

Calico Exporter

OpenStack 내재화

Exported on 08/22/2018

Table of Contents

1	Calico Exporter	4
1.1	설정 방법	4
1.1.1	K8S 설치 시, 아래의 kubescape 설정을 추가한다.....	4
1.1.2	Prometheus 설정.....	4
1.2	Metric 가져오는 법	6
1.3	Metric	11
1.3.1	Felix (calico node agent) Specific Metrics.....	11
1.4	Alert.....	16
1.4.1	Prometheus 차트에 기본으로 설정된 Alert	16
1.4.2	추가 Alert	18
1.5	References	18
2	Bird Exporter	19
2.1	설정 방법	19
2.1.1	Exporter 실행.....	19
2.1.2	Prometheus Scrape 설정	19
2.2	Metric	20
2.3	Alert.....	26

- Calico Exporter (see page 4)
 - 설정 방법 (see page 4)
 - K8S 설치 시, 아래의 kubespray 설정을 추가한다. (see page 4)
 - Prometheus 설정 (see page 4)
 - Metric 가져오는 법 (see page 6)
 - Metric (see page 11)
 - Felix (calico node agent) Specific Metrics (see page 11)
 - Alert (see page 16)
 - Prometheus 차트에 기본으로 설정된 Alert (see page 16)
 - 추가 Alert (see page 18)
 - References (see page 18)
- Bird Exporter (see page 19)
 - 설정 방법 (see page 19)
 - Exporter 실행 (see page 19)
 - Prometheus Scrape 설정 (see page 19)
 - Metric (see page 20)
 - Alert (see page 26)

1 Calico Exporter

1.1 설정 방법

1.1.1 K8S 설치 시, 아래의 kubespray 설정을 추가한다.

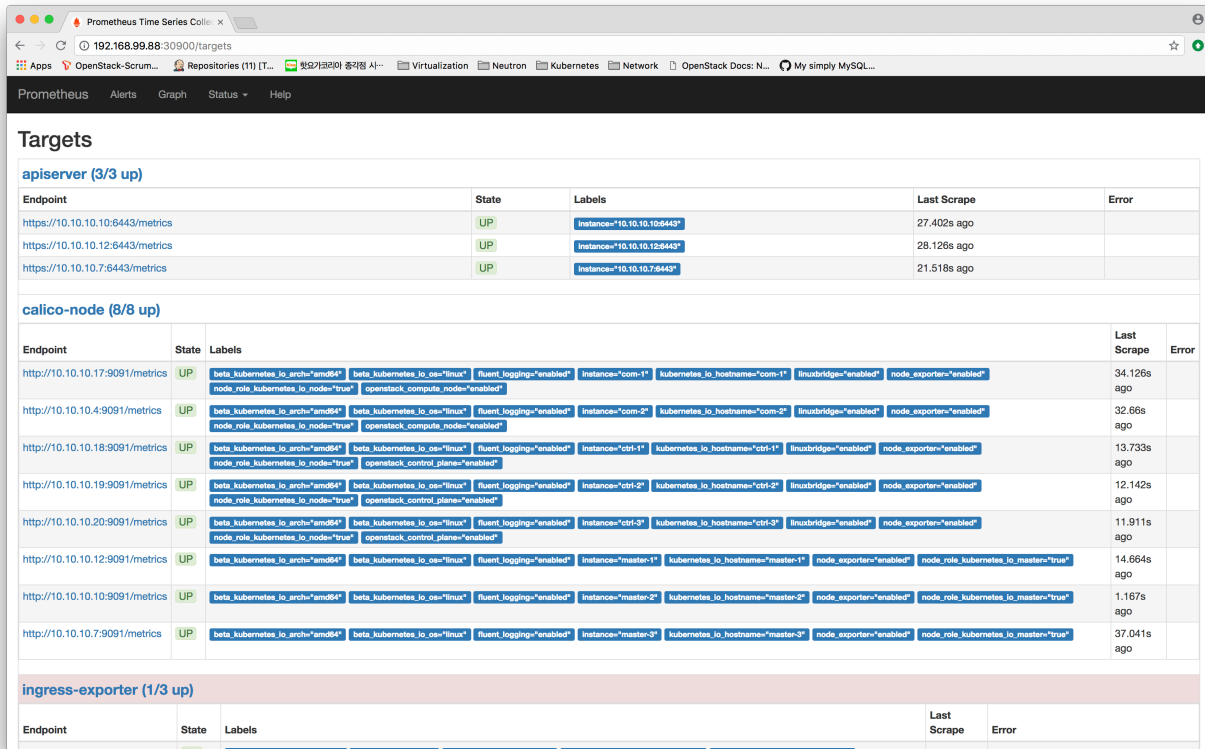
```
# export calico specific metric
calico_felix_prometheusmetricsenabled: "true"
calico_felix_prometheusmetricsport: 9091
# export cpu/memory metric
calico_felix_prometheusgometricsenabled: "true"
calico_felix_prometheusprocessmetricsenabled: "true"
```

1.1.2 Prometheus 설정

OSH-Infra prometheus 차트 배포 시 scrape_configs 리스트에 calico-node를 위한 아래 설정을 추가해 준다.

prometheus value overrides

```
conf:
  prometheus:
    scrape_configs:
      scrape_configs:
        - job_name: calico-node
          kubernetes_sd_configs:
            - role: node
          scrape_interval: 45s
          relabel_configs:
            - action: labelmap
              regex: __meta_kubernetes_node_label_(.+)
            # replace __address__ label with the [cluster node's IP address]:[calico scrape port]
            - source_labels: [__meta_kubernetes_node_address_InternalIP]
              regex: (.*)
              target_label: __address__
              replacement: ${1}:9091
            - source_labels:
                - __meta_kubernetes_node_name
              regex: (.+)
              target_label: __metrics_path__
              replacement: /metrics
            - source_labels:
                - __meta_kubernetes_node_name
              action: replace
              target_label: kubernetes_io_hostname
```



참고)

OSH-Infra의 prometheus 차트는 pod에 설정된 다음의 annotation을 이용해 scrape 타겟을 찾아낸다.

