Lab Work 1 - Cloud Computing

Charlotte Laclau
prenom.nom@univ-st-etienne.fr

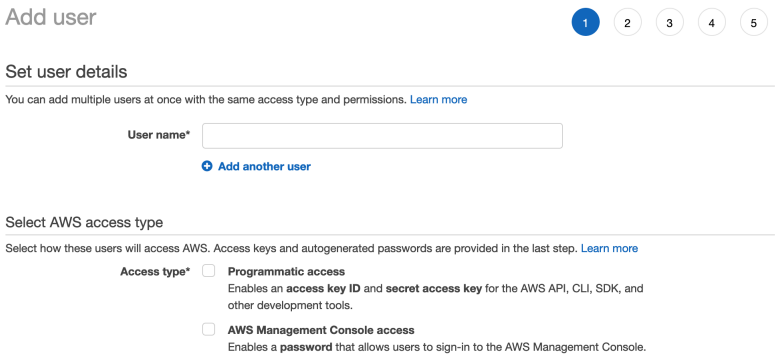
TSE - Big Data : management and analysis

Identity and access management (IAM)

This section provides you with the main steps to create an additional IAM user, and then add the user to an IAM group with administrative permissions. You can then access AWS using a special URL and the credentials for this new IAM user.

Task 1 : Create an IAM user and a group

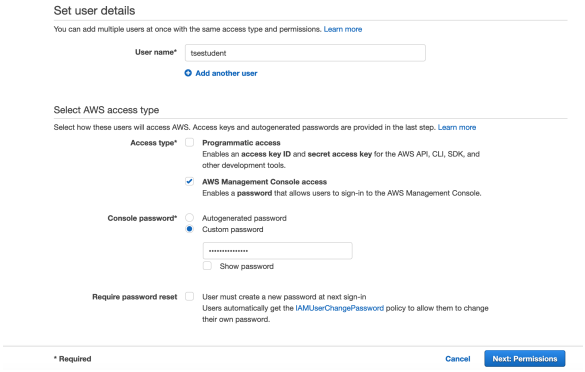
1. In the AWS Management Console, on the **Service** menu choose **IAM**.
2. From the navigation pane, click **Users**, and then **Add User**. The New User dialog will appear as shown here:



3. The **User name** field is required. For this example, we will all use the same user name : tsestudent.
   A valid user name can consist of letters, digits, and the following characters: plus (+), equal (=), comma (,), period (.), at (), underscore (_), and hyphen (-). The name is not case sensitive and can be a maximum of 64 characters in length.

   For **Access type**, select the box next to AWS Management Console access. By doing so, you are limiting the tsestudent account rights and privileges to reach solely the AWS console.
   Additional options will appear. For **Console Password**, select the choice **Custom password**. Enter a password you can remember into the textbox that appears.

   For **Require password** reset, **unselect** the box next to User must create a new password at next sign-in.
   In the end you should have something like this :

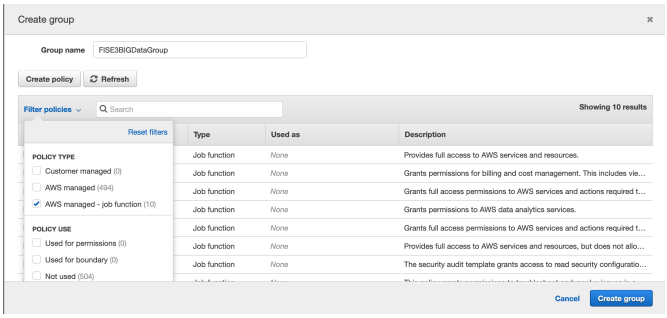


4. Click **Next: Permissions to continue**. A new dialog box should appear. Click on **Create group**. A Create group dialog will popup at this point.

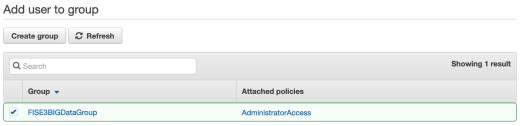
   The **Group name** field is required. For this example, we will all use the same group name, namely FISE3BIGDataGroup.

   What is a group ? How do you manage group membership / permissions ?

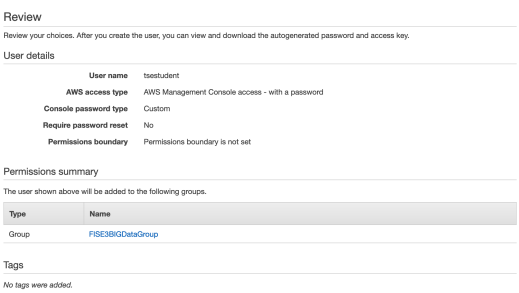
5. From the Filter menu choice, change the filter type from Policy type to Job Function as shown here:



6. In the policy list, select the check box for **AdministratorAccess.**
7. Click **Create group** to continue this process. You have now successfully created a group !



