

Waterhackweek: Cloud

20 minutes on cloud computing by myself, Rob Fatland

We have advance questions below... and please interrupt with your own questions

Objective: De-Mystify cloud use in relation to Jupyter notebooks

This content will be available (later today) in the cited GitHub repository

<https://github.com/cloudbank-project/image-research-computing-tutorial>

Program Today

Starting with <https://github.com/ChristinaB/dhsvm-opt>

From requirements.txt or environment.yml running on Binder.

Data on a public repository

- Which cloud platform can I use to run another version of this experiment? (any)
- Interactive example notebook: How to get from a repo to the Amazon cloud? (stay tuned)
- How do I configure and run this workflow on the public cloud without a JupyterHub community service? (ditto)
- What are the costs and benefits of building a custom kernel on a community jupyterhub vs an individual server in the cloud with access to Jupyter Notebooks and terminals? (horsepower)

Executive Summary in two parts

The path to your own (supercharged) Jupyter environment is not *trivial* but it is



- doable (help is available!)
 - <https://github.com/cloudbank-project/image-research-computing-tutorial>
- super flexible in relation to your data/compute needs
- available to you 24/7
- cost is proportional to use

Paying for it: Start with free credits (like \$100) and see if it is for you. If so: Set up a real account with the necessary dollars to pay for your use. Credits are available and you can always check in with me (rob5@uw.edu) and/or Cloudbank (<https://cloudbank.org>) for more support.

Demystify #1

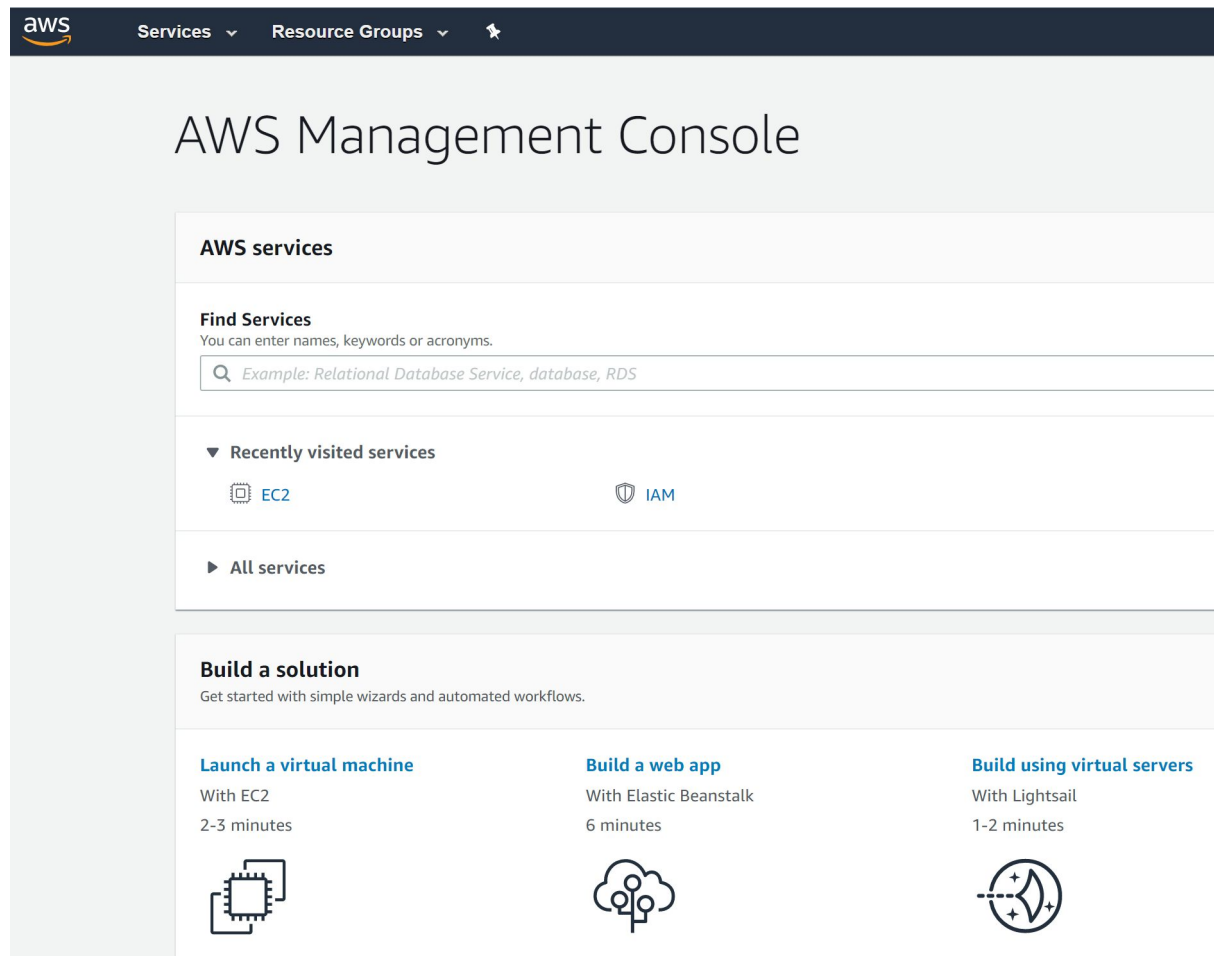
How do I interact with the cloud?

Usually first is through a browser, authenticate to a console.

Next is to learn the command line language of the cloud.

Third is to use a coding interface.

“Code is best that is invisible.”



De-mystify #2: The cloud is actually in Oregon. And Ohio. And Paris and Milan but not Prague.

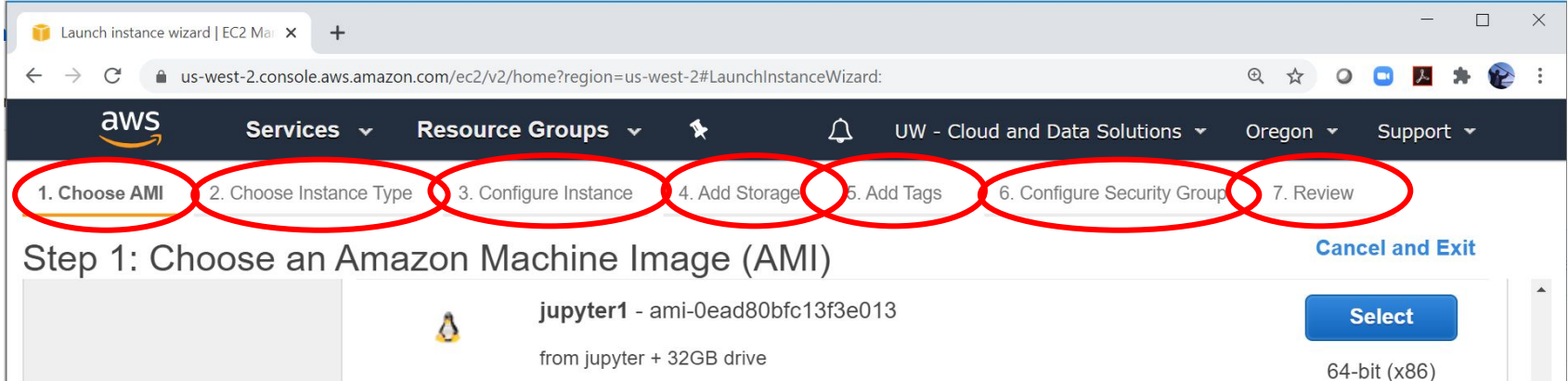
Once you know this and you are looking for something but can't find it: You might be in the wrong place on the right cloud.

Or you might be on the wrong cloud! What we are doing here today is done on AWS (Amazon Web Services) but could be done equally well on Azure or Google Cloud Platform for about the same cost.

UW - Cloud and Data Solutions ▾		Oregon ▴
US East (N. Virginia)	us-east-1	
US East (Ohio)	us-east-2	
US West (N. California)	us-west-1	
US West (Oregon)	us-west-2	
Africa (Cape Town)		af-south-1
Asia Pacific (Hong Kong)		ap-east-1
Asia Pacific (Mumbai)	ap-south-1	
Asia Pacific (Seoul)	ap-northeast-2	
Asia Pacific (Singapore)	ap-southeast-1	
Asia Pacific (Sydney)	ap-southeast-2	
Asia Pacific (Tokyo)	ap-northeast-1	
Canada (Central)		ca-central-1
Europe (Frankfurt)		eu-central-1
Europe (Ireland)	eu-west-1	
Europe (London)	eu-west-2	
Europe (Milan)	eu-south-1	
Europe (Paris)	eu-west-3	
Europe (Stockholm)	eu-north-1	
Middle East (Bahrain)		me-south-1
South America (São Paulo)		sa-east-1

De-mystify #3: It really is all about image.

The term ***image*** is cloud jargon meaning -- for all intents -- a tar file of an entire operating system including a working environment (our Jupyter notebook server) and possibly even data.





cost m5ad.12xlarge

\$2.472 per Hour

On-Demand Pricing

	vCPU	Linux/UNIX Usage
General Purpose - Current Generation		
m5ad.2xlarge	8	\$0.412 per Hour
m5ad.4xlarge	16	\$0.824 per Hour
m5ad.12xlarge	48	\$2.472 per Hour

m5ad.12xlarge

48

192

2 x 900 NVMe SSD

24 cores

192GB RAM

1.8TB SSD

De-mystify #4: Am I always paying \$2.70 / hour for my cloud computer?

Answer: No.

- Start / Stop
- SPOT market **Purchasing option** ⓘ ☐ Request Spot instances
- Development phase: Use a smaller / cheaper instance (VM)

...as cheap as pennies per hour...

key pair: A text file connecting your laptop to the cloud machine

Note: The selected key pair will be added to the set of keys authorized for this instance. Learn more about [removing existing key pairs from a public AMI](#).

Choose an existing key pair

Select a key pair

jupyter

☒ I acknowledge that I have access to the selected private key file (jupyter.pem), and that without this file, I won't be able to log into my instance.

Cancel

Launch Instances



Public IPv4 address

35.167.5.71


```
my computer> chmod 400 whwkeypair.pem
my computer> pwd
/home/kilroy
my computer> ssh -i whwkeypair.pem ubuntu@34.212.171.206
The authenticity of host '34.212.171.206 (34.212.171.206)' can't be established.
ECDSA key fingerprint is SHA256:o6eaCRVcWJEK9ntDjqzumPJHJg/mehXg09Lw9vQ7OXw.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added '34.212.171.206' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 16.04.6 LTS (GNU/Linux 4.4.0-1102-aws x86_64)
```

```
* Documentation:  https://help.ubuntu.com
* Management:    https://landscape.canonical.com
* Support:        https://ubuntu.com/advantage
```

```
* Kubernetes 1.19 is out! Get it in one command with:
```

```
sudo snap install microk8s --channel=1.19 --classic
```

```
https://microk8s.io/ has docs and details.
```

```
113 packages can be updated.
79 updates are security updates.
```

```
Last login: Wed Feb 19 20:49:34 2020 from 66.78.209.75
```

```
(base) ubuntu@ip-172-31-1-203:~$
(base) ubuntu@ip-172-31-1-203:~$
(base) ubuntu@ip-172-31-1-203:~$
(base) ubuntu@ip-172-31-1-203:~$
(base) ubuntu@ip-172-31-1-203:~$ ls
anaconda3  chlorophyll  data
(base) ubuntu@ip-172-31-1-203:~$
```

```
anaconda3 - conda/phy11 - data  
(base) ubuntu@ip-172-31-1-203:~$ (jupyter lab --no-browser --port=8889) &
```

token=7dc934bd0814cf8a425a5b058

```
my computer>  
my computer> ssh -N -f -i whwkeypair.pem -L localhost:7005:localhost:8889 ubuntu@34.212.  
my computer>
```



localhost:7005/



jupyter

Password or token:

.....

Log in

Conclusions

This presentation is an unvarnished summary of ‘what you need to know’ to operate your own Jupyter notebook server (Jupyter Lab) independent of other infrastructure. Some key takeaways...

- You choose to dial up / dial down the machine power and cost
- Setting up a pre-built image allows you to treat your work environment like a tar file
- This process applies to any cloud
- Using an ssh tunnel is one approach; remote desktop is another