

BIOWULF

HIGH PERFORMANCE COMPUTING AT THE NIH

Deep Learning in Biowulf

Jerez Te

HPC @ NIH Support Scientist [C]

Biowulf

- Largest HPC system solely devoted to biological computations
- Top 132 supercomputer world-wide
- Hardware: 90,000+ processors
- GPU compute cards: K20x, K80, P100, V100
- Software: Centrally installed apps for ML/DL/AI

The screenshot shows the Biowulf website interface. At the top, there's a navigation bar with links for Systems, Applications, Reference Data, Storage, User Guides, Training, User Dashboard, How To, and About. Below this is a main content area with a heading "The NIH HPC group plans, manages and supports high-performance computing systems specifically for the intramural NIH community." followed by a paragraph describing the systems and their applications. To the left, there's a "Quick Links" section with a list of links: System Status, How To..., Application/DB updates, User Guides, Policies, Training, and Contact Us. In the center, there's a "Biowulf Utilization" section with a bar chart showing "Allocated CPUs (K)" for "Monday, July 8th, 2019". Below the chart is a table for "Last 24 hrs" utilization:

Last 24 hrs	
50,315 jobs submitted	19 NIH Institutes
33,019 jobs completed	108 Principal Investigators
2,021,878 CPU hrs used	170 users

To the right, there's a "Recent Papers that used Biowulf & HPC Resources" section with a list of publications, each with a small thumbnail image and a brief description of the research.

