Using Python for Deep Learning in Triton

Deep Learning: General Information

- Nowadays Python has many popular deep learning packages like Google tensorflow, deeplearning.net's Theano, UC Berkeley's caffe etc.
- In Triton tensorflow and Theano are provided through anaconda modules
- Caffe is quite difficult to maintain
  (though in future containers might solve this)
- We're going to change our Anaconda version numbers in the future, currently they are not necessarily representative
- At least anaconda2/4.0.0 has tensorflow and Theno available

  → Check package versions before use

Deep Learning: Using Anaconda

- Anaconda can be used as a starting point to create your own Python environment

  module load anaconda2/4.0.0
  conda create -p <path to environment> <package1> <package2> ...

- Creating environments like this will save up your quota: packages are not really installed, they are just symlinked to main anaconda installation
- You can activate/deactivate the environment with

  source activate <path to environment>
  source deactivate <path to environment>

Deep Learning: Using Anaconda

- You can then install packages in the environment with conda or pip

```
conda install <package>
pip install <package>
```

- Remember to load all of the modules & environments within the sbatch scripts as well

Deep Learning: Example with Keras

- Lets run a simple interactive example from Keras repository (https://github.com/fchollet/keras)
- Getting the interactive queue:

  sinteractive -t 00:30:00 --gres=gpu:teslak80:1 --mem=4G -p gpushort

- On node:

  module load anaconda/4.0.0
  source activate <path to environment>
  module load CUDA

  srun --gres=gpu:teslak80:1 python addition_rnn.py

Deep Learning: Possible issues

- Conda has a bad habit of writing package cache in ~/.conda
  → You can symlink this away to folder in $WRKDIR
- If you use “conda create -n <name>”, the whole environment will go to ~/.conda (and fill your /home quota)
  → Use conda create -p <prefix> instead
- Theano uses ~/.theano for compiling code. This can cause problems if you run code on multiple machines at a same time.
  → This can be configured in ~/.theanorc

```
[global]
base_compiledir=/tmp/%(user)s/theano
```

Any questions?