NYU

NYU IT PROJECT

Analyzing the Differences Between NYU Prince and Google Cloud

Ruining Sun (rs6565), Alankrith Krishnan (ak7380), Xifei Li (xl2715)

Prince Cluster VS Google Cloud

Outline

- **Ease of Access**
  - User Interface: Login Process, Transferring Files, Jupyter Notebook Access
- **Storage Comparison**
  - Disks on Prince vs Buckets and Disks
- **Hardware Comparisons**
  - CPU, GPU and TPU
- **Time Taken**
  - Deep Learning Benchmark, GAN MNIST, ImageNet, Sentiment Analysis, Glove
- **Access Times**
  - Time taken to request resources on Prince and Google Cloud
- **Cost Reduction on Google Cloud**
  - Committed Use Discount, Preemptible VM
- **Other Resources on Google Cloud**

User Interface Comparison - Login

Prince Cluster

Access from **Terminal Command Line**:

- In NYU Campus:

  ssh NetID@prince.hpc.nyu.edu

- Out of Campus:

  ssh NetID@gw.hpc.nyu.edu
  ssh NetID@prince.hpc.nyu.edu

  Or Connecting to the NYU VPN

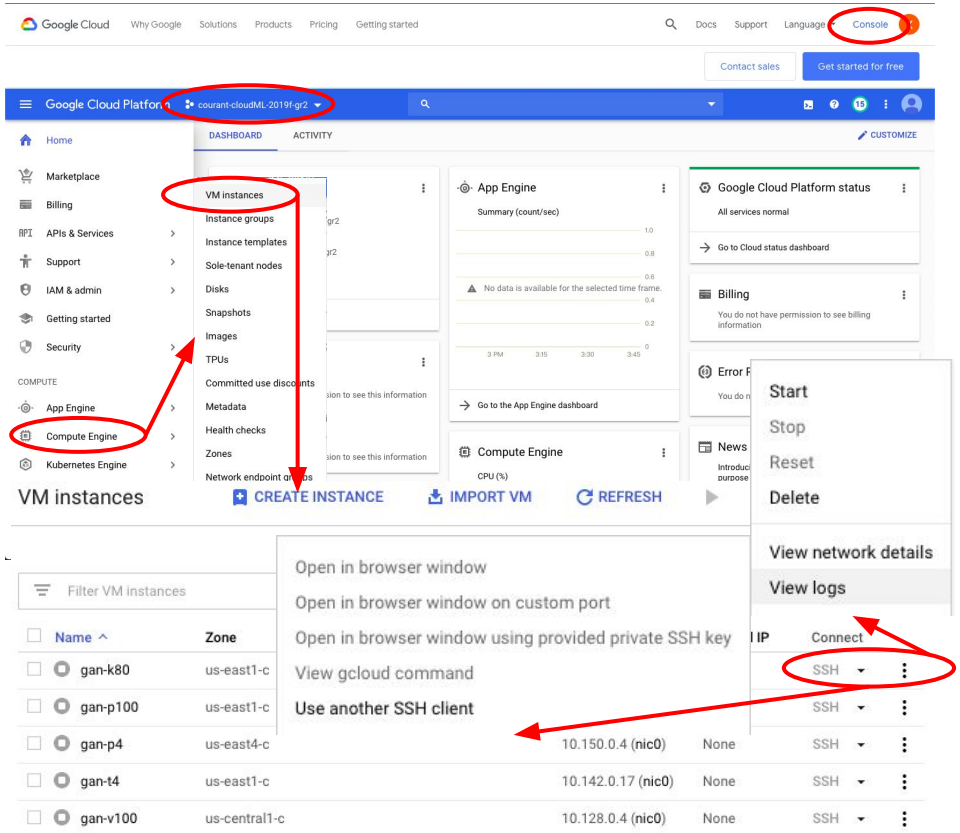
Google Cloud Platform

Access from **Terminal Command Line**:

- Download Google Cloud SDK:

  Essential tools for Google Cloud Platform

- Create Instances from Website
- gcloud compute config-ssh
- ssh <instance name>.<instance region>.<project name>



User Interface Comparison - Transfer & Jupyter Notebook

Prince Cluster

Access from **Terminal Command Line**:

Transferring Data: scp <file> NetID@prince.hpc.nyu.edu:/scratch/NetID

Running Jupyter Notebook:

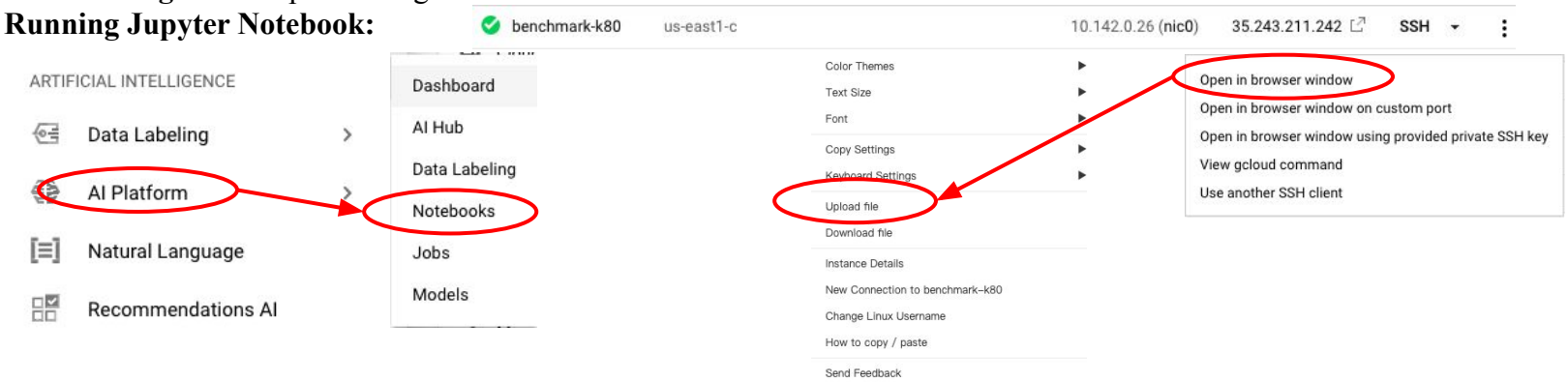
- Need a batch script -> copy run_jupyter.sbatch to scratch directory
  $ mkdir /scratch/<net_id>/myjupyter
  $ cp /share/apps/examples/jupyter/run-jupyter.sbatch /scratch/<net_id>/myjupyter
- submit run-jupyter.sbatch to job scheduler
  $ cd /scratch/<net_id>/myjupyter
  $ sbatch run-jupyter.sbatch
  Submitted batch job "job-number"

Google Cloud Platform

Access from **Terminal Command Line**:

Transferring Data: scp <file> <gcloud ssh link>:~

Running Jupyter Notebook:

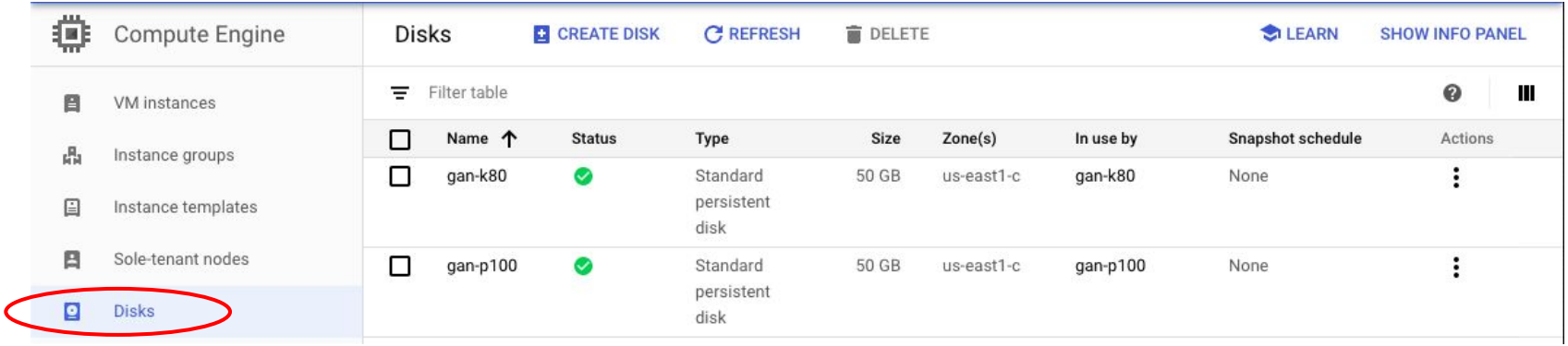


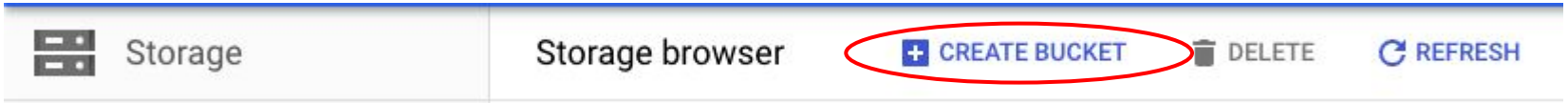
Storage Comparison

Prince Cluster - Disk

Space	Available on	Space purpose	Backed up?	Quota	Group quotas	Quota-expansion cost
/home	Prince	program-development space; storing small files you want to keep long term, e.g., source code, scripts	yes	20GB	no	N/A
/scratch	Prince, Dumbo	computational temporary/scratch space	no	5TB; inode quota: 1 million (cf. policy)	yes	N/A