Website: hpc.nih.gov

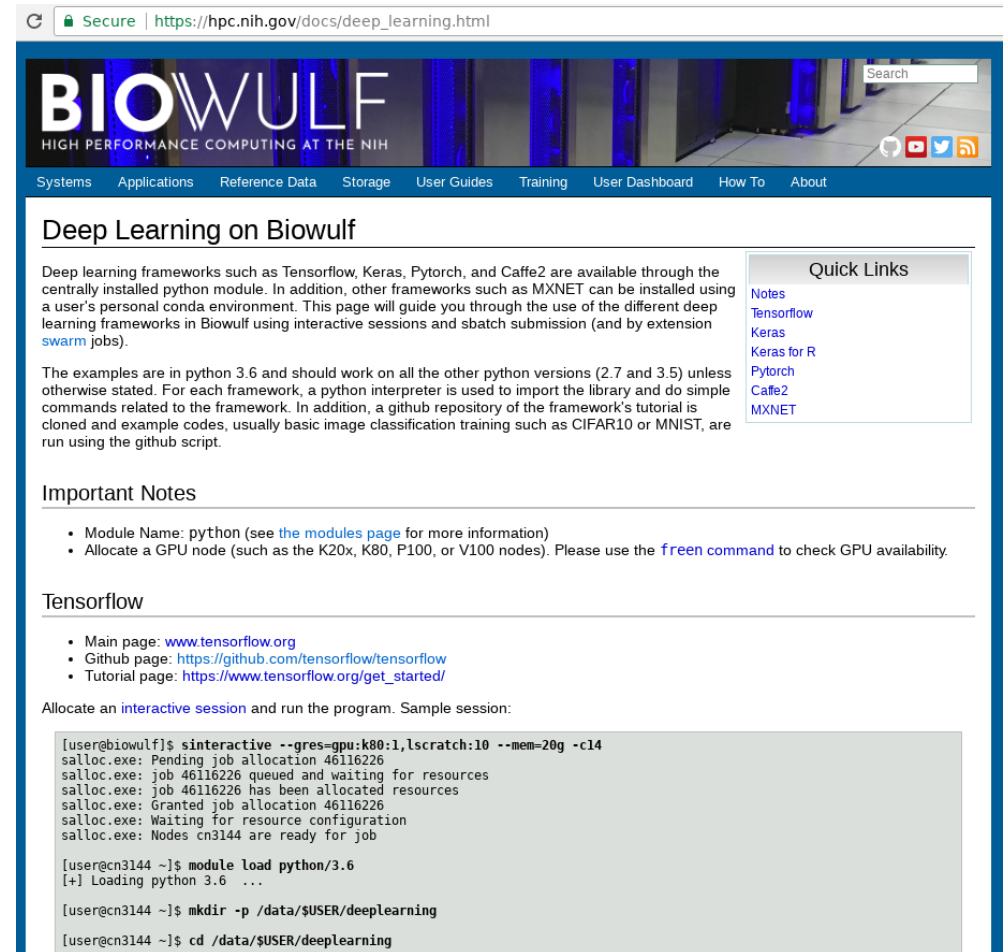
Hardware

		GPU Nodes			
• V100	8	28 x 2.4 GHz (Intel E5-2680v4) <i>hyperthreading enabled</i> 35 MB secondary cache 4 x NVIDIA V100 GPUs (16 GB VRAM, 5120 cores, 640 Tensor cores)	128 GB	56 Gb/s FDR Infiniband (1.11:1)	x2680 core28 cpu56 g256 gpuv100 ssd800 ibfdr
• P100	48	28 x 2.4 GHz (Intel E5-2680v4) <i>hyperthreading enabled</i> 35 MB secondary cache 4 x NVIDIA P100 GPUs (16 GB VRAM, 3584 cores)	128 GB	56 Gb/s FDR Infiniband (1.11:1)	x2680 core28 cpu56 g256 gpup100 ssd650 ibfdr
• K80	72	28 x 2.4 GHz (Intel E5-2680v4) <i>hyperthreading enabled</i> 35 MB secondary cache 2 x NVIDIA K80 GPUs with 2 x GK210 GPUs each (24 GB VRAM, 4992 cores)	256 GB	56 Gb/s FDR Infiniband (1.11:1)	x2680 core28 cpu56 g256 gpuk80 ssd800 ibfdr
• K20x	24	16 x 2.6 GHz (Intel E5-2650v2) <i>hyperthreading enabled</i> 20 MB secondary cache 2 x NVIDIA K20x GPUS (6 GB VRAM, 2688 cores)	128 GB	10 Gb/s Ethernet 56 Gb/s FDR Infiniband (1:11:1)	x2650 core16 cpu32 g128 ssd800 gpuk20x eth10g ibfdr

<https://hpc.nih.gov/systems/hardware.html>

Software

- 600+ apps
- Python and R as module
- Deep learning frameworks: TF, Keras, Pytorch
- Getting started with DL page
- Conda for personal packages
- Support singularity containers