- prometheus.io/scrape¹: scrape on/off 설정
- prometheus.io/scheme²: http or https
- prometheus.io/path³: metric path, 기본 값은 /metric
- prometheus.io/port⁴: metric을 가져올 수 있는 port

OSH-Infra의 calico 차트 또한 아래처럼 calico-node 데몬셋 설정에 위에 나열된 annotation을 포함시키고 있다. 그러나 우리가 사용 중인 kubespray의 calico-node template은 custom annotation을 허용하지 않고 있어 이처럼 kubernetes service discovery config를 prometheus 설정에 추가해 주는 방법을 선택했다.

```

openstack-helm-infra/calico/template/daemonset-calico-node.yaml

{{- $prometheus_annotations := $envAll.Values.monitoring.prometheus.calico_node }}

kind: DaemonSet
apiVersion: apps/v1
metadata:
  name: calico-node
  namespace: kube-system
  
```

- 1 <http://prometheus.io/scrape>
- 2 <http://prometheus.io/scheme>
- 3 <http://prometheus.io/path>
- 4 <http://prometheus.io/port>

```

labels:
  k8s-app: calico-node
spec:
  template:
    metadata:
      annotations:
{{- if .Values.monitoring.prometheus.enabled }}
{{ tuple $prometheus_annotations | include "helm-toolkit.snippets.prometheus_pod_annotations" | indent 8 }}
{{- end }}

```

1.2 Metric 가져오는 법

다음의 명령으로 metric을 수집할 수 있다.

```
$ curl -s http://[node_ip]:9091/metrics
```

metric

```

# HELP felix_active_local_endpoints Number of active endpoints on this host.
# TYPE felix_active_local_endpoints gauge
felix_active_local_endpoints 0
# HELP felix_active_local_policies Number of active policies on this host.
# TYPE felix_active_local_policies gauge
felix_active_local_policies 0
# HELP felix_active_local_selectors Number of active selectors on this host.
# TYPE felix_active_local_selectors gauge
felix_active_local_selectors 0
# HELP felix_active_local_tags Number of active tags on this host.
# TYPE felix_active_local_tags gauge
felix_active_local_tags 0
# HELP felix_calc_graph_output_events Number of events emitted by the calculation graph.
# TYPE felix_calc_graph_output_events counter
felix_calc_graph_output_events 9
# HELP felix_calc_graph_update_time_seconds Seconds to update calculation graph for each datastore OnUpdate
call.
# TYPE felix_calc_graph_update_time_seconds summary
felix_calc_graph_update_time_seconds{quantile="0.5"} NaN
felix_calc_graph_update_time_seconds{quantile="0.9"} NaN
felix_calc_graph_update_time_seconds{quantile="0.99"} NaN
felix_calc_graph_update_time_seconds_sum 0.0007077069999999999
felix_calc_graph_update_time_seconds_count 29
# HELP felix_calc_graph_updates_processed Number of datastore updates processed by the calculation graph.
# TYPE felix_calc_graph_updates_processed counter
felix_calc_graph_updates_processed{type="GlobalConfigKey"} 5
felix_calc_graph_updates_processed{type="HostConfigKey"} 7
felix_calc_graph_updates_processed{type="HostIPKey"} 7
felix_calc_graph_updates_processed{type="IPPoolKey"} 1
felix_calc_graph_updates_processed{type="ProfileLabelsKey"} 1
felix_calc_graph_updates_processed{type="ProfileRulesKey"} 1
felix_calc_graph_updates_processed{type="ProfileTagsKey"} 1

```

```

felix_calc_graph_updates_processed{type="ReadyFlagKey"} 1
felix_calc_graph_updates_processed{type="WorkloadEndpointKey"} 4
# HELP felix_cluster_num_host_endpoints Total number of host endpoints cluster-wide.
# TYPE felix_cluster_num_host_endpoints gauge
felix_cluster_num_host_endpoints 0
# HELP felix_cluster_num_hosts Total number of calico hosts in the cluster.
# TYPE felix_cluster_num_hosts gauge
felix_cluster_num_hosts 7
# HELP felix_cluster_num_workload_endpoints Total number of workload endpoints cluster-wide.
# TYPE felix_cluster_num_workload_endpoints gauge
felix_cluster_num_workload_endpoints 4
# HELP felix_exec_time_micros Summary of time taken to fork/exec child processes
# TYPE felix_exec_time_micros summary
felix_exec_time_micros{quantile="0.5"} 348.517
felix_exec_time_micros{quantile="0.9"} 804.987
felix_exec_time_micros{quantile="0.99"} 1520.949
felix_exec_time_micros_sum 542392.4769999993
felix_exec_time_micros_count 831
# HELP felix_host Configured Felix hostname (as a label), typically used in grouping/aggregating stats; the
label defaults to the hostname of the host but can be overridden by configuration. The value of the gauge
is always set to 1.
# TYPE felix_host gauge
felix_host{host="master-1"} 1
# HELP felix_int_dataplane_addr_msg_batch_size Number of interface address messages processed in each
batch. Higher values indicate we're doing more batching to try to keep up.
# TYPE felix_int_dataplane_addr_msg_batch_size summary
felix_int_dataplane_addr_msg_batch_size{quantile="0.5"} NaN
felix_int_dataplane_addr_msg_batch_size{quantile="0.9"} NaN
felix_int_dataplane_addr_msg_batch_size{quantile="0.99"} NaN
felix_int_dataplane_addr_msg_batch_size_sum 20
felix_int_dataplane_addr_msg_batch_size_count 14
# HELP felix_int_dataplane_apply_time_seconds Time in seconds that it took to apply a dataplane update.
# TYPE felix_int_dataplane_apply_time_seconds summary
felix_int_dataplane_apply_time_seconds{quantile="0.5"} 0.001896493
felix_int_dataplane_apply_time_seconds{quantile="0.9"} 0.003718345
felix_int_dataplane_apply_time_seconds{quantile="0.99"} 0.004805772
felix_int_dataplane_apply_time_seconds_sum 2.9646724240000015
felix_int_dataplane_apply_time_seconds_count 1293
# HELP felix_int_dataplane_failures Number of times dataplane updates failed and will be retried.
# TYPE felix_int_dataplane_failures counter
felix_int_dataplane_failures 0
# HELP felix_int_dataplane_iface_msg_batch_size Number of interface state messages processed in each batch.
Higher values indicate we're doing more batching to try to keep up.
# TYPE felix_int_dataplane_iface_msg_batch_size summary
felix_int_dataplane_iface_msg_batch_size{quantile="0.5"} NaN
felix_int_dataplane_iface_msg_batch_size{quantile="0.9"} NaN
felix_int_dataplane_iface_msg_batch_size{quantile="0.99"} NaN
felix_int_dataplane_iface_msg_batch_size_sum 13
felix_int_dataplane_iface_msg_batch_size_count 10
# HELP felix_int_dataplane_messages Number dataplane messages by type.
# TYPE felix_int_dataplane_messages counter
felix_int_dataplane_messages{type="ConfigUpdate"} 1
felix_int_dataplane_messages{type="HostMetadataUpdate"} 7
felix_int_dataplane_messages{type="IPAMPoolUpdate"} 1