Google Cloud Platform - Disk & Buckets





Hardware Comparisons

	Prince	Google Cloud
CPU architectures	Broadwell Haswell Skylake **Ivy Bridge**	Broadwell Haswell Skylake **Sandy Bridge**
GPU	NVIDIA V100 NVIDIA P100 NVIDIA K80 **NVIDIA P40**	NVIDIA V100 NVIDIA P100 NVIDIA K80 **NVIDIA P4** **NVIDIA T4**
TPU	NA	**Provided**

Time Taken

Computer vision applications

Natural language processing applications

Deep learning Benchmark

GAN MNIST

Prince		Google Cloud	
k80	**32m55.403s**	k80	36m44.027s
p100	**10m28.113s**	p100	12m18.119s
v100	**7m42.563s**	v100	8m11.290s
p40	**15m7.923s**	p4	24m48.815s
		t4	20m21.956s

- TensorFlow framework
- 50 epochs of training on MNIST dataset

IMAGENET

Prince		Google Cloud	
k80	**89m29.484s**	k80	92m54.042s
p100	54m25.404s	p100	**25m18.315s**
v100	57m8.564s	v100	**19m31.144s**
p40	55m38.370s	p4	59m25.841s
		t4	**52m55.920s**

- PyTorch framework
- 10 epochs of training with Resnet-18 on 100K images

Sentiment Analysis

Prince		Google Cloud	
k80	**34m38.850s**	k80	40m17.536s
p100	**14m2.130s**	p100	23m21.367s
v100	**19m48.365s**	v100	21m32.302s
p40	**20m37.119s**	p4	29m50.465s
		t4	24m22.108s

- PyTorch framework
- 50 epochs of training on IMDB listings

GLOVE

Prince		Google Cloud	
k80	**44m53.011s**	k80	48m47.352s
p100	**18m12.880s**	p100	19m23.915s
v100	**12m44.187s**	v100	14m21.830s
p40	**27m37.078s**	p4	44m15.027s
		t4	28m52.857s

- PyTorch framework
- 100 epochs of training on Scikit-learn (20 newsgroups)

Deep Learning Benchmark

Prince		Google Cloud	
k80	**8m8.993s**	k80	8m31.262s
p100	**3m5.580s**	p100	3m13.127s
v100	**2m10.413s**	v100	2m12.207s
p40	**3m23.709s**	p4	4m51.964s
		t4	3m49.471s

- Deep Learning Benchmark from GitHub.
- MLPerf, MLBench and others require nvidia-docker/gpu based docker
- 20 epochs of training with VGG16, ResNet152 and DenseNet161 models and two frameworks - PyTorch and TensorFlow

Access Times

- Average access times on Google Cloud is always ~2 minutes (this includes VM setup and driver installation time). If resources are not enough, the VM fails to create and you can create the VM in a different zone instead.
- Access times on Prince vary based on the time of day and the day itself

Day	Time	Avg Access Time
Normal Days	Morning	Instant
	Afternoon	~2 minutes
	Evening	~5 minutes
	Night	Instant
Heavy Days (Finals Week)	Morning	~5 minutes
	Afternoon	~13 minutes
	Evening	~25 minutes
	Night	~20 minutes

Cost Reduction on Google Cloud

- Committed Use Discounts
  - Ideal for workloads with predictable resources.
  - Paying for 1 year or 3 years.
  - Billed monthly.
  - Able to specify CPUs, memory, GPUs and storages.
  - Discount up to 57% for most resources.
- Preemptible VM
  - Much less price than normal instances.
  - Lasts only 24 hours.
  - Compute Engine might preempt these instances.
  - Ideal for fault-tolerant app.

Other Resources/Benefits of Google Cloud

- Cloud TPU's - Deep Learning Accelerators
- Superuser/sudo rights - Root Access
- Docker/Kubernetes - Google Kubernetes Engine and Docker Containers
- AI Platform - Auto ML/Predictions from Google based on your data
- Easy JupyterLab access - Create notebooks with a VM attached
- Static IP's for instances if necessary
- Premade images to use in VM's for deep learning jobs

Conclusion

	Prince	Google Cloud
Ease of Access		✓
Storage		✓
Performance	✓	✓
Access Time	✓	✓
Hardware	✓	✓
Flexibility		✓

Google Cloud is the choice in our analysis - much better User Experience, Buckets and Disk based storage, comparable performance and hardware, and the access times are constant and do not vary, and is much more flexible than Prince while offering more features.Can scale infinitely without needing to deal with the hardware directly (through the UI), making HPC admin work much easier for the future.

Questions?

Demo (if time permits)