The screenshot shows a web browser window with the URL https://hpc.nih.gov/docs/deep_learning.html. The page header features the Biowulf logo and navigation links: Systems, Applications, Reference Data, Storage, User Guides, Training, User Dashboard, How To, and About. The main content area is titled "Deep Learning on Biowulf" and includes a "Quick Links" sidebar with links to Notes, Tensorflow, Keras, Keras for R, Pytorch, Caffe2, and MXNET. The main text explains that deep learning frameworks like Tensorflow, Keras, Pytorch, and Caffe2 are available through the centrally installed python module. It also provides "Important Notes" regarding module names and GPU allocation. A "Tensorflow" section lists links to the main page, github page, and tutorial page. A terminal session is shown at the bottom, demonstrating how to allocate an interactive session, load the python/3.6 module, and create a directory for deep learning.

```
[user@biowulf]$ sinteractive --gres=gpu:k80:1,lscratch:10 --mem=20g -c14
salloc.exe: Pending job allocation 46116226
salloc.exe: job 46116226 queued and waiting for resources
salloc.exe: job 46116226 has been allocated resources
salloc.exe: Granted job allocation 46116226
salloc.exe: Waiting for resource configuration
salloc.exe: Nodes cn3144 are ready for job

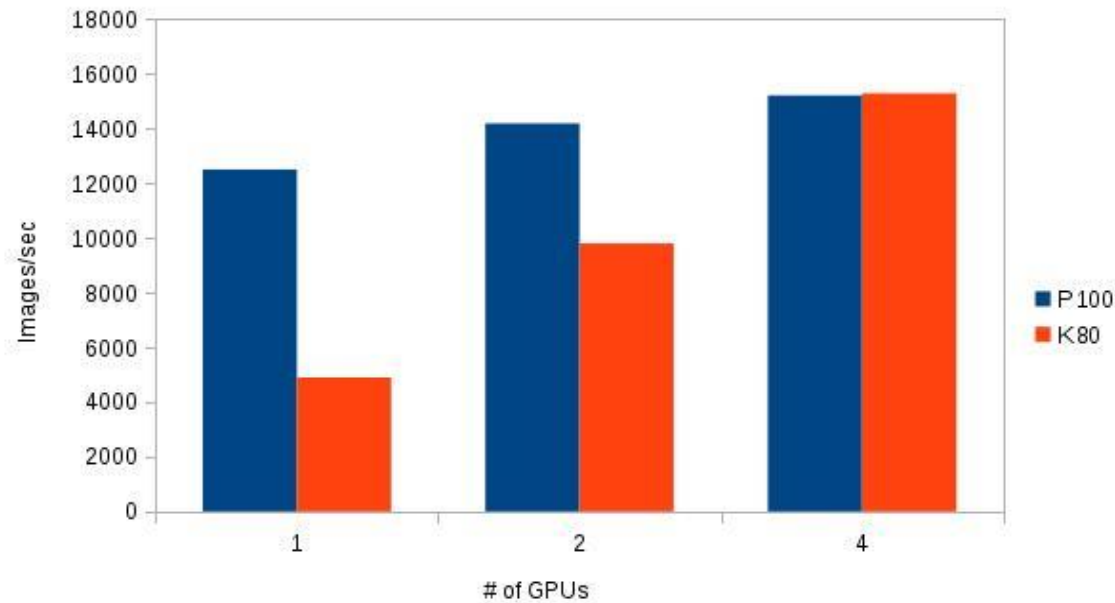
[user@cn3144 ~]$ module load python/3.6
[+] Loading python 3.6 ...

[user@cn3144 ~]$ mkdir -p /data/$USER/deeplearning

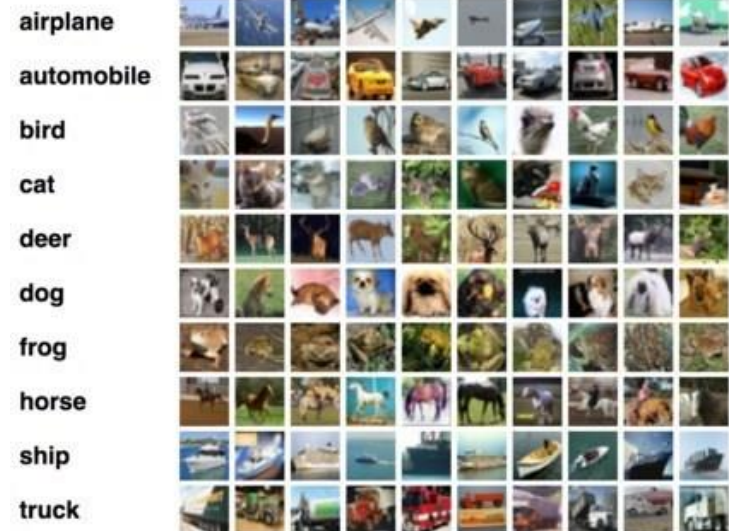
[user@cn3144 ~]$ cd /data/$USER/deeplearning
```

https://hpc.nih.gov/docs/deep_learning.html

Effects of sample size on training



- Cifar 10 – 32x32 color images in 10 classes
- 50K training set
- 10K test set



Source: appliedmachinelearning.blog

Monitoring

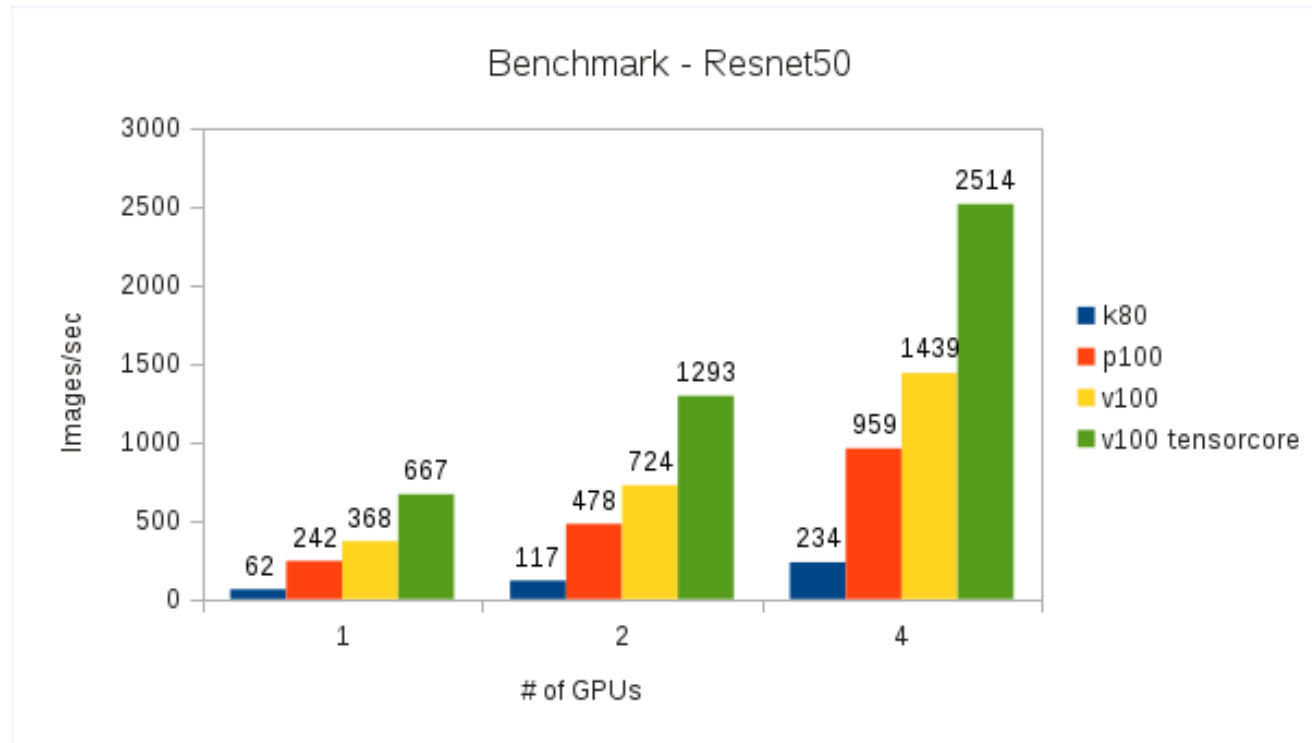
You can check the behaviour of your job with the 'nvidia-smi' utility. Determine the GPU node on which your job is running via [jobload](#) or [sjobs](#). Suppose your job is on node cn0626, and is using 2 GPUs:

```
biowulf% ssh cn0626 nvidia-smi
Mon Jan 16 17:52:13 2017
+-----+
| NVIDIA-SMI 367.48                Driver Version: 367.48                |
+-----+-----+
| GPU  Name      Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp   Perf   Pwr:Usage/Cap|      Memory-Usage | GPU-Util  Compute M. |
+-----+-----+
|  0  Tesla K80      On          | 0000:83:00.0  Off  |           Off       |
| N/A   41C    P0    115W / 149W | 145MiB / 12205MiB |    75%    Default   |
+-----+-----+
|  1  Tesla K80      On          | 0000:84:00.0  Off  |           Off       |
| N/A   27C    P8     33W / 149W |    0MiB / 12205MiB |     0%    Default   |
+-----+-----+
|  2  Tesla K80      On          | 0000:8A:00.0  Off  |           Off       |
| N/A   70C    P0    128W / 149W | 145MiB / 12205MiB |    80%    Default   |
+-----+-----+
|  3  Tesla K80      On          | 0000:8B:00.0  Off  |           Off       |
| N/A   45C    P8     33W / 149W |    0MiB / 12205MiB |     0%    Default   |
+-----+-----+

+-----+-----+
| Processes:                         GPU Memory |
| GPU       PID  Type  Process name                               Usage |
+-----+-----+
|  0        51187  C    ...cal/apps/amber/amber16/bin/pmemd.cuda.MPI  61MiB |
|  0        51188  C    ...cal/apps/amber/amber16/bin/pmemd.cuda.MPI  79MiB |
|  2        51187  C    ...cal/apps/amber/amber16/bin/pmemd.cuda.MPI  79MiB |
|  2        51188  C    ...cal/apps/amber/amber16/bin/pmemd.cuda.MPI  61MiB |
+-----+-----+
```