8. Click on **Next : Tags**. We will skip tags and then click on **Next : Review**. Check that you have the same info as shown bellow.



9. Click on **Create User**

Task 2 : using the IAM User

Why using the IAM User, rather then your root account ?

Connect to your AWS account using the IAM User created in the previous part.

Amazon Elastic Compute Cloud (EC2)

This section provides you with a basic overview of launching, resizing, managing, and monitoring an Amazon EC2 instance. Amazon EC2 is a web service that provides resizable compute capacity in the cloud.

What is the pricing policy of Amazon EC2?

By the end of this section, you will be able to :

- Launch a web server with termination protection enabled
- Monitor your EC2 instance
- Modify the security group that your web server is using to allow HTTP access
- Resize your EC2 instance to scale
- Terminate your EC2 instance

Task 1 : launch your first EC2 instance

1. In the AWS Management Console, on the **Services** menu click on **EC2** (Compute)
2. Before you launch the instance, you will create a **Security Group**.

   What is a Security Group ?

   - Click on **Security Groups** on the left list.
   - Click on **Create Security Group**
   - Name the security group *SSH* and add a brief description
   - Add a security rule by clicking on **Add Rule**. Specify the rule **type** to be *SSH* and set the **source** to *anywhere*
   - Click on **Create**.
3. Go back to the EC2 Dashboard, and click on **Launch Instance**
4. Select the image to be *Amazon Linux*.

   What is an AMI ? How many AMI's are available in AWS

5. Select the instance type to be **t2.micro**

   What are the characteristics of this instance ?
   What does EBS stand for ?
   What type of EBS volume is your AMI based on ?
6. Click on **Next**
7. Look at the different options and reply to the following questions
   - What is CloudWatch ? Why is it useful ?
   - True or False : monitoring is only available if you pay additionnal fees
   - What is the role of the tenancy option ?
8. Without selecting any boxes, click on **Next**

9. This page allows you to add storage capacity. Do not change anything and click on **Next**
10. In this step, add a **tag** to your instance (i.e. key : lab1, value: webserver). Tagging is useful to
    - Name the instance and what is the task undertaken by it.
    - Keep track of costs of the instances, their health, and status
11. Click on **Next**.
12. You are now going to add a security group to your instance
    - Choose the option **Select an existing security group**
    - Select the group that you created before (SSH)
13. Click on **Review and Launch**
14. Click on **Launch**
15. Before you are able to launch the instance, you need to create a new key pair.
    - Select the option **Create a New Key Pair** and name the Key Pair `mykey`.
    - During the key-pair creation process, you download a file called `mykey.pem`.

      You will always need this key to log into your instance. If you lose it, your instance becomes useless. What is a key pair ?
16. Click on **View Instances**.

Task 2 : connect to your instance

Now you will connect to your instance. First using SSH, then using the web (HTTP).

1. SSH Connection
    - Click on **Connect** displayed at the top of the window.
    - Follow the intructions given in the window.

      For **Linux** and **MacOS** the only thing you need to do is change the access rights of `mykey.pem` so that only you can read the file. To do so, run `chmod 400 mykey.pem` in the terminal.

      If you work under **Windows**, you will need to use PuTTY as Windows doesn't shop an SSH client. PuTTY comes with a tool called PuTTYgen that can convert the `mykey.pem` file to `mykey.ppk`.
2. SFTP Connection

   To create an SFTP connection, you need an SFTP client. you can use fileZilla. It works on Windows and Mac, etc. If you are have a windows computer, you can also use WinSCP.
    - For fileZilla, you need to convert the key from .pem to .ppk extension. To proceed, open a terminal
        - brew install putty
        - puttygen lab1Instance.pem -o lab1Instance.ppk
    - Go in **Filezilla** → **Settings**
    - From the left list click on **SFTP**
    - Click on **Add a key file** and select the .ppk key associated to your EC2 instance. Click on **OK**
    - Now, go to **Files** → **Site Manager**
    - Click on **New Site**. Name your site to be **amazon-ec2**
    - You should provide the host. To find it, go on EC2 list of your instances. If you select the instance, check the **Desription** displayed at the bottom. Copy the value of **public DNS** and paste it to the host field of fileZilla.

- You also need to fill the **User** field. Usually the user name of an Amazon Linux instance is **ec2-user**.
- Click on **Connect**. Now, using fileZilla explorer, you can navigate through the file system of your instance (do not close the connection, we will use it later).

Task 3 : turn the instance into a Web Server

In this part, you will see how you can easily provide a web interface to your instance.

1. From EC2 panel, choose your instance. Go down to the Description section. Copy the Public IP address. Paste it in your navigator. The navigator will not be able to open the instance. This is due to 2 reasons
   - No web server software is installed on your instance (e.g. Apache).
   - The security group controlling the access to your instance does not allow HTTP access.
2. **Install Web Server**. Before you start, check the monitoring tab. What do you observe ?
   - SSH your instance, following the steps from Task 2
   - