

AIXPRT Community Preview user guide

Table of Contents

AIXPRT Community Preview user guide	1
Introduction	2
Getting started (all network models)	2
Intel OpenVINO (CPU and GPU)	2
System requirements	2
Installation and system configuration	3
Building the workloads	3
Running the benchmark	3
TensorFlow (CPU and GPU)	4
System requirements	4
Installation and system configuration	4
Running the benchmark	5
TensorFlow with NVIDIA TensorRT (GPU)	6
System requirements	6
Installation and system configuration	6
Running the benchmark	7
Results	8
Support.....	8

Introduction

AIXPRT has the capability of running on the following platforms:

- AMD CPU
- AMD GPU
- Intel CPU
- Intel GPU
- Nvidia GPU

The workloads are implemented using the publicly available libraries and SDKs for each platform.

Getting started (all network models)

The first step after receiving access to AIXPRT on GitHub is to clone the AIXPRT repository.

1. Download the AIXPRT repository. Since the repository contains large files, over 50MB in size, the git-lfs package must be installed and the repository must be cloned (A zip file of the repository will not include the large files).
 - a. Install git lfs. The instructions are found at <https://packagecloud.io/github/git-lfs/install> and are listed in the following 2 steps.

```
curl -s https://packagecloud.io/install/repositories/github/git-lfs/script.deb.sh | sudo bash
sudo apt-get install git-lfs
```

- b. Clone the AIXPRT repository (You may need to enter your credentials for each large file).

```
git clone https://github.com/BenchmarkXPRT/AIXPRT.git
```

2. Navigate to AIXPRT/Modules/Deep-Learning/README.md of the cloned branch and follow instructions to run the benchmark.
3. AIXPRT runs on a default configuration provided by the workloads. However, once a default config file is generated, a user can edit the config to change workload behavior.

NOTE : A config file will be generated at AIXPRT/Config/{config_name}.json after the first run of index.py. One can edit this config file to run the specific workload in a specific way or use one of the provided config files.

Intel OpenVINO (CPU and GPU)

This module contains workloads to evaluate the system performance of use cases related to image classification and object detection using the Intel Distribution of OpenVINO toolkit for Linux. It contains the resnet50_v1 and ssd-mobilenet workloads and runs single and multi-batch size scenarios.

System requirements

Operating System

- Ubuntu 16.04 LTS (64-bit)

CPU

- 6th to 8th generation Intel Core and Intel Xeon processors
- Intel Pentium processor N4200/5, N3350/5, N3450/5 with Intel HD Graphics

Installation and system configuration

1. Download the AIXPRT repository. Since the repository contains large files, over 50MB in size, the git-lfs package must be installed and the repository must be cloned (A zip file of the repository will not include the large files).

- a. Install git lfs. The instructions are found at <https://packagecloud.io/github/git-lfs/install> and are listed in the following 2 steps.

```
curl -s https://packagecloud.io/install/repositories/github/git-lfs/script.deb.sh | sudo bash
sudo apt-get install git-lfs
```

- b. Clone the AIXPRT repository (You may need to enter your credentials for each large file).

```
git clone https://github.com/BenchmarkXPRT/AIXPRT.git
```

- c. Move the AIXPRT/OpenVINO/AIXPRT repository directory to /home/[user]/AIXPRT.

2. Install all dependencies

- a. Register on the Intel OpenVINO toolkit page (link below) and download the OpenVINO for Linux toolkit.

<https://software.intel.com/en-us/openvino-toolkit/choose-download/free-download-linux>

- b. Install the full package version of Intel OpenVINO by following the instructions at <https://software.intel.com/en-us/articles/OpenVINO-Install-Linux>.

