

```

{
  "cells": [
    {
      "cell_type": "markdown",
      "id": "d2ab2199",
      "metadata": {},
      "source": [
        "# Lab2: Cloud Computing Services and Setup\n",
        "\n",
        "Maryam R.Aliabadi, August 7th, 2023"
      ]
    },
    {
      "cell_type": "markdown",
      "id": "16c7018f",
      "metadata": {},
      "source": [
        "\n",
        "\n",
        "## Learning objectives\n",
        "\n",
        "The goal of this lab is to launch and configure an EC2 instance for
use within your team. You will learn:\n",
        "\n",
        "- Setting up an EC2 instance\n",
        "- Logging into EC2\n",
        "- Setting up a common space in EC2\n",
        "- Setup your JupyterHub\n",
        "- Installing packages in TLJH"
      ]
    },
    {
      "cell_type": "markdown",
      "id": "efc4550e",
      "metadata": {},
      "source": [
        "## Category of services in AWS\n",
        "\n",
        "Here are the main categories of services available in AWS:\n",
        "\n",
        "- [Compute](https://aws.amazon.com/products/compute/)\n",
        "- [Storage](https://aws.amazon.com/products/storage/)\n",
        "- [Database](https://aws.amazon.com/products/databases/)\n",
        "- [Networking and Content Delivery](https://aws.amazon.com/vpc/) \n",
        "- [Security, Identity, and Compliance](https://aws.amazon.com/iam/)
\n",
\n"
      ]
    }
  ]
}

```

```
"\n",
```

```
"Having an idea of different cloud services is essential in deciding the most suitable solution for your business. While it's impossible to cover every category and service within AWS, we will be focusing on the fundamental and commonly used services that serve as the building blocks for most AWS services. Every year they add around 10 new services, so again the most important skill is to adapt to new technologies and tools.\n",
```

```
"\n",
```

```
"For a comprehensive list of services and their details, please refer to [this](https://d1.awsstatic.com/whitepapers/aws-overview.pdf) link.
```

```
Alternatively, for a simplified version of the list with brief descriptions, you may visit
```

```
[this](https://adayinthelifeof.nl/2020/05/20/aws.html) link. Please note that the latter article is from 2020.\n",
```

```
"\n",
```

```
"## Compute (EC2)\n",
```

```
"A web service that provides secure, resizable compute capacity in the cloud. It is designed to make ***web-scale computing*** easier for developers.\n",
```

```
"\n",
```

```
"### The littlest JupyterHub (TLJH)\n",
```

```
"\n",
```

```
"We are all using JupyterLab on our personal computers, but when we move to the industry, things change, and you will mostly use JupyterHub, which is a hub of Jupyter where employees log in with their username and password. We have a similar setup here at UBC known as https://ubc.syzygy.ca/. In our example, we will set up a JupyterHub so that your team members can use it to log in and create a collaborative environment. This entire JupyterHub setup is made easy by [TLJH](https://tljh.jupyter.org/en/latest/install/amazon.html), which we will be using. TLJH...\n",
```

```
"\n",
```

```
"- Greatly simplifies the deployment process.\n",
```

```
"- Enables a much faster setup/teardown.\n",
```

```
"- Provides the same ability to customize authentication, etc.\n",
```

```
"- Defines the user environment by installing packages as admin."
```

```
]
```

```
},
```

```
{
```

```
"cell_type": "markdown",
```

```
"id": "b5e8abd1",
```

```
"metadata": {},
```

```
"source": [
```

```
"### Setting up an EC2 instance\n",
```

```
"\n",
```

```
"[Setting up EC2
```

```

instance](https://ubc.zoom.us/rec/share/j17o-yumnCBgkKh6_-LaQI3Zg-dph9tAgdo
7i4AHDXMsziDLs-qj6Q4675qWVh_H.EEtm1jqD_gBe5dRK?startTime=1647906858000)
(Passcode: jSR1T!QN)\n",
  "\n",
  "You can use
[this](https://tljh.jupyter.org/en/latest/install/amazon.html) doc as a
reference for setup. You can find the minimal setup instructions below:"
]
},
{
  "cell_type": "markdown",
  "id": "f82d19b6",
  "metadata": {},
  "source": [
    "1. (Name and tags) Give your instance a name.\n",
    "2. (Application and OS Images) Choose an Amazon Machine Image
(AMI).\n",
    "3. (Instance type) Choose an Instance Type.\n",
    "4. (Key pair (login)) Select \"Create a new key pair\" if setting up
the cluster for the first time, and give it a name. Download the private
key and keep it secure. Next time, select \"Choose an existing keypair\"
and choose the key you already created.\n",
    "5. (Network Settings) Make sure to select a public (VPC) and
subnet.\n",
    "6. (Network Settings) Select or create a security group. When you
first setup create a security group\n",
    "7. (Configure storage) Use Root as the storage with the desired size
(30 GB size is perfect). Make sure to select \"General purpose SSD (gp2)\"
as the Volume Type to save costs.\n",
    "   - You should have: 1x 30 GiB gp2 Root volume (Not encrypted) \n",
    "9. (Under Advanced details) Get the configuration code from step 7 in
the [above link](https://tljh.jupyter.org/en/latest/install/amazon.html),
or copy the example below. Scroll to where it says User data and paste your
code into the User data text box. Remember to replace \"admin-user-name\"
(remove < > as well) with your jupyter username. For example, this is how
my `EC2 user data` looks like (I replaced <admin-user-name> with `gittu`.
Once my hub is setup I am going to use this username to login to my
jupyterhub."
  ]
},
{
  "cell_type": "raw",
  "id": "96b66e73",
  "metadata": {},
  "source": [
    "#!/bin/bash\n",

