Cloud Native Batch Computing Platform Volcano

Yang Wang, Contributor of Volcano Community

Volcano 安装部署

1. 安装k8s集群:

- git clone git@github.com:volcano-sh/volcano.git
- kind create cluster --name my-k8s -config hack/e2e-kind-config.yaml
- Kind的安装与使用参见: https://kind.sigs.k8s.io/

2. 安装Volcano:

使用helm安装:

- 添加Volcano helm仓: helm repo add volcano-sh https://volcano-sh.github.io/helm-charts
- 安装Volcano: helm install volcano volcano-sh/volcano -n volcano-system --create-namespace

通过yaml安装:

kubectl apply -f https://raw.githubusercontent.com/volcano-sh/volcano/master/installer/volcano-development.yaml









02 运行Pytorch作业

03 运行Spark作业







TensorFlow 作业配置

apiVersion: batch.volcano.sh/v1alpha1 kind: Job metadata: name: tensorflow-benchmark _labels: "volcano.sh/job-type": "Tensorflow" 根据业务需求设置作业类型 spec: minAvailable: 3 schedulerName: volcano plugins: 开启SSH免密码认证 env: [] svc: [] 用于创建headless service解决ps和worker之间的通信问题 policies: - event: PodEvicted 配置作业重启策略 action: RestartJob



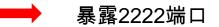




TensorFlow 作业配置

```
tasks:
  - replicas: 1
   name: ps
   template:
    spec:
     imagePullSecrets:
       - name: default-secret
     containers:
       - command:
         - sh
          PS_HOST=`cat /etc/volcano/ps.host | sed 's/$/&:2222/g' | tr "\n" ","`;
          WORKER_HOST=`cat /etc/volcano/worker.host | sed 's/$/&:2222/g' | tr "\n" ","`;
          python tf_cnn_benchmarks.py --batch_size=32 --model=resnet50 --
variable_update=parameter_server --flush_stdout=true --num_gpus=1 --
local_parameter_device=cpu --device=cpu --data_format=NHWC --job_name=ps --
task_index=${VK_TASK_INDEX} --ps_hosts=${PS_HOST} --worker_hosts=${WORKER_HOST}
        image: volcanosh/example-tf:0.0.1
        name: tensorflow
        ports:
         - containerPort: 2222
          name: tfjob-port
        workingDir: /opt/tf-benchmarks/scripts/tf_cnn_benchmarks
      restartPolicy: OnFailure
```

PS的运行配置









TensorFlow 作业配置

restartPolicy: OnFailure

```
- replicas: 2
   name: worker
   policies:
     - event: TaskCompleted
                                                                                                          配置作业终止策略
     action: CompleteJob
   template:
     spec:
      imagePullSecrets:
       - name: default-secret
      containers:
       - command:
         - sh
          PS_HOST=`cat /etc/volcano/ps.host | sed 's/$/&:2222/g' | tr "\n" ","`;
          WORKER_HOST=`cat /etc/volcano/worker.host | sed 's/$/&:2222/g' | tr "\n" ","`;
                                                                                                            Worker运行信息
          python tf cnn benchmarks.py --batch size=32 --model=resnet50 --
variable_update=parameter_server --flush_stdout=true --num_gpus=1 --
local_parameter_device=cpu --device=cpu --data_format=NHWC --job_name=worker --
task_index=${VK_TASK_INDEX} --ps_hosts=${PS_HOST} --worker_hosts=${WORKER_HOST}
        image: volcanosh/example-tf:0.0.1
        name: tensorflow
                                                                                                            暴露2222端口
        ports:
         - containerPort: 2222
          name: tfjob-port
        workingDir: /opt/tf-benchmarks/scripts/tf_cnn_benchmarks
```

TensorFlow 演示

Tensorflow作业状态:

NAME	READY	STATUS	RESTARTS	AGE
tensorflow-benchmark-ps-0	1/1	Running	0	6s
tensorflow-benchmark-worker-0	1/1	Running	0	6s
tensorflow-benchmark-worker-1	1/1	Running	0	6s

Tensorflow作业结果:

```
2023-07-03 11:25:20.462920: I tensorflow/core/distributed runtime/rpc/grpc server lib.cc:324] Started server with tar
get: grpc://localhost:2222
TensorFlow: 1.5
            resnet50
             training
SingleSess: False
Batch size: 32 global
             32 per device
            ['/job:worker/task:0/cpu:0']
Devices:
Data format: NHWC
Optimizer: sgd
            parameter server
Running parameter server 0
Running warm up
d3122e8662 with config: intra_op_parallelism_threads: 1 gpu_options { force_gpu_compatible: true } allow_soft_placeme
2023-07-03 11:25:23.826165: I tensorflow/core/distributed_runtime/master_session.cc:1008] Start master session 2f5f97
Generating model
```

作业输出







01 运行TensorFlow 作业



03 运行Spark作业







Pytorch 作业配置

apiVersion: batch.volcano.sh/v1alpha1 kind: Job metadata: 根据业务需求设置作业类型 name: pytorch-job Tabels: "volcano.sh/job-type": "Pytorch" spec: minAvailable: 1 schedulerName: volcano 通过Volcano提供的通用插件配置Pytorch作业的角色和端口 plugins: 信息 pytorch: ["--master=master","-worker=worker", "--port=23456"] 配置作业重启策略 policies: - event: PodEvicted action: RestartJob







Pytorch 作业配置

```
tasks:
  - replicas: 1
   name: master
   policies:
     - event: TaskCompleted
      action: CompleteJob
   template:
     spec:
      containers:
       - image: gcr.io/kubeflow-ci/pytorch-dist-sendrecv-test:1.0
        imagePullPolicy: IfNotPresent
        name: master
      restartPolicy: OnFailure
  - replicas: 2
   name: worker
   template:
     spec:
      containers:
       - image: gcr.io/kubeflow-ci/pytorch-dist-sendrecv-test:1.0
        imagePullPolicy: IfNotPresent
        name: worker
        workingDir: /home
      restartPolicy: OnFailure
```

配置作业终止策略。当pytorch master 完成时,整个Pytorch作业被视为完成。

Worker角色配置







Pytorch 演示

Pytorch作业状态:

```
[root@ecs-4b42-0002 demo-kind-test]# kubectl get pod
NAME
                        READY
                                STATUS
                                          RESTARTS
                                                      AGE
pytorch-job-master-0
                        1/1
                                Running
                                                      4s
pytorch-job-worker-0
                        1/1
                                Running
                                                      4s
pytorch-job-worker-1
                        1/1
                                Running
                                                      4s
[root@ecs-4b42-0002 demo-kind-test]# kubectl get pod
NAME
                      READY
                               STATUS
                                           RESTARTS
                                                      AGE
pytorch-job-master-0
                               Completed
                                                      36s
                       0/1
pytorch-job-worker-0
                       0/1
                               Completed
                                                      36s
pytorch-job-worker-1
                       0/1
                               Completed
                                                      36s
```

Pytorch作业结果:

Pytorch 作业调度完成

Pytorch 作业运行完成

作业输出







01 运行TensorFlow 作业

02 运行Pytorch作业



03 运行Spark作业







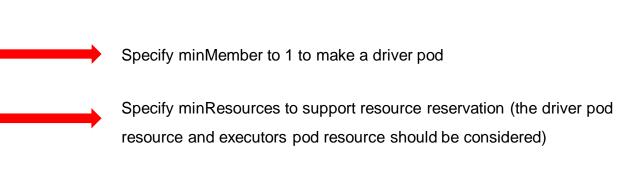
Spark 作业配置

Podgroup config:

apiVersion: scheduling.volcano.sh/v1beta1
kind: PodGroup
spec:
minMember: 1

minResources:
cpu: "4"
memory: "5Gi"

priorityClassName: system-node-critical
queue: default









Spark 作业配置

```
./spark-3.3.1/bin/spark-submit \
--master k8s://https://127.0.0.1:40883 \
--deploy-mode cluster \
--driver-cores 1 \
--driver-memory 2G \
--num-executors 1 \
--executor-cores 1 \
--executor-memory 1G \
--name spark-volcano-wy1 \
--class org.apache.spark.examples.SparkPi \
--conf spark.executor.instances=3 \
--conf spark.kubernetes.container.image=wangyang0616/spark:3.3.1-volcano.v1 \
--conf spark.kubernetes.authenticate.driver.serviceAccountName=spark \
--conf spark.kubernetes.scheduler.name=volcano \
--conf spark.kubernetes.scheduler.volcano.podGroupTemplateFile=.../podgroup-template.yaml \
--conf spark.kubernetes.driver.pod.featureSteps=org.apache.spark.deploy.k8s.features.VolcanoFeatureStep \
--conf spark.kubernetes.executor.pod.featureSteps=org.apache.spark.deploy.k8s.features.VolcanoFeatureStep \
local:///opt/spark/examples/jars/spark-examples_2.12-3.3.1.jar
```

- 1. Specify custom scheduler
- 2. Specify scheduler hints (podgroup template)
- 3. Specify custom feature step







Spark 演示

Spark 作业状态:

NAME spark-pi-43882d891f8c6886-exec-1 spark-pi-43882d891f8c6886-exec-2 spark-pi-43882d891f8c6886-exec-3 spark-volcano-wy1-b98f3d891f8c5504-driver	READY 1/1 1/1 1/1 1/1	Running Running	RESTARTS 0 0 0 0	AGE 7s 7s 7s 12s
NAME	READY	STATUS	RESTARTS	AGE
spark-volcano-wy1-b98f3d891f8c5504-driver	0/1	Completed	0	35s







Thanks