Advanced Simulations using Volunteer Computing:

JANE ARLETH DELA CRUZ

Supervisors:
Laurence Field, Ben Segal, Francois Grey

Volunteer Computing

LHC @home

BOINC

Berkeley Open Infrastructure for Network Computing

ATLAS@home	ATLAS@home (Event Simulations)
	Beauty(Beauty Physics Simulations)
CMS	CMS@Home(Event Simulations)
	SixTrack(Beam Simulations)
	Test4Theory(Monte Carlo Events)

LHC@Home Project for “SDG”

- Help some non-CERN scientists use the same LHC@Home infrastructure to solve health-related problems such as epidemiology modelling

- Run Machine Learning using Keras and Tensorflow Test on BOINC for the first time.

Machine Learning for SDG

- Combining satellite imagery and machine learning to predict poverty (Stanford University, USA, National Bureau of Economic Research, USA)
- Classifying Smoking Urges Via Machine Learning (University of Pittsburgh, USA)
- Machine Learning Based Big Data Processing Framework for Cancer Diagnosis Using Hidden Markov Model and GM Clustering

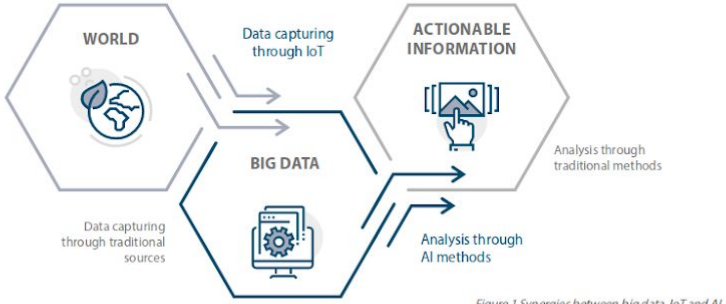


Figure 1 Synergies between big data, IoT and AI

Jane Arleth dela Cruz

Explored Solutions	Pros	Cons
Virtualization  Virtualbox	Hosts are heterogeneous in terms of OS Provide a unified software environment	Requires Hypervisor GPU support untested
Containerization  Singularity	Runs natively on SLC6/CentOS7(Can run on GPU's)	Only runs on Linux

- Build Singularity Container
  Python2.7, Keras 2.2 , Tensorflow 1.70
- Test Container with ML jobs
  Training and Prediction

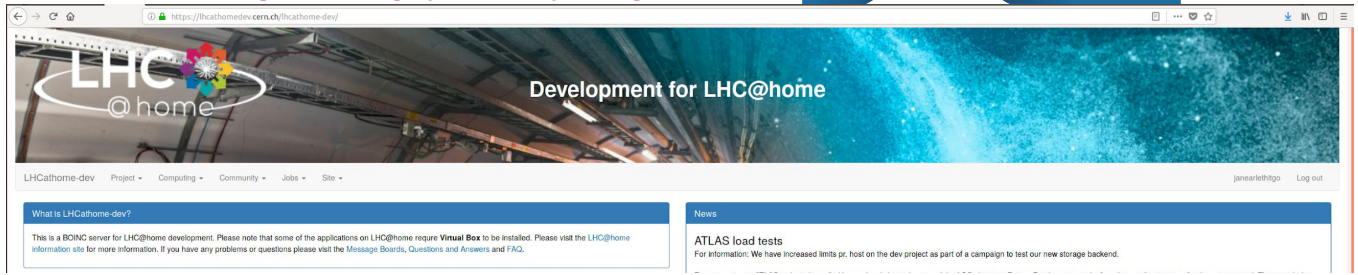

```
$ singularity run container.img  learn.py
LD_LIBRARY_PATH: /usr/local/cuda/lib64:/.singularity.d/libs
PATH: /usr/local/cuda/bin:/bin:/sbin:/usr/bin:/usr/sbin:/usr/local/bin:/usr/local/sbin
Arguments received: learn.py

  from .. import h5g, h5i, h5o, h5r, h5t, h5l, h5p
Using TensorFlow backend.

Preparing model, please wait...
('x_train shape:', (50000, 32, 32, 3))
(50000, 'train samples')
(10000, 'test samples')
Train on 50000 samples, validate on 10000 samples
Epoch 1/1
50000/50000 [==============================] - 210s 4ms/step - loss: 1.8072 - acc: 0.3364 - val_loss: 1.5128 - val_acc:
0.4546
model.h5
```

Make the BOINC Application

- Submit the application to BOINC-server via Condor Submission

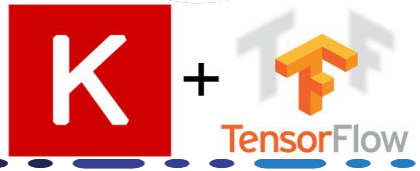


- Run application on Volunteer Computer via BOINC-Manager

Conclusion

- LHC@Home has potential to run ML jobs.
- Potential Internal ML Computing Resource using Desktop Computers
- First time to run ML (Keras+Tensorflow) on VC.
- (For now,) ML jobs can be run on Linux only with Singularity.