- c. Install pillow using the following script:

```
sudo apt-get install python-pip python-numpy
pip install Pillow
```

Building the workloads

1. Make sure you are connected to an active internet connection.
2. Grant the necessary executable permissions to the compile_AIXPRT_sources.sh file which is located at ~AIXPRT/install:

```
chmod 777 compile_AIXPRT_sources.sh
```

3. Run the AIXPRT install script:

```
cd AIXPRT/install
sudo ./compile_AIXPRT_sources.sh
```

Running the benchmark

1. Navigate to AIXPRT harness directory:

```
cd AIXPRT/Harness
```

2. Run the benchmark using the following script:

```
python3 index.py
```

TensorFlow (CPU and GPU)

This module contains workloads to evaluate the system performance of use cases related to image classification and object detection using TensorFlow. It contains the resnet50_v1 and ssd-mobilenet workloads and runs single and multi-batch size scenarios.

Workloads are built and tested using the TensorFlow (version 1.12) framework. For more information about TensorFlow please go to <https://www.tensorflow.org>. Workloads run with fp32 precision by default.

System requirements

Operating System

- Ubuntu 16.04 LTS (64-bit)

CPU and GPU

- Please consult the TensorFlow website for minimum hardware and software requirement information: <https://www.tensorflow.org/install>.

Installation and system configuration

1. Download the AIXPRT repository. Since the repository contains large files, over 50MB in size, the git-lfs package must be installed and the repository must be cloned (A zip file of the repository will not include the large files).

- a. Install git lfs. The instructions are found at <https://packagecloud.io/github/git-lfs/install> and are listed in the following 2 steps.

```
curl -s https://packagecloud.io/install/repositories/github/git-lfs/script.deb.sh | sudo bash
sudo apt-get install git-lfs
```

- b. Clone the AIXPRT repository (You may need to enter your credentials for each large file).

```
git clone https://github.com/BenchmarkXPRT/AIXPRT.git
```

- c. Move the AIXPRT/OpenVINO/AIXPRT repository directory to /home/[user]/AIXPRT.

2. Install all dependencies.

```
sudo apt-get update
sudo apt-get install python3 python3-numpy python3-pil
sudo apt-get install python python-numpy python-pil
```

3. Install TensorFlow.

- For Intel and AMD CPUs, please read the instructions here: <https://www.tensorflow.org/install>.

```
sudo apt-get install python-pip
```

```
pip install tensorflow
```

- For NVIDIA GPUs, follow the instructions here: <https://www.tensorflow.org/install/gpu>.

- a. Remove any installed drivers.

```
sudo apt-get purge nvidia*
```

- b. Retrieve the repository.

```
sudo add-apt-repository ppa:graphics-drivers/ppa
```

- c. Update the apt request.

```
sudo apt update
```

- d. Confirm that the desired driver is present.

```
ubuntu-drivers devices
```

- e. Install the desired driver.

```
sudo apt install nvidia-410
```

```
reboot
```

- f. Verify that the desired driver is installed.

```
nvidia-smi
```

```
sudo apt-get install python-pip
```

```
pip install tensorflow-gpu==1.12.0
```

- To install TensorFlow with AMD ROCm support, follow the instructions here: <https://rocm.github.io/dl.html>.

NOTE: For AMD-GPUs, Ubuntu 18.04 supports the latest drivers, so we recommend using Ubuntu 18.04 in those instances.

Running the benchmark

1. Navigate to AIXPRT harness directory:

```
cd AIXPRT/Harness
```

2. Run the benchmark using the following script:

```
python3 index.py
```

3. If running on a GPU target, please edit the AIXPRT/Config/{filename.json} to set "hardware" to gpu.

TensorFlow with NVIDIA TensorRT (GPU)

This module contains workloads to evaluate the system performance of use cases related to Image Classification and Object Detection using Tensorflow with TensorRT optimizations. It contains the resnet50_v1 and ssd-mobilenet workloads and runs single and multi-batch size scenarios.

System requirements

Operating System

- Ubuntu 16.04 LTS (64-bit)

GPU

- CUDA-enabled NVIDIA GPUs. For more info, visit <https://developer.nvidia.com/cuda-gpus>.

For NVIDIA Tegra Xavier

- JetPACK 4.1.1

Installation and system configuration

1. Download the AIXPRT repository. Since the repository contains large files, over 50MB in size, the git-lfs package must be installed and the repository must be cloned (A zip file of the repository will not include the large files).

- a. Install git lfs. The instructions are found at <https://packagecloud.io/github/git-lfs/install> and are listed in the following 2 steps.

```
curl -s https://packagecloud.io/install/repositories/github/git-lfs/script.deb.sh | sudo bash
sudo apt-get install git-lfs
```

- b. Clone the AIXPRT repository (You may need to enter your credentials for each large file).

```
git clone https://github.com/BenchmarkXPRT/AIXPRT.git
```

- c. Move the AIXPRT/OpenVINO/AIXPRT repository directory to /home/[user]/AIXPRT.

2. Install all dependencies.

- a. For NVIDIA discrete graphics
 - i. Install CUDA 10 from <https://developer.nvidia.com/cuda-downloads>.
 - ii. Restart the system.
 - iii. Install docker from <https://docs.docker.com/install/linux/docker-ce/ubuntu/#install-from-a-package>. Run `lsb_release -a` to show you the Ubuntu version and name. You will also need to install the cli and container packages.
 - iv. Test the installation of the docker image.

```
sudo docker run hello-world
```

- v. Run the following command to be able to run as a regular user. If it does not work on the first try, reboot the system.

```
sudo usermod -a -G docker $USER
```

- vi. Install NVIDIA-docker from <https://github.com/NVIDIA/nvidia-docker>.
- vii. Pull and run TensorRT docker container v19.01

```
Docker pull nvcr.io/nvidia/tensorflow:19.01-py3
```

- b. If using NVIDIA Tegra Xavier, flash JetPACK v4.1.1 from <https://developer.nvidia.com/embedded/jetpack>.

Running the benchmark

1. Choose the target machine and run the following commands:
 - a. If using NVIDIA discrete graphics, run the docker image.

```
nvidia-docker run -v <Path_to_AIXPRT_directory>:/workspace/AIXPRT --shm-size=1g --ulimit memlock=-1 --ulimit stack=67108864 -it --rm nvcr.io/nvidia/tensorflow:19.01-py3
```

- b. If using NVIDIA Tegra Xavier, install the necessary prerequisites:

```
sudo apt-get install libhdf5-serial-dev hdf5-tools
sudo apt-get install python3-pip
pip3 install --extra-index-url https://developer.download.nvidia.com/compute/redist/jp/v411 tensorflow-gpu==1.12.0-rc2+nv18.11
sudo apt-get install python3-matplotlib
```

2. Add the models directory to PYTHONPATH to install tensorflow/models and Run the TF Slim setup.
3. Run script setup.sh inside /workspace/AIXPRT.
 - a. If the script fails, run the following commands manually:

```
git clone https://github.com/tensorflow/models.git
cd models
export PYTHONPATH="$PYTHONPATH:$PWD"
cd research
export PYTHONPATH="$PYTHONPATH:$PWD"
wget -O protobuf.zip https://github.com/google/protobuf/releases/download/v3.0.0/protoc-3.0.0-linux-x86_64.zip
unzip protobuf.zip
./bin/protoc object_detection/protos/*.proto --python_out=.
cd slim
python setup.py install
pip install requests pillow
```

4. Navigate to AIXPRT harness directory:

```
cd AIXPRT/Harness
```

5. Run the benchmark using the following script:

```
python3 index.py
```

Results

When the test is complete, the benchmark saves results to AIXPRT/Results. To submit results, please follow the instructions in ResultSubmission.md or at

<https://www.principledtechnologies.com/benchmarkxpert/aixprt/2019/submit-results.php>.

Support

If you need technical support or have any questions, please send a message to BenchmarkXPRTsupport@principledtechnologies.com.