

Tensorflow2.0 with grpc+verbs

环境准备

name	version
tensorflow	2.0
gcc,g++	7.3.1 20180303 (默认：4.8.5需要升级)
python	2.7.5
bazel	0.26.1
centos	CentOS Linux release 7.8.2003 (Core)
git	2.31.1 (默认: 1.8.3需要升级)

编译

tensorflow源码准备

1. git clone https://github.com/tensorflow/tensorflow.git
2. git checkout -b r2.0 origin/r2.0

bazel安装

1. wget https://releases.bazel.build/0.26.1/release/bazel-0.26.1-installer-linux-x86_64.sh
2. sh bazel-0.26.1-installer-linux-x86_64.sh
3. ln -s /usr/local/bin/bazel /usr/bin/bazel

git升级

1. yum install http://opensource.wandisco.com/centos/7/git/x86_64/wandisco-git-release-7-2.noarch.rpm
2. yum remove git
3. yum install git

git 1.8.3 相关错误：

```
+ git -C /root/.cache/bazel/_bazel_root/79b7ee1a7abc6e1e5dea514986007/external/io_bazel_rules_docker reset --hard 251f6a68b439744094faff800cd029798edf9faa
Unknown option: -C
usage: git [--version] [--help] [-c name=value]
        [--exec-path[=<path>]] [--html-path] [--man-path] [--info-path]
        [-p|--paginate|--no-pager] [--no-replace-objects] [--bare]
        [--git-dir=<path>] [--work-tree=<path>] [--namespace=<name>]
        <command> [<args>]
+ git -C /root/.cache/bazel/_bazel_root/79b7ee1a7abc6e1e5dea514986007/external/io_bazel_rules_docker fetch '' origin 251f6a68b439744094faff800cd029798edf9faa:251f6a68b439744094faff800cd029798edf9faa
Unknown option: -C
usage: git [--version] [--help] [-c name=value]
        [--exec-path[=<path>]] [--html-path] [--man-path] [--info-path]
        [-p|--paginate|--no-pager] [--no-replace-objects] [--bare]
        [--git-dir=<path>] [--work-tree=<path>] [--namespace=<name>]
        <command> [<args>]
+ git -C /root/.cache/bazel/_bazel_root/79b7ee1a7abc6e1e5dea514986007/external/io_bazel_rules_docker fetch origin 251f6a68b439744094faff800cd029798edf9faa:251f6a68b439744094faff800cd029798edf9faa
Unknown option: -C
usage: git [--version] [--help] [-c name=value]
        [--exec-path[=<path>]] [--html-path] [--man-path] [--info-path]
        [-p|--paginate|--no-pager] [--no-replace-objects] [--bare]
        [--git-dir=<path>] [--work-tree=<path>] [--namespace=<name>]
        <command> [<args>]
INFO: Elapsed time: 6.460s
INFO: 0 processes.
FAILED: Build did NOT complete successfully (0 packages loaded)
```

gcc

1. `wget http://ftp.gnu.org/gnu/gcc/gcc-7.3.0/gcc-7.3.0.tar.gz`
2. `tar -zxvf gcc-7.3.0.tar.gz`
3. `cd gcc-7.3.0`
4. `./contrib/download_prerequisites`
5. `mkdir build && cd build`
6. `../configure --prefix=/usr/local/gcc --enable-languages=c,c++ --disable-multilib`
7. `make && make install`
8. `mv /usr/bin/{gcc,gcc.bak}`
9. `mv /usr/bin/{g++,g++.bak}`
10. `ln -s /usr/local/gcc/bin/gcc /usr/bin/gcc`
11. `ln -s /usr/local/gcc/bin/g++ /usr/bin/g++`

依赖包安装

name	version
pip	
numpy	
builtins	
python-devel	
libibverbs-devel	
enum	
keras_preprocessing	
mock	

pip

1. `wget https://bootstrap.pypa.io/pip/2.7/get-pip.py`
2. `python get-pip.py`

numpy

```
1. pip install numpy
```

builtins

```
1. pip install future
```

libibverbs

```
1. yum install libibverbs-devel
```

enum

```
1. pip install enum34
```

keras_preprocessing

```
1. pip install keras_preprocessing
```

mock

```
1. pip install mock
```

编译tensorflow

```
1. cd ./tensorflow #进入tensorflow源码目录
2. bazel build --config=opt --config=verbs --distdir=/home/bazel_package
   //tensorflow/tools/pip_package:build_pip_package --verbose_failures
```

编译成功如下：

```
INFO: Analyzed target //tensorflow/tools/pip_package:build_pip_package (0 packages loaded, 0 targets configured).
INFO: Found 1 target...
Target //tensorflow/tools/pip_package:build_pip_package up-to-date:
  bazel-bin/tensorflow/tools/pip_package/build_pip_package
INFO: Elapsed time: 1853.521s, Critical Path: 276.04s
INFO: 2833 processes; 2833 local.
INFO: Build completed successfully, 2840 total actions
```

制作安装包

```
1. bazel-bin/tensorflow/tools/pip_package/build_pip_package /tmp/tensorflow_pkg
```

```
[root@k8s-3 tensorflow]# bazel-bin/tensorflow/tools/pip_package/build_pip_package /tmp/tensorflow_pkg
Thu Apr 14 16:20:13 CST 2022 : === Preparing sources in dir: /tmp/tmp.warnUXzcnK
/home/deeproute/tensorflow /home/deeproute/tensorflow
/home/deeproute/tensorflow
/tmp/tmp.warnUXzcnK/tensorflow/include /home/deeproute/tensorflow
/home/deeproute/tensorflow
Thu Apr 14 16:20:26 CST 2022 : === Building wheel
warning: no files found matching 'README'
warning: no files found matching '*.pyd' under directory '*'
warning: no files found matching '*.pd' under directory '*'
warning: no files found matching '*.dylib' under directory '*'
warning: no files found matching '*.dll' under directory '*'
warning: no files found matching '*.lib' under directory '*'
warning: no files found matching '*.csv' under directory '*'
warning: no files found matching '*.h' under directory 'tensorflow_core/include/tensorflow'
warning: no files found matching '*' under directory 'tensorflow_core/include/third_party'
Thu Apr 14 16:20:45 CST 2022 : === Output wheel file is in: /tmp/tensorflow_pkg
```

安装

```
1. pip install /tmp/tensorflow_pkg/tensorflow-2.0.4-cp27-cp27mu-linux_x86_64.whl
```

安装失败：

```
Failed to build grpcio
Installing collected packages: six, h5py, absl-py, pyasn1, pyasn1-modules, rsa, cachetools, google-auth, markdown, protobuf, oauthlib, certifi, urllib3, idna, chardet, requests, requests-oauthlib, google-auth-oauthlib, grpcio, Werkzeug, tensorboard, keras-applications, wrapt, opt-einsum, google-pasta, tensorflow-estimator, gast, termcolor, backports.weakref, astor, functools32, tensorflow
Attempting uninstall: six
  Found existing installation: six 1.9.0
  Uninstalling six-1.9.0:
    Successfully uninstalled six-1.9.0
Attempting uninstall: urllib3
  Found existing installation: urllib3 1.10.2
  Uninstalling urllib3-1.10.2:
    Successfully uninstalled urllib3-1.10.2
Attempting uninstall: chardet
  Found existing installation: chardet 2.2.1
  Uninstalling chardet-2.2.1:
    Successfully uninstalled chardet-2.2.1
Attempting uninstall: requests
  Found existing installation: requests 2.6.0
ERROR: Cannot uninstall 'requests'. It is a distutils installed project and thus we cannot accurately determine which files belong to it which would lead to only a partial uninstall.
```

解决：

```
1. sudo pip install icrawler --ignore-installed requests
```

再次安装tensorflow:

```
1. pip install /tmp/tensorflow_pkg/tensorflow-2.0.4-cp27-cp27mu-linux_x86_64.whl
```

测试

下载benchmark源码（server端和client端都需要）

```
1. git clone https://github.com/tensorflow/benchmarks.git
2. git checkout -b cnn_tf_v1.12_compatible origin/cnn_tf_v1.12_compatible
```

运行：

ps 节点（10.9.9.146），由于没有gpu，使用cpu代替：

```
1. cd benchmarks/scripts/tf_cnn_benchmarks
2. python tf_cnn_benchmarks.py --device=CPU --batch_size=64 --model=resnet50 --
variable_update=parameter_server --data_format=NHWC --server_protocol=grpc+verbs
--variable_update=distributed_replicated --job_name=ps --ps_hosts=10.9.9.146:5001
--worker_hosts=10.9.9.147:5001 --local_parameter_device=cpu
```

worker 节点（10.9.9.147），由于没有gpu，使用cpu代替：

```
1. cd benchmarks/scripts/tf_cnn_benchmarks
2. python tf_cnn_benchmarks.py --device=CPU --batch_size=64 --model=resnet50 --
variable_update=parameter_server --data_format=NHWC --server_protocol=grpc+verbs
--variable_update=distributed_replicated --job_name=worker --
ps_hosts=10.9.9.146:5001 --worker_hosts=10.9.9.147:5001 --
local_parameter_device=cpu
```

```
[root@k8s-1 tf_cnn_benchmarks]# python tf_cnn_benchmarks.py --device=CPU --batch size=64 --model=resnet50 --variable update=parameter server --data format=NHWC --server_protocol=grpc-verbs --variable update=distributed_replicated --job name=ps --ps hosts=10.9.9.146:5001 --worker hosts=10.9.9.147:5001 --local parameter device=cpu
WARNING:tensorflow:From /usr/lib64/python2.7/site-packages/tensorflow_core/python/compat/v2_compat.py:65: disable_resource_variables (from tensorflow.python.ops.variable_scope) is deprecated and will be removed in a future version.
Instructions for updating:
non-resource variables are not supported in the long term
2022-04-14 19:03:44.402456: I tensorflow/core/platform/cpu_feature_guard.cc:142] Your CPU supports instructions that this TensorFlow binary was not compiled to use: SSE4.2
2022-04-14 19:03:44.412673: I tensorflow/core/platform/profile_utils/cpu_utils.cc:94] CPU Frequency: 2593910000 Hz
2022-04-14 19:03:44.414912: I tensorflow/compiler/xla/service/service.cc:168] XLA service 0x47e0fc0 executing computations on platform Host. Devices:
2022-04-14 19:03:44.414970: I tensorflow/compiler/xla/service/service.cc:175] StreamExecutor device (0): Host, Default Version
2022-04-14 19:03:44.419317: I tensorflow/core/distributed_runtime/rpc/grpc_channel.cc:258] Initialize GrpcChannelCache for job ps -> {0 -> localhost:5001}
2022-04-14 19:03:44.419756: I tensorflow/core/distributed_runtime/rpc/grpc_channel.cc:258] Initialize GrpcChannelCache for job worker -> {0 -> 10.9.9.147:5001}
2022-04-14 19:03:44.423340: I tensorflow/core/distributed_runtime/rpc/grpc_channel.cc:258] Initialize GrpcChannelCache for job ps -> {0 -> localhost:5001}
2022-04-14 19:03:44.423384: I tensorflow/core/distributed_runtime/rpc/grpc_channel.cc:258] Initialize GrpcChannelCache for job worker -> {0 -> 10.9.9.147:5001}
2022-04-14 19:04:55.652198: I tensorflow/contrib/verbs/rdma_mgr.cc:130] Connected to remote node /job:worker/replica:0/task:0
TensorFlow: 2.0
Model: resnet50
Dataset: imagenet (synthetic)
Mode: training
SingleSess: False
Batch size: 64 global
64 per device
Num batches: 100
Num epochs: 0.00
Devices: ['/job:worker/replica:0/task:0/CPU:0']
NUMA bind: False
Data format: NHWC
Optimizer: sgd
Variables: distributed_replicated
Sync: True
=====
Running parameter server 0
```

参考链接

- https://github.com/tensorflow/benchmarks/tree/cnn_tf_v2.0_compatible
- https://github.com/tensorflow/benchmarks/tree/master/scripts/tf_cnn_benchmarks
- <https://github.com/bytedance/byteps/blob/master/docs/step-by-step-tutorial.md>