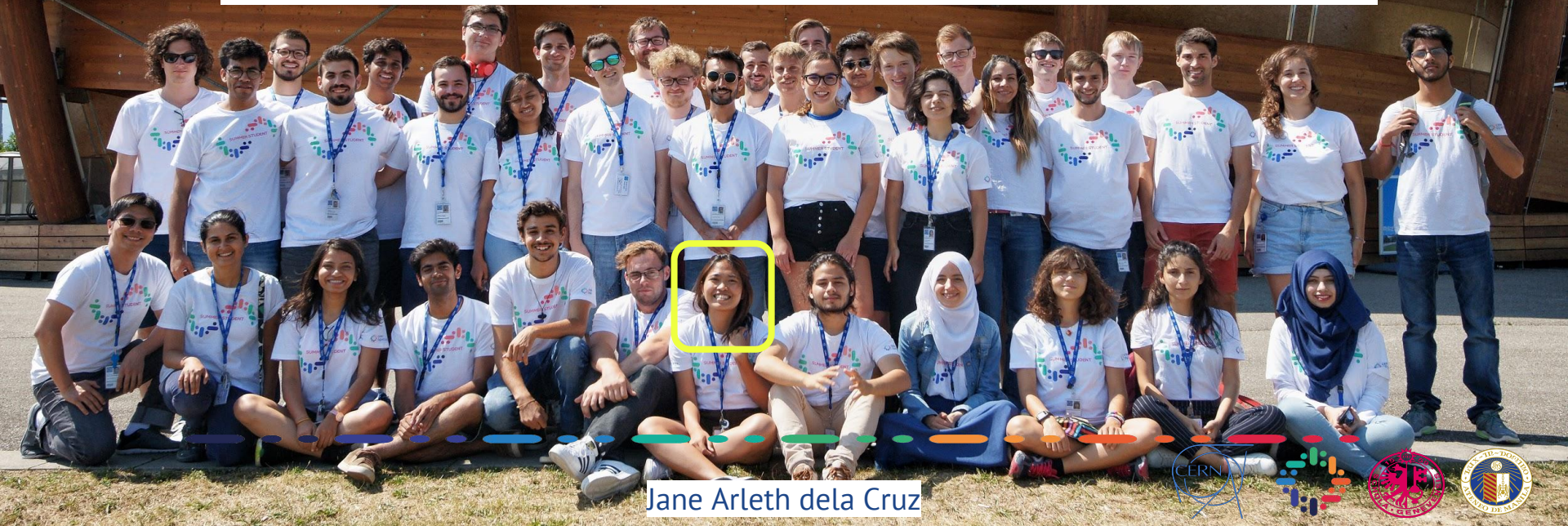




Thank You!

Questions?

jane.delacruz@obf.ateneo.edu



References

LHC@Home: http://lhcathome.web.cern.ch/

BOINC: http://boinc.berkeley.edu

Singularity Python Methods, University of Arizona HPC: https://docs.hpc.arizona.edu/display/UAHPC/Singularity+-+CentOS7%2C+Tensorflow1.4.1%2C+Keras%2C+Python3.5%2C+Cuda8.0%2C+cuDNN6

LHC@Home: a BOINC-based volunteer computing infrastructure for physics studies at CERN: http://ceur-ws.org/Vol-1973/paper02.pdf

Advances in ATLAS@Home towards a major ATLAS computing resource: http://cds.cern.ch/record/2626282/files/ATL-SOFT-SLIDE-2018-409.pdf

Backfilling the Grid with Containerized BOINC in the ATLAS computing: https://indico.cern.ch/event/587955/contributions/2937192/attachments/1683819/2706562/Backfilling_the_Grid_with_Containerized_BOINC_in_the_ATLAS_computing.pdf

Innovative Big Data Approaches for Capturing and Analyzing Data to Monitor and Achieve the SDGs: https://reliefweb.int/sites/reliefweb.int/files/resources/Innovative%20Big%20Data%20Approaches%20for%20Capturing%20and%20Analyzing%20Data%20to%20Monitor%20and%20Achieve%20the%20SDGs.pdf