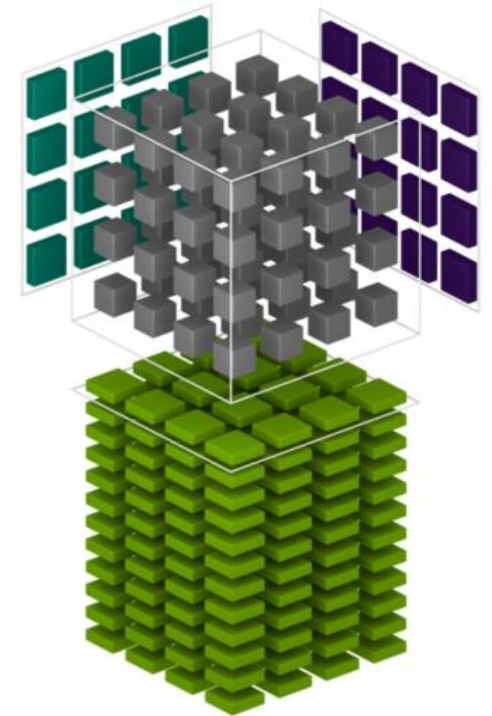
<https://hpc.nih.gov/apps/AMBER.html>

Benchmarks



Tensorcores

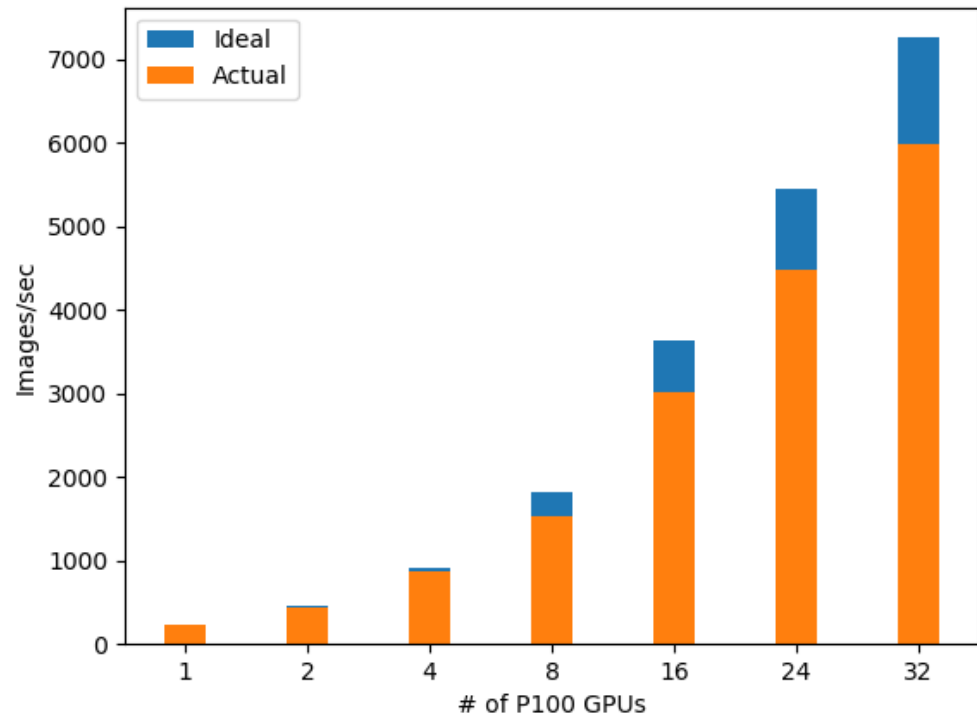
- Processing core for specialized matrix math
- 4x4 matrix multiply and accumulate units
- Suitable for deep learning and certain types of HPC
- Allow larger batch size; comparable accuracy
- Model uses tf.float16 data type to automatically take advantage of Tensor Core hardware whenever possible
- `TF_ENABLE_AUTO_MIXED_PRECISION=1`



Source: <https://developer.nvidia.com>

Horovod

- Multi-node training
- Using horovod container*
- Synthetic data
- P100 GPUs
- Optimization needed (?)



*docker://horovod/horovod:0.16.4-tf1.14.0-torch1.1.0-mxnet1.4.1-py3.6

Expansion plans

HPC Downtime from Thursday, Jul 25 through Monday, Jul 29

Jobs submitted with walltimes which extend into the downtime will remain pending until after its conclusion.



BIOWULF
HIGH PERFORMANCE COMPUTING AT THE NIH

216 standard nodes:

- Skylake architecture
- 64 cpus/node
- 384 Gb memory
- 100 Gb/s infiniband

56 GPU nodes:

- 4 Nvidia V100/node
- 32 Gb GPU memory
- Nvlink

- Phase 5 target date:
September 2019

Questions?

Contact staff@hpc.nih.gov with any questions about the NIH HPC Systems