```

```

    "curl -L https://tljh.jupyter.org/bootstrap.py \\n",
    " | sudo python3 - \\n",
    "   --admin maryam"
  ]
},
{
  "cell_type": "markdown",
  "id": "4a1e5b1e",
  "metadata": {},
  "source": [
    "9. Click on Launch instance.\n",
    "10. Search for your EC2 instance under instances to see if it's
running. Give it 15 - 20 minutes as it takes time to set up JupyterHub even
if it shows running. You can also check system logs during this time."
  ]
},
{
  "cell_type": "markdown",
  "id": "9a30d836",
  "metadata": {},
  "source": [
    "#### Set up public IPV4 address\n",
    "1. On the EC2 dashboard, go to Elastic IPs\n",
    "2. Allocate Elastic IP Address\n",
    "3. Make sure the Network Boarder Group is ca-central-1\n",
    "4. Click on Allocate\n",
    "5. Select the newly created IP address and click on Actions ->
Associate Elastic IP Address. Select your Instance from the dropdown, and
choose a Private IP address.\n",
    "7. On the EC2 dashboard, go to Instances and select your Instance.\n",
    "8. Click the subnet ID\n",
    "9. Select your subnet, then click on Actions -> Edit subnet
settings\n",
    "10. Check Enable auto-assign public IPv4 address"
  ]
},
{
  "cell_type": "markdown",
  "id": "212d9c34",
  "metadata": {},
  "source": [
    "### Interacting with the EC2 instance \n",
    "\n",
    "Now we have our EC2 instance ready let's see some things that we can
do; You can Stop, Terminate, Start the instance and Monitor System logs.
You can also create an image of your instance and scale it up if you need

```

more processing power.\n",

"Assuming someone from your team has set up this EC2 instance, there are several questions we need to consider. In the following three sections, you will find answers to these questions:\n",

"\n",

"- How can you and your team members log in to the EC2 instance?\n",

"- How can you make directories accessible to your team members\n",

"- How can team members log in to the JupyterHub?\n",

"\n",

"In the following, you will find the answer to the above questions:"

]

},

{

"cell_type": "markdown",

"id": "8be4b904",

"metadata": {},

"source": [

Logging into the EC2 instance \n",

"\n",

"- [Logging into your EC2

instance](https://ubc.zoom.us/rec/share/j17o-yumnCBgkqh6_-LaQI3Zg-dph9tAgdo7i4AHDXMszIDLs-qj6Q4675qWVh_H.EEtm1jqD_gBe5dRK?startTime=1647908464000)

(Passcode: jSR1T!QN)"

]

},

{

"cell_type": "markdown",

"id": "e33deaa4",

"metadata": {},

"source": [

"You can use different ways to connect to an EC2 instance and use it.

Check the ***\`Connect\`*** button to determine how to connect to the instance. Select the ***\`SSH Client\`*** tab and follow the instructions there. "

]

