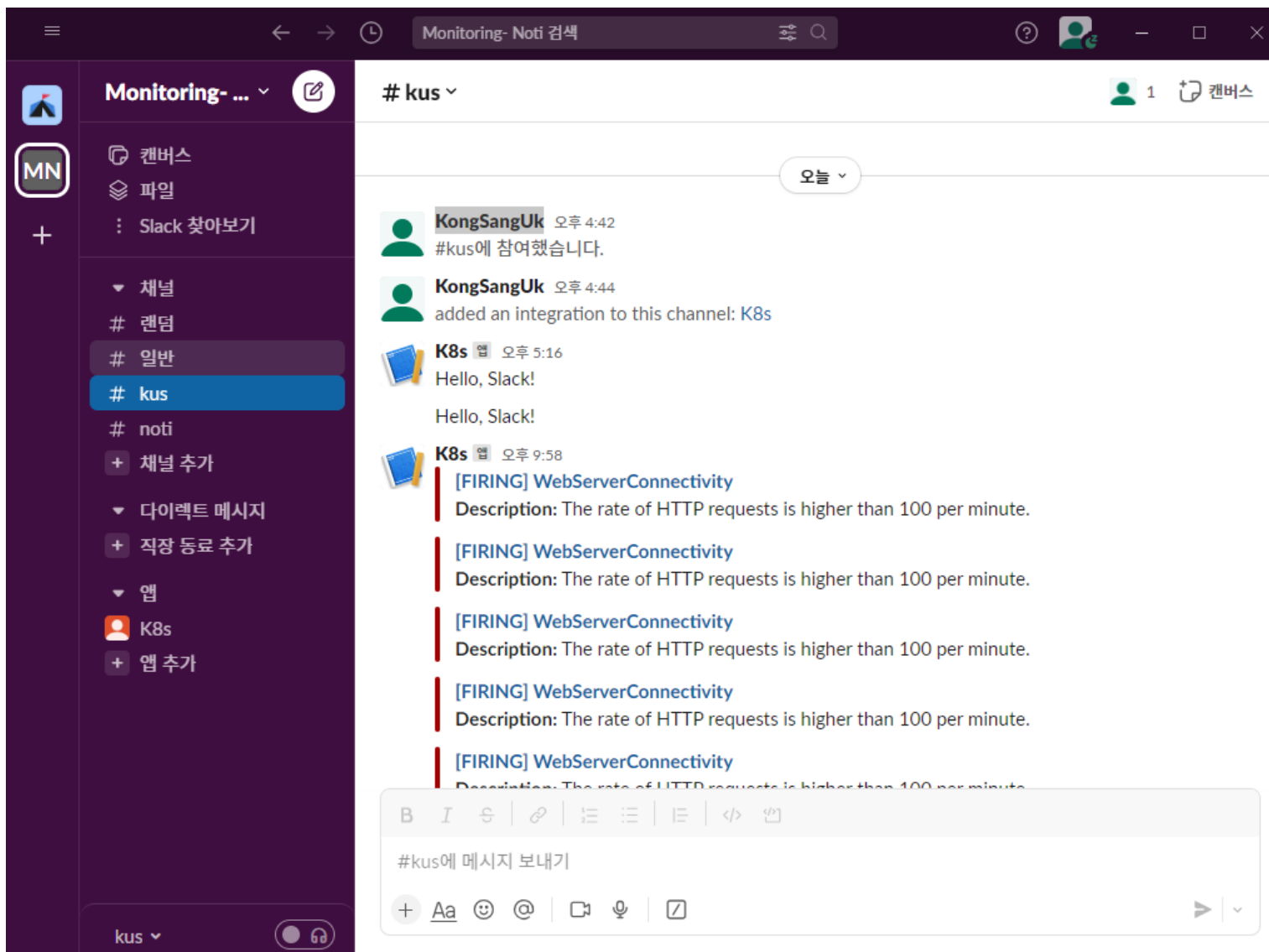
description

The rate of HTTP requests is higher than 100 per minute.

