

## vmha\_worker

### 功能需求

- hypervisor 宕机：由prometheus-alertmanager端触发告警，vmha\_worker接收告警并进行evacuate操作
- hypervisor 业务网口全面故障：由prometheus-alertmanager端触发告警，vmha\_worker不做处理
- VIM 存储网口全面故障：由prometheus-alertmanager端触发告警，vmha\_worker接收告警并进行nova live-migration操作

### 设计方案

#### 一、prometheus rules

##### hypervisor 宕机rules

```
- name: hypervisor_error
  rules:
    - alert: hypervisor_error
      expr: probe_success{type = 'Hypervisor'} == 0
      for: 3s
      labels:
        severity: Critical
      annotations:
        description: >-
          hypervisor {{ $labels.hostname }} error.
        summary: >-
          hypervisor {{ $labels.hostname }} error.
        description_template: >-
          the hypervisor %(hostname)s is error.
        summary_template: >-
          the hypervisor %(hostname)s is error.
        suggestion: >-
          if the hypervisor {{ $labels.hostname }} is down, please check
the evacuate progress.
        suggestion_template: >-
          if the hypervisor %(hostname)s is down, please check the
evacuate progress.
        resource_type: physical.node
        sub_resource_type: system
        alert_target: "{{ $labels.hostname }}"
```

##### hypervisor 业务网口全面故障

```
- name: business_link
  rules:
    - alert: business_link_interruption
      expr: >
```

```

        # eth0,eth1 are business-device members
        sum(node_network_up{interface=~'eth0|eth1'}) by (hostname) == 0
for: 1m
labels:
    severity: Critical
annotations:
    description: >-
        the business network link of {{ $labels.hostname }}
interruption.
    summary: >-
        the business network link of {{ $labels.hostname }}
interruption.
    description_template: >-
        the business network link of %(hostname)s interruption.
    summary_template: >-
        the business network link of %(hostname)s interruption.
    suggestion: >-
        Please check the link of the {{ $labels.hostname }}, may be
one of the
        business interfaces of the {{ $labels.hostname }} host is
down.
    suggestion_template: >-
        Please check the link of the %(hostname)s, may be one of the
[[ network_name ]]
        interfaces on the %(hostname)s is down.
    resource_type: physical.network
    sub_resource_type: link
    alert_target: "{{ $labels.hostname }}"

```

## VIM 存储网口全面故障

```

- name: storage_link
  rules:
    - alert: storage_link_interruption
      expr: >
        # eth0,eth1 are storage-device members
        sum(node_network_up{interface=~'eth0|eth1'}) by (hostname) == 0
for: 1m
labels:
    severity: Critical
annotations:
    description: >-
        the storage network link of {{ $labels.hostname }}
interruption.
    summary: >-
        the storage network link of {{ $labels.hostname }}
interruption.
    description_template: >-
        the storage network link of %(hostname)s interruption.
    summary_template: >-
        the storage network link of %(hostname)s interruption.
    suggestion: >-

```

```

    Please check the link of the {{ $labels.hostname }}, may be
one of the
    storage interfaces of the {{ $labels.hostname }} host is down.
suggestion_template: >-
    Please check the link of the %(hostname)s, may be one of the
[[ network_name ]]
    interfaces on the %(hostname)s is down.
resource_type: physical.network
sub_resource_type: link
alert_target: "{{ $labels.hostname }}"

```

## 二、alertmanager inhibit\_rules

由于hypervisor宕机不仅会触发它本身的告警，还会触发由hypervisor宕机引起的其他告警，比如存储/管理/业务网络故障，此时需要抑制后者告警，抑制的作用有两方面：1. 接受者不会接收到大量告警，有利于快速的定位故障发生的位置。2. 便于针对由hypervisor宕机引起的其他告警做出相应的action

### 抑制由于hypervisor宕机产生的其他告警

```

- source_match:
  severity: critical
  target_match:
    severity: Major
  equal:
    - alertname
    - instance
- source_match:
  alertname: hypervisor_error
  severity: critical
  target_match:
    severity: critical
  equal:
    - hostname

```

## 三、vmha\_worker 处理

根据alertname和severity标识一个特定的告警，进而做出action

```

alertname == 'hypervisor_error' and severity == 'critical' ==>
evacuate

alertname == 'storage_link' and severity == 'critical' ==> live-
migration

```

### evacuate风险:

1. 当集群中大量的 hypervisor 宕机时，剩余资源不足，此时去做 evacuate 会有很大风险。

解决方案1: 设置max\_evacuate\_host\_num阈值, 已疏散过的hypervisor超过 max\_evacuate\_host\_num 抛出告警;

解决方案2: 统计不可用的hypervisor个数failed\_hypervisor\_num, 超过总数的1/3则不做evacuate操作, 并抛出告警

## 2. check evacuate has failed ?

方案1: 检查vm状态是否active

方案2: 检查evacuate返回指, 判断request是否被accept

## 接收alertmanager抛出的告警信息, 做出action

```
alerts_items = json.loads(request.data).get('alerts')
for alert in alerts_items:
    labels = alert['labels']
    if labels['alertname'] == 'host_down':
        # evacuate
        # 根据service.state统计failed_hypervisor_num, 超过三分之一hypervisor
        # down 不做evacuate并触发hypervisor cluster warning
        if is_evacuate():
            # alert hypervisor cluster warning
            # break
            pass
        evacuate()
    elif labels['alertname'] == 'storage_net_down':
        # live-migration
        live-migration()
```

## evacuate

```
hypervisors = novaclient.hypervisors.search(labels['hostname'])
for hypervisor in hypervisors:
    for server in hypervisor.servers:
        evacuate_server(node, server)
```

## 根据service.state统计failed\_hypervisor\_num, 更新计算节点集群状态

```
def is_evacuate(self):
    failed_hypervisor_num = 0
    all_hypervisor_num = 0
    for service in self.novaclient.services.list():
        if service.binary == u'nova-compute':
            if service.state != 'up':
                failed_hypervisor_num += 1
            all_hypervisor_num += 1
    if failed_hypervisor_num*3 > all_hypervisor_num:
```

```

    return True
else:
    return False

```

### alertmanager抛出的告警格式

```

{
  "receiver": "admin",
  "status": "resolved",
  "alerts": [
    {
      "status": "resolved",
      "labels": {
        "alertname": "InstanceIsDown",
        "instance": "localhost:9100",
        "severity": "critical",
        "job": "node"
      },
      "annotations": {
        "description": "localhost:9100 of job node has been down for more
than 5 minutes.\n",
        "summary": "Instance localhost:9100 down"
      },
      "startsAt": "2019-02-28T14:12:31.142457796+08:00",
      "endsAt": "2019-02-28T14:15:31.142457796+08:00",
      "generatorURL": "http://t470p:9090/graph?
g0.expr=up%7Bjob%3D%22node%22%7D+%3D%3D+0&g0.tab=1"
    }
  ],
  "groupLabels": {
    "instance": "localhost:9100"
  },
  "commonLabels": {
    "alertname": "InstanceIsDown",
    "instance": "localhost:9100",
    "job": "node"
  },
  "commonAnnotations": {
    "description": "localhost:9100 of job node has been down for more than
5 minutes.\n",
    "summary": "Instance localhost:9100 down"
  },
  "externalURL": "http://t470p:9093",
  "version": "4",
  "groupKey": "{}:{instance=\"localhost:9100\"}"
}

```