},

{

"cell_type": "markdown",

"id": "5e7dd6c8",

"metadata": {},

"source": [

"The person who sets up the EC2 instance (Maryam in this case) will only have access to the server, as that person will have the private key. Actually, Maryam has the private key of the root user `ubuntu` (the one I created when setting up EC2 instance and downloaded it at the last stage) and is logging into the EC2 instance as `ubuntu`. If you have not yet

logged into the EC2 instance, please check out the last step in the EC2 creation process. The following is how you logged into the EC2 instance:"

```
]
},
{
  "cell_type": "raw",
  "id": "36e2d2d8",
  "metadata": {},
  "source": [
    "ssh -i <path_to_the_privatekey> <username>@<ec2_hostname>\n",
    "\n",
    "# example\n",
    "ssh -i ~/Downloads/mytestkey.pem
ubuntu@ec2-54-555-555-555.compute-1.amazonaws.com"
  ]
},
{
  "cell_type": "markdown",
  "id": "8957aa89",
  "metadata": {},
  "source": [
    "As you can see, I logged into the instance as Ubuntu. If others need to access the instance, a new user account must be created, and a keypair should be generated for them to use.\n",
    "\n",
    "For example, imagine a team with 3 members: Maryam, Maria, and Nicole. Maryam creates an EC2 instance and therefore has exclusive access to it, using the private key. But how can Maria and Nicole access it? It is not advisable for Maryam to share his private key with others. Instead, Maryam needs to create individual user accounts for Maria and Nicole and register their public keys into the system.\n",
    "\n",
    "Here is the conversation that happens:\n",
    "\n",
    "Maryam: Hey Maria - I have created an EC2 instance for us to use. I want to give you access to it. Can you provide me with your public key?\n",
    "Maria: Sure, I don't have anything at the moment but I will create a keypair and provide you with the public key. \n",
    "Maryam: Great, Once I have it, I will add your public key to the server so you can access it.\n",
    "\n",
    "I found [this document in AWS](https://aws.amazon.com/premiumsupport/knowledge-center/new-user-accounts-linux-instance/) to set up user accounts for Maria and Nicole. Make sure to follow the instructions carefully. Suppose you encounter any issues or confusing steps. In that case, you can refer to the video that I created
```

(although it is not necessary to watch it to complete the setup process successfully)."

```
]
},
{
  "cell_type": "markdown",
  "id": "7c7626f9",
  "metadata": {},
  "source": [
    "### Adding a new user to your EC2 instance\n",
    "\n",
    "1.   Connect to your Linux instance using SSH.\n",
    "\n",
    "2.   Use the adduser command to add a new user account to an EC2
instance (replace new_user with the new account name). The following
example creates an associated group, home directory, and an entry in the
/etc/passwd file of the instance.\n"
  ]
},
{
  "cell_type": "raw",
  "id": "a176d389",
  "metadata": {},
  "source": [
    "$ sudo adduser new_user"
  ]
},
{
  "cell_type": "markdown",
  "id": "044dd092",
  "metadata": {},
  "source": [
    "The home directory might not be created by default in some
configurations. Verify that the home directory was created before
continuing."
  ]
},
{
  "cell_type": "markdown",
  "id": "ddf2cd6e",
  "metadata": {},
  "source": [
    "Note: If you add the new_user to an Ubuntu instance, then include the
--disabled-password option to avoid adding a password to the new account:"
  ]
},
}
```

```

{
  "cell_type": "raw",
  "id": "0af6bafe",
  "metadata": {},
  "source": [
    "$ sudo adduser new_user --disabled-password"
  ]
},
{
  "cell_type": "markdown",
  "id": "37f59d24",
  "metadata": {},
  "source": [
    "3.    Change the security context to the new_user account so that
folders and files you create have the correct permissions:"
  ]
},
{
  "cell_type": "markdown",
  "id": "117c102c",
  "metadata": {},
  "source": [
    "$ sudo su - new_user"
  ]
},
{
  "cell_type": "markdown",
  "id": "e25e04eb",
  "metadata": {},
  "source": [
    "Note: When you run the sudo su - new_user command, the name at the top
of the command shell prompt changes to reflect the new user account context
of your shell session."
  ]
},
{
  "cell_type": "markdown",
  "id": "b692cc58",
  "metadata": {},
  "source": [
    "\n",
    "4.    Create a .ssh directory in the new_user home directory:"
  ]
},
{
  "cell_type": "raw",

```

```

    "id": "bf1badcd",
    "metadata": {},
    "source": [
      "$ mkdir .ssh"
    ]
  },
  {
    "cell_type": "markdown",
    "id": "b3c31f9a",
    "metadata": {},
    "source": [
      "5. Use the chmod command to change the .ssh directory's permissions
to 700. Changing the permissions restricts access so that only the new_user
can read, write, or open the .ssh directory"
    ]
  },
  {
    "cell_type": "raw",
    "id": "e0eacb84",
    "metadata": {},
    "source": [
      "$ chmod 700 .ssh"
    ]
  },
  {
    "cell_type": "markdown",
    "id": "58f09701",
    "metadata": {},
    "source": [
      "6. Use the touch command to create the authorized_keys file in the
.ssh directory:"
    ]
  },
  {
    "cell_type": "raw",
    "id": "4df112ae",
    "metadata": {},
    "source": [
      "$ touch .ssh/authorized_keys"
    ]
  },
  {
    "cell_type": "markdown",
    "id": "0241e6c5",
    "metadata": {},
    "source": [

```

"7. Use the chmod command to change the .ssh/authorized_keys file permissions to 600. Changing the file permissions restricts read or write access to the new_user."

```
]
},
{
  "cell_type": "raw",
  "id": "3b98041f",
  "metadata": {},
  "source": [
    "$ chmod 600 .ssh/authorized_keys"
  ]
},
{
  "cell_type": "markdown",
  "id": "14006eea",
  "metadata": {},
  "source": [
    "\n",
    "After the successful completion of the process, nicole and maria will
be able to login like below:\n"
  ]
},
{
  "cell_type": "raw",
  "id": "acb9e551",
  "metadata": {},
  "source": [
    "# example nicole\n",
    "ssh -i <path_to_nicole_private_key>
nicole@ec2-54-555-555-555.compute-1.amazonaws.com\n",
    "# example maria\n",
    "ssh -i <path_to_maria_private_key>
maria@ec2-54-555-555-555.compute-1.amazonaws.com"
  ]
},
{
  "cell_type": "markdown",
  "id": "9a801bce",
  "metadata": {},
  "source": [
    "\n",
    "- You can also watch
[this](https://ubc.zoom.us/rec/share/VnhSjTmqSA0fLFzfFqAWht_zi1TXw6CkZqVnzf
QfjoBdNjShU4o1SgE6fkUcj1Q.4BJufU87vGu60nmR?startTime=1649177849000) short
video (Passcode: `r$*A10%&`) to learn some basics, which may help you
```

```

better understand the Unix world. \n"
]
},
{
  "cell_type": "markdown",
  "id": "a63e43ad",
  "metadata": {},
  "source": [
    "\n",
    "### Setting up a common space in EC2\n",
    "\n",
    "[Setting up a common space in
EC2](https://ubc.zoom.us/rec/share/ZjuksMDLRv09jw0-HFLTid-7DA2jsJ_d5yjjzjaSQ
CqrnSVmghzd3qrjKQgqoaFjm.49XxL5ueW7EDClH4?startTime=1679605324000)
(Passcode: j4DrAf@!).\n",
    "\n",
    "After creating accounts for your team members, you'll notice folders
for each user in the \"/home\" directory. However, the person who set up
the EC2 instance (in my case maryam) doesn't have a user account. Instead,
Maryam is using the ubuntu account (and has access to the ubuntu private
key). Also, here ubuntu is the `root user`.\n"
  ]
},
{
  "cell_type": "raw",
  "id": "27efbe33",
  "metadata": {},
  "source": [
    "maria@ip-172-31-21-101:~$ cd /home\n",
    "maria@ip-172-31-21-101:/home$ ls -ltrh\n",
    "total 12K\n",
    "drwxr-xr-x 2 nicole nicole 4.0K Mar 21 07:32 nicole\n",
    "drwxr-xr-x 5 ubuntu ubuntu 4.0K Mar 21 07:39 ubuntu\n",
    "drwxr-xr-x 5 maria maria 4.0K Mar 21 07:39 maria"
  ]
},
{
  "cell_type": "markdown",
  "id": "c40419d0",
  "metadata": {},
  "source": [
    "Setup a common data folder to download data, and this folder should be
accessible by all users in the JupyterHub. Following commands, make a
folder and make it accessible to everyone. \n"
  ]
},
}