```

```

felix_int_dataplane_messages{type="InSync"} 1
# HELP felix_int_dataplane_msg_batch_size Number of messages processed in each batch. Higher values
indicate we're doing more batching to try to keep up.
# TYPE felix_int_dataplane_msg_batch_size summary
felix_int_dataplane_msg_batch_size{quantile="0.5"} NaN
felix_int_dataplane_msg_batch_size{quantile="0.9"} NaN
felix_int_dataplane_msg_batch_size{quantile="0.99"} NaN
felix_int_dataplane_msg_batch_size_sum 10
felix_int_dataplane_msg_batch_size_count 4
# HELP felix_ipset_calls Number of ipset commands executed.
# TYPE felix_ipset_calls counter
felix_ipset_calls 1
# HELP felix_ipset_errors Number of ipset command failures.
# TYPE felix_ipset_errors counter
felix_ipset_errors 0
# HELP felix_ipset_lines_executed Number of ipset operations executed.
# TYPE felix_ipset_lines_executed counter
felix_ipset_lines_executed 8
# HELP felix_ipsets_calico Number of active Calico IP sets.
# TYPE felix_ipsets_calico gauge
felix_ipsets_calico{ip_version="inet"} 2
# HELP felix_ipsets_total Total number of active IP sets.
# TYPE felix_ipsets_total gauge
felix_ipsets_total 2
# HELP felix_iptables_chains Number of active iptables chains.
# TYPE felix_iptables_chains gauge
felix_iptables_chains{ip_version="4",table="filter"} 10
felix_iptables_chains{ip_version="4",table="mangle"} 3
felix_iptables_chains{ip_version="4",table="nat"} 6
felix_iptables_chains{ip_version="4",table="raw"} 6
# HELP felix_iptables_lines_executed Number of iptables rule updates executed.
# TYPE felix_iptables_lines_executed counter
felix_iptables_lines_executed{ip_version="4",table="filter"} 33
felix_iptables_lines_executed{ip_version="4",table="mangle"} 0
felix_iptables_lines_executed{ip_version="4",table="nat"} 8
felix_iptables_lines_executed{ip_version="4",table="raw"} 0
# HELP felix_iptables_lock_acquire_secs Time in seconds that it took to acquire the iptables lock(s).
# TYPE felix_iptables_lock_acquire_secs summary
felix_iptables_lock_acquire_secs{quantile="0.5"} NaN
felix_iptables_lock_acquire_secs{quantile="0.9"} NaN
felix_iptables_lock_acquire_secs{quantile="0.99"} NaN
felix_iptables_lock_acquire_secs_sum 0
felix_iptables_lock_acquire_secs_count 0
# HELP felix_iptables_lock_retries Number of times the iptables lock was held by someone else and we had to
retry.
# TYPE felix_iptables_lock_retries counter
felix_iptables_lock_retries{version="1.4"} 0
felix_iptables_lock_retries{version="1.6"} 0
# HELP felix_iptables_restore_calls Number of iptables-restore calls.
# TYPE felix_iptables_restore_calls counter
felix_iptables_restore_calls 2
# HELP felix_iptables_restore_errors Number of iptables-restore errors.
# TYPE felix_iptables_restore_errors counter
felix_iptables_restore_errors 0