summary

High request rate on nginx

```
curl -X POST -H 'Content-type: application/json' --data '{"text":"Hello, Slack!"}' [ 토큰 ]
```



alertmanager VS alertmanagers


```
podLabels: {}
```

```
## Prometheus AlertManager configuration
```

```
##
```

```
alertmanagers: []
```

```
## Specify if a Pod Security Policy for node-exporter must be created
```

```
## Ref: https://kubernetes.io/docs/concepts/policy/pod-security-policy/
```

```
##
```

```
podSecurityPolicy:
```

```
  annotations: {}
```

```
    ## Specify pod annotations
```

```
    ## Ref: https://kubernetes.io/docs/concepts/policy/pod-security-policy/#apparmor
```

```
    ## Ref: https://kubernetes.io/docs/concepts/policy/pod-security-policy/#seccomp
```

```
    ## Ref: https://kubernetes.io/docs/concepts/policy/pod-security-policy/#sysctl
```

```
    ##
```

```
    # seccomp.security.alpha.kubernetes.io/allowedProfileNames: '*'
```

```
    # seccomp.security.alpha.kubernetes.io/defaultProfileName: 'docker/default'
```

```
    # apparmor.security.beta.kubernetes.io/defaultProfileName: 'runtime/default'
```

```
alertmanager:
  ## If false, alertmanager will not be installed
  ##
  enabled: true
  service:
    enabled: true
    type: LoadBalancer
  config:
    global:
      slack_api_url: 'https://hooks.slack.com/services/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'

    receivers:
      - name: 'slack'
      - name: 'KongSangUk'
        slack_configs:
          - channel: '#ksu'
            send_resolved: true
            title: '[{{.Status | toUpper}}] {{.CommonLabels.alertname }}'
            text: >-
              | *Description:* {{.CommonAnnotations.description }}

    route:
      group_wait: 10s
      group_interval: 10s
      repeat_interval: 10s
      receiver: 'slack'
      routes:
        - receiver: 'KongSangUk'

  persistence:
    size: 2Gi

podSecurityContext:
  runAsUser: 65534
  runAsNonRoot: true
  runAsGroup: 65534
  fsGroup: 65534
```

Prometheus Query(PromQL)

Prometheus Query(PromQL)

sum

인스턴트 벡터의 모든 시계열 값의 합을 계산

avg

인스턴트 벡터의 모든 시계열 값의 평균을 계산

min

인스턴트 벡터의 모든 시계열 값 중에서 최소값을 반환

max

인스턴트 벡터의 모든 시계열 값 중에서 최대값을 반환

rate

시계열 데이터의 순간 변화율을 계산

irate

rate()와 유사하지만, 마지막 데이터 포인트에서의 변화율을 계산

increase

범위 벡터에서 각 시계열의 증가량을 계산

delta

범위 벡터에서 각 시계열의 변화량을 계산

Metrics for NGINX OSS

Name	Type	Description	Labels
nginx_up	Gauge	Shows the status of the last metric scrape: 1 for a successful scrape and 0 for a failed one	

Stub status metrics

Name	Type	Description	Labels
nginx_connections_accepted	Counter	Accepted client connections.	
nginx_connections_active	Gauge	Active client connections.	
nginx_connections_handled	Counter	Handled client connections.	
nginx_connections_reading	Gauge	Connections where NGINX is reading the request header.	
nginx_connections_waiting	Gauge	Idle client connections.	
nginx_connections_writing	Gauge	Connections where NGINX is writing the response back to the client.	
nginx_http_requests_total	Counter	Total http requests.	

Stub status metrics

Name	Type	Description	Labels
nginx_connections_accepted	Counter	Accepted client connections.	
nginx_connections_active	Gauge	Active client connections.	
nginx_connections_handled	Counter	Handled client connections.	
nginx_connections_reading	Gauge	Connections where NGINX is reading the request header.	
nginx_connections_waiting	Gauge	Idle client connections.	
nginx_connections_writing	Gauge	Connections where NGINX is writing the response back to the client.	
nginx_http_requests_total	Counter	Total http requests.	