```

```

{
  "cell_type": "code",
  "execution_count": null,
  "id": "37caa351-e178-4ed1-aae8-0fbbd641c04b",
  "metadata": {},
  "outputs": [],
  "source": [
    "sudo mkdir -p /srv/data/my_shared_data_folder\n",
    "sudo chmod 777 /srv/data/my_shared_data_folder/"
  ]
},
{
  "cell_type": "markdown",
  "id": "a16d70f4",
  "metadata": {},
  "source": [
    "\n",
    "#### Setup your JupyterHub\n",
    "\n",
    "[Setup a
notebook](https://ubc.zoom.us/rec/share/ZjuksMDLRv09jw0-HFLTID-7DA2jsJ_d5yj
zjaSQCqrnSVmghzd3qrjKQgqoaFjm.49XxL5ueW7EDClH4?startTime=1679605620000)
(Passcode: `j4DrAf@!`).\n",
    "\n",
    "We installed our TLJH while [Setting up EC2
instance](https://ubc.zoom.us/rec/share/j17o-yumnCBgkhh6_-LaQI3Zg-dph9tAgdo
7i4AHDXMsziDLs-qj6Q4675qWVh_H.EEtm1jqD_gBe5dRK?startTime=1647906858000)
(Passcode: jSR1T!QN). Currently, we only have one admin user - the person
who set up the EC2 instance - and the admin username is what you provided
in your `EC2 user data` (in my case maryam). We will use that username to
log in to JupyterHub. \n",
    "\n",
    "1) In the EC2 dashboard, select your instance and then in the
description panel below, find where it says Public IPv4 address and click
on `Open address`. This will open up your browser for JupyterHub. (NOTE:
Make sure it is not automatically prepended with `https://`, we want to
access it over `http://`)\n",
    "\n",
    "2) Enter your username (what you provided in your `EC2 user data`) and
use a strong password & remember it, as what you enter here will be the
admin password.\n"
  ]
},
{
  "cell_type": "markdown",
  "id": "e18c9daa",

```

```

    "metadata": {},
    "source": [
        "We want to add other team members as admin users to the jupyterhub.
[Here](https://tljh.jupyter.org/en/latest/howto/admin/admin-users.html?high
light=adding%20user#adding-admin-users-from-the-jupyterhub-interface) are
the instructions that you can use. I strongly recommend you check it out,
but I will list down the main points."
    ]
},
{
    "cell_type": "markdown",
    "id": "997d4a25",
    "metadata": {},
    "source": [
        "In your JupyterHub, go to \"Control Panel\" --> \"admin.\" Here add
other members of your group, use their names and make them admins.\n",
        "\n",
        "3) Check if other members can log in to the JupyterHub from their
machines by giving them the URL to connect. Step 2 is applicable here for
other members."
    ]
},
{
    "cell_type": "markdown",
    "id": "de3228ee",
    "metadata": {},
    "source": [
        "\n",
        "##### Install packages in TLJH \n",
        "\n",
        "You have to install the packages that are needed. You ***should***
refer this TLJH [document](
https://tljh.jupyter.org/en/latest/howto/env/user-environment.html).\n",
        "\n",
        "Don't forget to add option -E. This way, all packages that you install
will be available to other users in your JupyterHub. For example, open up
the terminal from the jupyter, and install packages that you need;"
    ]
},
{
    "cell_type": "code",
    "execution_count": null,
    "id": "f6360a48-af53-4e08-890f-1e9d08969cc1",
    "metadata": {},
    "outputs": [],
    "source": [

```

```

    "sudo -E pip install pandas\n",
    "sudo -E pip install s3fs\n",
    "sudo -E pip install pyarrow"
  ]
},
{
  "cell_type": "markdown",
  "id": "4a34e81f",
  "metadata": {},
  "source": [
    "You surely want to install the packages mentioned above. Also, install
any other packages that are required for your data wrangling. Please note
that the setup and server we have are too small for a fully-fledged
collaborative space. Therefore, please only install what is necessary, or
it may lead to a system crash.\n"
  ]
},
{
  "cell_type": "markdown",
  "id": "1fe9f1b2",
  "metadata": {},
  "source": [
    "### How to create an image ?\n",
    "\n",
    "Well, at this point, you have set up your EC2 instance and TLJH. You
have also installed all the necessary packages that you need for doing your
analysis in jupyter. You might have also installed some other packages in
your Ubuntu OS. Whatever it is, if you want to obtain a copy of your
computer with everything that it has, then you can make use of EC2 images.
You can create the image by ***selecting the EC2 instance --> Click Actions
--> Image and Templates --> Create Image***.\n"
  ]
},
{
  "cell_type": "code",
  "execution_count": null,
  "id": "e48fe0e3-32a5-4011-8bca-42ac96f8752c",
  "metadata": {},
  "outputs": [],
  "source": []
}
],
"metadata": {
  "kernelspec": {
    "display_name": "Python 3 (ipykernel)",
    "language": "python",

```

```
  "name": "python3"
},
"language_info": {
  "codemirror_mode": {
    "name": "ipython",
    "version": 3
  },
  "file_extension": ".py",
  "mimetype": "text/x-python",
  "name": "python",
  "nbconvert_exporter": "python",
  "pygments_lexer": "ipython3",
  "version": "3.11.5"
},
"vscode": {
  "interpreter": {
    "hash":
"6e9e0baa62560f8a3b402c12d339bdad33c58a25305700ec7e7682c0b6251f68"
  }
},
"nbformat": 4,
"nbformat_minor": 5
}
```