```



```

# HELP felix_iptables_rules Number of active iptables rules.
# TYPE felix_iptables_rules gauge
felix_iptables_rules{ip_version="4",table="filter"} 33
felix_iptables_rules{ip_version="4",table="mangle"} 8
felix_iptables_rules{ip_version="4",table="nat"} 8
felix_iptables_rules{ip_version="4",table="raw"} 17
# HELP felix_iptables_save_calls Number of iptables-save calls.
# TYPE felix_iptables_save_calls counter
felix_iptables_save_calls 424
# HELP felix_iptables_save_errors Number of iptables-save errors.
# TYPE felix_iptables_save_errors counter
felix_iptables_save_errors 0
# HELP felix_log_errors Number of errors encountered while logging.
# TYPE felix_log_errors counter
felix_log_errors 0
# HELP felix_logs_dropped Number of logs dropped because the output stream was blocked.
# TYPE felix_logs_dropped counter
felix_logs_dropped 0
# HELP felix_resync_state Current datastore state.
# TYPE felix_resync_state gauge
felix_resync_state 3
# HELP felix_resyncs_started Current datastore state.
# TYPE felix_resyncs_started counter
felix_resyncs_started 1
# HELP felix_route_table_list_seconds Time taken to list all the interfaces during a resync.
# TYPE felix_route_table_list_seconds summary
felix_route_table_list_seconds{quantile="0.5"} 0.000136635
felix_route_table_list_seconds{quantile="0.9"} 0.000170352
felix_route_table_list_seconds{quantile="0.99"} 0.000170352
felix_route_table_list_seconds_sum 0.01827725
felix_route_table_list_seconds_count 93
# HELP felix_route_table_per_iface_sync_seconds Time taken to sync each interface
# TYPE felix_route_table_per_iface_sync_seconds summary
felix_route_table_per_iface_sync_seconds{quantile="0.5"} NaN
felix_route_table_per_iface_sync_seconds{quantile="0.9"} NaN
felix_route_table_per_iface_sync_seconds{quantile="0.99"} NaN
felix_route_table_per_iface_sync_seconds_sum 0
felix_route_table_per_iface_sync_seconds_count 0
# HELP go_gc_duration_seconds A summary of the GC invocation durations.
# TYPE go_gc_duration_seconds summary
go_gc_duration_seconds{quantile="0"} 4.4641e-05
go_gc_duration_seconds{quantile="0.25"} 0.000101019
go_gc_duration_seconds{quantile="0.5"} 0.000128754
go_gc_duration_seconds{quantile="0.75"} 0.000180665
go_gc_duration_seconds{quantile="1"} 0.097972091
go_gc_duration_seconds_sum 0.427844401
go_gc_duration_seconds_count 234
# HELP go_goroutines Number of goroutines that currently exist.
# TYPE go_goroutines gauge
go_goroutines 55
# HELP go_info Information about the Go environment.
# TYPE go_info gauge
go_info{version="go1.8.5"} 1
# HELP go_memstats_alloc_bytes Number of bytes allocated and still in use.

```

```

# TYPE go_memstats_alloc_bytes gauge
go_memstats_alloc_bytes 7.208776e+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 2.25703984e+08
# 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 1.484369e+06
# HELP go_memstats_frees_total Total number of frees.
# TYPE go_memstats_frees_total counter
go_memstats_frees_total 1.989655e+06
# HELP go_memstats_gc_cpu_fraction The fraction of this program's available CPU time used by the GC since
the program started.
# TYPE go_memstats_gc_cpu_fraction gauge
go_memstats_gc_cpu_fraction 7.044068265573924e-05
# 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 618496
# 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 7.208776e+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 1.417216e+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 9.068544e+06
# HELP go_memstats_heap_objects Number of allocated objects.
# TYPE go_memstats_heap_objects gauge
go_memstats_heap_objects 39188
# 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 0
# HELP go_memstats_heap_sys_bytes Number of heap bytes obtained from system.
# TYPE go_memstats_heap_sys_bytes gauge
go_memstats_heap_sys_bytes 1.048576e+07
# HELP go_memstats_last_gc_time_seconds Number of seconds since 1970 of last garbage collection.
# TYPE go_memstats_last_gc_time_seconds gauge
go_memstats_last_gc_time_seconds 1.5337042089640448e+09
# HELP go_memstats_lookups_total Total number of pointer lookups.
# TYPE go_memstats_lookups_total counter
go_memstats_lookups_total 17940
# HELP go_memstats_mallocs_total Total number of mallocs.
# TYPE go_memstats_mallocs_total counter
go_memstats_mallocs_total 2.028843e+06
# HELP go_memstats_mcache_inuse_bytes Number of bytes in use by mcache structures.
# TYPE go_memstats_mcache_inuse_bytes gauge
go_memstats_mcache_inuse_bytes 4800
# HELP go_memstats_mcache_sys_bytes Number of bytes used for mcache structures obtained from system.
# TYPE go_memstats_mcache_sys_bytes gauge
go_memstats_mcache_sys_bytes 16384
# HELP go_memstats_mspan_inuse_bytes Number of bytes in use by mspan structures.
# TYPE go_memstats_mspan_inuse_bytes gauge
go_memstats_mspan_inuse_bytes 124032

```

```

# HELP go_memstats_mspan_sys_bytes Number of bytes used for mspan structures obtained from system.
# TYPE go_memstats_mspan_sys_bytes gauge
go_memstats_mspan_sys_bytes 147456
# HELP go_memstats_next_gc_bytes Number of heap bytes when next garbage collection will take place.
# TYPE go_memstats_next_gc_bytes gauge
go_memstats_next_gc_bytes 7.79663e+06
# HELP go_memstats_other_sys_bytes Number of bytes used for other system allocations.
# TYPE go_memstats_other_sys_bytes gauge
go_memstats_other_sys_bytes 1.118887e+06
# HELP go_memstats_stack_inuse_bytes Number of bytes in use by the stack allocator.
# TYPE go_memstats_stack_inuse_bytes gauge
go_memstats_stack_inuse_bytes 1.048576e+06
# HELP go_memstats_stack_sys_bytes Number of bytes obtained from system for stack allocator.
# TYPE go_memstats_stack_sys_bytes gauge
go_memstats_stack_sys_bytes 1.048576e+06
# HELP go_memstats_sys_bytes Number of bytes obtained from system.
# TYPE go_memstats_sys_bytes gauge
go_memstats_sys_bytes 1.4919928e+07
# HELP go_threads Number of OS threads created.
# TYPE go_threads gauge
go_threads 14
# HELP process_cpu_seconds_total Total user and system CPU time spent in seconds.
# TYPE process_cpu_seconds_total counter
process_cpu_seconds_total 107.94
# HELP process_max_fds Maximum number of open file descriptors.
# TYPE process_max_fds gauge
process_max_fds 1024
# HELP process_open_fds Number of open file descriptors.
# TYPE process_open_fds gauge
process_open_fds 14
# HELP process_resident_memory_bytes Resident memory size in bytes.
# TYPE process_resident_memory_bytes gauge
process_resident_memory_bytes 2.260992e+07
# HELP process_start_time_seconds Start time of the process since unix epoch in seconds.
# TYPE process_start_time_seconds gauge
process_start_time_seconds 1.53369550289e+09
# HELP process_virtual_memory_bytes Virtual memory size in bytes.
# TYPE process_virtual_memory_bytes gauge
process_virtual_memory_bytes 4.6350336e+07

```

1.3 Metric

1.3.1 Felix (calico node agent) Specific Metrics

* 빨간색: 알람이 설정된 metric

Metric	Type	Description	Note
felix_active_local_endpoints	gauge	Number of active endpoints on this host.	로컬 노드에 존재하는 workload endpoint("cali"로 시작하는 인터페이스)의 수. host network를 사용하지 않는 pod의 수와 동일하다.
felix_active_local_policies	gauge	Number of active policies on this host.	Policy 리소스의 수.
felix_active_local_selectors	gauge	Number of active selectors on this host.	정의된 selector의 수. 참고로 selector는 특정 policy를 적용할 endpoint를 지정하는 데 사용됨.
felix_active_local_tags	gauge	Number of active tags on this host.	정의된 tag의 수. 참고로 tag는 특정 profile을 적용할 endpoint를 지정하는 데 사용가능하며, label로 대체가능하여 deprecated 됨.
felix_calc_graph_output_events	counter	Number of events emitted by the calculation graph.	Calculation graph가 발생시킨 메시지 수. Calc graph는 data model의 변화를 감지해 실제 노드의 네트워크 상태를 업데이트하는데 필요한 정보를 포함한 메시지를 생성하는 역할을 함.
felix_route_table_per_iface_sync_seconds	summary	Seconds to update calculation graph for each datastore OnUpdate call.	Datastore OnUpdate 이벤트 발생 시, calc graph를 업데이트하는데 걸린 시간.
felix_calc_graph_updates_processed	counter	Number of datastore updates processed by the calculation graph.	WorkloadEndpointKey, HostEndpointKey와 같은 datastore key 별 업데이트 처리 횟수.
felix_cluster_num_host_endpoints	gauge	Total number of host endpoints cluster-wide.	Calico에 등록된 hostEndpoint 리소스의 수. 보통 보안 정책 적용을 위해 노드의 물리 인터페이스를 hostEndpoint로 등록한다.
felix_cluster_num_hosts	gauge	Total number of calico hosts in the cluster.	클러스터 내 노드의 수. 노드의 상태와는 관계가 없다.
felix_cluster_num_workload_endpoints	gauge	Total number of workload endpoints cluster-wide.	클러스터 내 전체 cali* 인터페이스의 수. hostNetwork를 사용하지 않는 전체 pod의 수와 동일함.

Metric	Type	Description	Note
felix_exec_time_micros	summary	Summary of time taken to fork/exec child processes.	노드의 네트워크 상태 업데이트에 필요한 iptables, ipset, route 등의 명령 실행 혹은 child process fork에 걸린 시간.
felix_host	gauge	Configured Felix hostname (as a label), typically used in grouping/aggregating stats; the label defaults to the hostname of the host but can be overridden by configuration. The value of the gauge is always set to 1.	호스트네임. 값은 항상 1.
felix_int_data_plane_addr_msg_batch_size	summary	Number of interface address messages processed in each batch. Higher values indicate we're doing more batching to try to keep up.	한 batch 사이클 동안 처리한 interface address 업데이트 메시지 수. 단, "cali"로 시작하는 인터페이스 외에는 무시함. 참고로, WorkloadEndpoint 생성/삭제 시 발생하는 메시지임.
felix_int_data_plane_apply_time_seconds	summary	Time in seconds that it took to apply a dataplane update.	Dataplane 업데이트를 적용하는데 걸린 시간.
felix_int_data_plane_failures	counter	Number of times dataplane updates failed and will be retried.	Dataplane 업데이트 처리 실패 횟수.
felix_int_data_plane_iface_msg_batch_size	summary	Number of interface state messages processed in each batch. Higher values indicate we're doing more batching to try to keep up.	한 batch 사이클 동안 처리한 interface 상태 업데이트 메시지 수. 단, "cali"로 시작하는 인터페이스 외에는 무시함. 참고로, WorkloadEndpoint 생성/삭제 시 발생하는 메시지임.
felix_int_data_plane_messages	counter	Number dataplane messages by type.	Dataplane 메시지의 종류 별 발생 횟수.

Metric	Type	Description	Note
felix_int_data_plane_msg_batch_size	summary	Number of message processed in each batch. Higher values indicate we're doing more batching to try to keep up.	한 batch 사이클 동안 처리한 dataplane 메시지 수. 참고로 dataplane 메시지는 calc graph가 발생시킴. 한번에 많은 k8s 리소스 업데이트가 일어난 경우 이 값이 클 수 있다.
felix_ipset_calls	counter	Number of ipset commands executed.	ipset restore 명령 실행 횟수.
felix_ipset_errors	counter	Number of ipset command failures.	ipset 명령 실패 횟수.
felix_ipset_lines_executed	counter	Number of ipset operations executed.	ipset restore 시에 실제 수행 된 ipset operation의 수.
felix_ipsets_calico	gauge	Number of active Calico IP sets.	ipset의 수. 단, "cali"로 시작하는 ipset의 수만 카운트함.
felix_ipsets_total	gauge	Total number of active IP sets.	로컬 노드에 존재하는 전체 ipset의 수.
felix_iptables_chains	gauge	Number of active iptables chains.	iptables chain의 수. 단, "cali-"로 시작하는 chain의 수만 카운트 함.
felix_iptables_lines_executed	counter	Number of iptables rule updates executed.	iptables rule update 실행 횟수.
felix_iptables_lock_acquire_secs	summary	Time in seconds that it took to acquire the iptables lock(s).	iptables lock을 획득하는데 걸린 시간.

Metric	Type	Description	Note
felix_iptables_lock_retries	counter	Number of times the iptables lock was held by someone else and we had to retry.	iptables lock이 걸린 횟수.
felix_iptables_restore_calls	counter	Number of iptables-restore calls.	iptables-restore 명령 실행 횟수.
felix_iptables_restore_errors	counter	Number of iptables-restore errors.	iptables-restore 명령 실행 실패 횟수.
felix_iptables_rules	gauge	Number of active iptables rules.	iptables rule 수. 단, "cali-"로 시작하는 rule의 수만 카운트함.
felix_iptables_save_calls	counter	Number of iptables-save calls.	iptables-save 실행 횟수.
felix_iptables_save_errors	counter	Number of iptables-save errors.	iptables-save 명령 실행 실패 횟수.
felix_log_errors	counter	Number of errors encountered while logging.	로깅 에러
felix_logs_dropped	counter	Number of logs dropped because the output stream was blocked.	로깅 에러
felix_resync_state	gauge	Current datastore state.	Datstore sync 상태. <ul style="list-style-type: none"> • WaitForDatastore: 1 • ResyncInProgress: 2 • InSync: 3

Metric	Type	Description	Note
felix_resyncs_started	counter	Current datastore state.	Datastore sync 상태. <ul style="list-style-type: none"> • WaitForDatastore: 1 • ResyncInProgress: 2 • InSync: 3
felix_route_table_list_seconds	summary	Time taken to list all the interfaces during a resync.	노드의 인터페이스 목록 조회에 걸린 시간.
felix_route_table_per_iface_sync_seconds	summary	Time taken to sync each interface.	인터페이스 별 라우팅 테이블 업데이트에 걸린 시간.

1.4 Alert

1.4.1 Prometheus 차트에 기본으로 설정된 Alert

Alert	Expr	Duration	Summary
calico_dataplane_failures_high_1h	absent(felix_int_dataplane_failures) OR increase(felix_int_dataplane_failures[1h]) > 5		A high number of dataplane failures within Felix are happening 최근 1시간 동안 dataplane failure가 5회 이상 발생하면 알람.

Alert	Expr	Duration	Summary
calico_dataplane_address_msg_batch_size_high_5m	absent(felix_int_dataplane_addr_msg_batch_size_sum) OR absent(felix_int_dataplane_addr_msg_batch_size_count) OR (felix_int_dataplane_addr_msg_batch_size_sum / felix_int_dataplane_addr_msg_batch_size_count) > 5	5m	Felix address message batch size is higher 5분 동안 한 batch 사이클 당 인터페이스 업데이트가 5회 이상 발생하면 알람.
calico_dataplane_iface_msg_batch_size_high_5m	absent(felix_int_dataplane_iface_msg_batch_size_sum) OR absent(felix_int_dataplane_iface_msg_batch_size_count) OR (felix_int_dataplane_iface_msg_batch_size_sum / felix_int_dataplane_iface_msg_batch_size_count) > 5	5m	Felix interface message batch size is higher 5분 동안 한 batch 사이클 당 인터페이스 업데이트가 평균 5회 이상 발생하면 알람.
calico_ipset_errors_high_1h	absent(felix_ipset_errors) OR increase(felix_ipset_errors[1h]) > 5		A high number of ipset errors within Felix are happening 최근 1시간 동안 ipset 명령 실행 오류가 5회 이상 발생하면 알람.
calico iptable_save_errors_high_1h	absent(felix iptables_save_errors) OR increase(felix iptables_save_errors[1h]) > 5		A high number of iptable save errors within Felix are happening 최근 1시간 동안 iptables-save 명령 실행 오류가 5회 이상 발생하면 알람.
calico iptable_restore_errors_high_1h	absent(felix iptables_restore_errors) OR increase(felix iptables_restore_errors[1h]) > 5		A high number of iptable restore errors within Felix are happening 최근 1시간 동안 iptables-restore 명령 실행 오류가 5회 이상 발생하면 알람.

1.4.2 추가 Alert

Alert	Expr	Duration	Summary
calico_node_down_5m	up{job="calico-node"} == 0	5m	Felix has been down for more than 5 minutes calico-node가 5분 이상 down 상태면 알람.

```

conf:
  prometheus:
    rules:
      calico:
        groups:
          - name: calico.rules
            rules:
              - alert: calico_node_down_5m
                expr: up{job="calico-node"} == 0
                for: 5m
                labels:
                  severity: critical
                annotations:
                  description: 'Felix on {{ $labels.instance }} has seen down for more than 5 minutes'
                  summary: 'Felix is down'

```

1.5 References

<https://prometheus.io/docs/introduction/overview/>

<https://en.fabernovel.com/insights/tech-en/alerting-in-prometheus-or-how-i-can-sleep-well-at-night>

https://prometheus.io/docs/prometheus/latest/configuration/configuration/#%3Cscrape_config%3E

https://prometheus.io/docs/prometheus/latest/configuration/configuration/#%3Crelabelfilter_config%3E

https://prometheus.io/docs/prometheus/latest/configuration/configuration/#%3Ckubernetes_sd_configs%3E

<https://docs.projectcalico.org/v3.0/getting-started/bare-metal/bare-metal>

<https://www.tigera.io/securing-host-endpoints-with-project-calico-2591113ab99c/>

<https://github.com/projectcalico/felix/tree/master/k8sfv>

<https://prometheus.io/docs/prometheus/latest/querying/functions/>

<https://docs.projectcalico.org/v2.6/reference/calicoctl/resources/profile#nav-calicoctl-resourcedefinitions>

2 Bird Exporter

2.1 설정 방법

2.1.1 Exporter 실행

```

apiVersion: extensions/v1beta1
kind: DaemonSet
metadata:
  name: calico-bird-exporter
  namespace: kube-system
  labels:
    app: calico-bird-exporter
spec:
  selector:
    matchLabels:
      app: calico-bird-exporter
  template:
    metadata:
      labels:
        app: calico-bird-exporter
    spec:
      restartPolicy: "Always"
      hostNetwork: true
      tolerations:
        - effect: NoSchedule
          operator: Exists
      containers:
        - name: bird-exporter
          image: tacorepo:5000/birdwatcher:latest
          volumeMounts:
            - name: var-run-calico
              mountPath: /var/run/calico
      volumes:
        - name: var-run-calico
          hostPath:
            path: /var/run/calico

```

2.1.2 Prometheus Scrape 설정

```

conf:
  prometheus:
    scrape_configs:
      scrape_configs:
        - job_name: bird-exporter

```

```

kubernetes_sd_configs:
- role: node
scrape_interval: 45s
relabel_configs:
- action: labelmap
  regex: __meta_kubernetes_node_label_(.+)
# replace __address__ label with the [cluster node's IP address]:[calico scrape port]
- source_labels: [__meta_kubernetes_node_address_InternalIP]
  regex: (.*)
  target_label: __address__
  replacement: ${1}:6502
- source_labels:
  - __meta_kubernetes_node_name
  regex: (.+)
  target_label: __metrics_path__
  replacement: /
- source_labels:
  - __meta_kubernetes_node_name
  action: replace
  target_label: kubernetes_io_hostname

```

2.2 Metric

다음의 명령으로 metric을 수집할 수 있다.

```
$ curl -s http://[node_ip]:6502
```

bird metrics

```

bird_up{bird_protocol_instance="kernel1"} 1
bird_preference{bird_protocol_instance="kernel1"} 10
bird_routes{bird_protocol_instance="kernel1",bird_route_type="imported"} 4
bird_routes{bird_protocol_instance="kernel1",bird_route_type="exported"} 18
bird_routes{bird_protocol_instance="kernel1",bird_route_type="preferred"} 3
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="import",bird_action="updates",bird_outcome="received"} 4
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="import",bird_action="updates",bird_outcome="rejected"} 0
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="import",bird_action="updates",bird_outcome="filtered"} 0
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="import",bird_action="updates",bird_outcome="ignored"} 0
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="import",bird_action="updates",bird_outcome="accepted"} 4
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="import",bird_action="withdraws",bird_outcome="received"} 0
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="import",bird_action="withdraws",bird_outcome="rejected"} 0
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="import",bird_action="withdraws",bird_outcome="ignored"} 0

```

```

bird_route_changes{bird_protocol_instance="kernel1",bird_direction="import",bird_action="withdraws",bird_outcome="accepted"} 0
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="export",bird_action="updates",bird_outcome="received"} 25
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="export",bird_action="updates",bird_outcome="rejected"} 6
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="export",bird_action="updates",bird_outcome="filtered"} 0
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="export",bird_action="updates",bird_outcome="accepted"} 19
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="export",bird_action="withdraws",bird_outcome="received"} 1
bird_route_changes{bird_protocol_instance="kernel1",bird_direction="export",bird_action="withdraws",bird_outcome="accepted"} 1
bird_up{bird_protocol_instance="device1"} 1
bird_preference{bird_protocol_instance="device1"} 240
bird_routes{bird_protocol_instance="device1",bird_route_type="imported"} 0
bird_routes{bird_protocol_instance="device1",bird_route_type="exported"} 0
bird_routes{bird_protocol_instance="device1",bird_route_type="preferred"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="import",bird_action="updates",bird_outcome="received"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="import",bird_action="updates",bird_outcome="rejected"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="import",bird_action="updates",bird_outcome="filtered"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="import",bird_action="updates",bird_outcome="ignored"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="import",bird_action="updates",bird_outcome="accepted"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="import",bird_action="withdraws",bird_outcome="received"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="import",bird_action="withdraws",bird_outcome="rejected"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="import",bird_action="withdraws",bird_outcome="ignored"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="import",bird_action="withdraws",bird_outcome="accepted"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="export",bird_action="updates",bird_outcome="received"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="export",bird_action="updates",bird_outcome="rejected"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="export",bird_action="updates",bird_outcome="filtered"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="export",bird_action="updates",bird_outcome="accepted"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="export",bird_action="withdraws",bird_outcome="received"} 0
bird_route_changes{bird_protocol_instance="device1",bird_direction="export",bird_action="withdraws",bird_outcome="accepted"} 0
bird_up{bird_protocol_instance="direct1"} 1
bird_preference{bird_protocol_instance="direct1"} 240
bird_routes{bird_protocol_instance="direct1",bird_route_type="imported"} 3
bird_routes{bird_protocol_instance="direct1",bird_route_type="exported"} 0
bird_routes{bird_protocol_instance="direct1",bird_route_type="preferred"} 3

```

```

bird_route_changes{bird_protocol_instance="direct1",bird_direction="import",bird_action="updates",bird_outcome="received"} 3
bird_route_changes{bird_protocol_instance="direct1",bird_direction="import",bird_action="updates",bird_outcome="rejected"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="import",bird_action="updates",bird_outcome="filtered"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="import",bird_action="updates",bird_outcome="ignored"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="import",bird_action="updates",bird_outcome="accepted"} 3
bird_route_changes{bird_protocol_instance="direct1",bird_direction="import",bird_action="withdraws",bird_outcome="received"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="import",bird_action="withdraws",bird_outcome="rejected"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="import",bird_action="withdraws",bird_outcome="ignored"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="import",bird_action="withdraws",bird_outcome="accepted"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="export",bird_action="updates",bird_outcome="received"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="export",bird_action="updates",bird_outcome="rejected"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="export",bird_action="updates",bird_outcome="filtered"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="export",bird_action="updates",bird_outcome="accepted"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="export",bird_action="withdraws",bird_outcome="received"} 0
bird_route_changes{bird_protocol_instance="direct1",bird_direction="export",bird_action="withdraws",bird_outcome="accepted"} 0
bird_up{bird_protocol_instance="Node_192_168_97_2"} 1
bird_preference{bird_protocol_instance="Node_192_168_97_2"} 100
bird_routes{bird_protocol_instance="Node_192_168_97_2",bird_route_type="imported"} 18
bird_routes{bird_protocol_instance="Node_192_168_97_2",bird_route_type="exported"} 2
bird_routes{bird_protocol_instance="Node_192_168_97_2",bird_route_type="preferred"} 17
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="import",bird_action="updates",bird_outcome="received"} 19
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="import",bird_action="updates",bird_outcome="rejected"} 0
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="import",bird_action="updates",bird_outcome="filtered"} 0
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="import",bird_action="updates",bird_outcome="ignored"} 0
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="import",bird_action="updates",bird_outcome="accepted"} 19
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="import",bird_action="withdraws",bird_outcome="received"} 1
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="import",bird_action="withdraws",bird_outcome="rejected"} 0
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="import",bird_action="withdraws",bird_outcome="ignored"} 0
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="import",bird_action="withdraws",bird_outcome="accepted"} 1

```

```

bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="export",bird_action="updates",
bird_outcome="received"} 47
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="export",bird_action="updates",
bird_outcome="rejected"} 35
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="export",bird_action="updates",
bird_outcome="filtered"} 10
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="export",bird_action="updates",
bird_outcome="accepted"} 2
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="export",bird_action="withdraw
s",bird_outcome="received"} 1
bird_route_changes{bird_protocol_instance="Node_192_168_97_2",bird_direction="export",bird_action="withdraw
s",bird_outcome="accepted"} 21
bird_bgp_state{bird_protocol_instance="Node_192_168_97_2",bird_bgp_state="Established"} 1
bird_hold_timer_current{bird_protocol_instance="Node_192_168_97_2"} 95
bird_hold_timer_initial{bird_protocol_instance="Node_192_168_97_2"} 180
bird_keepalive_timer_current{bird_protocol_instance="Node_192_168_97_2"} 53
bird_keepalive_timer_initial{bird_protocol_instance="Node_192_168_97_2"} 60
bird_up{bird_protocol_instance="static1"} 1
bird_preference{bird_protocol_instance="static1"} 200
bird_routes{bird_protocol_instance="static1",bird_route_type="imported"} 1
bird_routes{bird_protocol_instance="static1",bird_route_type="exported"} 0
bird_routes{bird_protocol_instance="static1",bird_route_type="preferred"} 1
bird_route_changes{bird_protocol_instance="static1",bird_direction="import",bird_action="updates",bird_outc
ome="received"} 1
bird_route_changes{bird_protocol_instance="static1",bird_direction="import",bird_action="updates",bird_outc
ome="rejected"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="import",bird_action="updates",bird_outc
ome="filtered"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="import",bird_action="updates",bird_outc
ome="ignored"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="import",bird_action="updates",bird_outc
ome="accepted"} 1
bird_route_changes{bird_protocol_instance="static1",bird_direction="import",bird_action="withdraws",bird_ou
tcome="received"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="import",bird_action="withdraws",bird_ou
tcome="rejected"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="import",bird_action="withdraws",bird_ou
tcome="ignored"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="import",bird_action="withdraws",bird_ou
tcome="accepted"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="export",bird_action="updates",bird_outc
ome="received"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="export",bird_action="updates",bird_outc
ome="rejected"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="export",bird_action="updates",bird_outc
ome="filtered"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="export",bird_action="updates",bird_outc
ome="accepted"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="export",bird_action="withdraws",bird_ou
tcome="received"} 0
bird_route_changes{bird_protocol_instance="static1",bird_direction="export",bird_action="withdraws",bird_ou
tcome="accepted"} 0

```

Metric	Type	D i m e n s i o n s	Description
bird_ up	gau ge	b i r d - p r o t o c o l - i n s t a n c e	birdc show protocol all 결과에 포함된 모든 항목에 대해, state가 "up"인지 확인하여 up 이면 1을 반환. 결과에는 기본적으로 Kernel, Device, Direct, Static 라우팅 프로토콜이 포함되어 있고, calico 설정에 따라 다른 calico node들과 route reflector들이 BGP로 추가된다.

Metric	Type	D i m e n s i o n s	Description
bird_bgp_state	gauge	bird-protocol-instance bird-bgp-state	birdc show protocol all 결과 중 BGP 프로토콜로 연결된 항목에 대해 BGP state 값을 "bird_bgp_state"에 넣어준다. 게이지 값은 항상 1이다.

2.3 Alert

Alert	Expr	Duration	Summary
bird_bgp_state_inactive_3m	bird_bgp_state{bird_bgp_state!="Established"} == 1	3m	BGP 연결이 끊어진 상태가 3분 이상 지속되면 알람. RR을 사용하는 경우에 설정.

```

conf:
  prometheus:
    rules:
      bird:
        groups:
          - name: bird.rules
            rules:
              - alert: bird_bgp_state_inactive_3m
                expr: bird_bgp_state{bird_bgp_state!="Established"} == 1
                for: 3m
                labels:
                  severity: warn
                annotations:
                  description: '{{ $labels.instance }} has inactive BGP connection to
                    {{ $labels.bird_protocol_instance }} for more than 3 minutes'
                  summary: 'BGP connection fails'

```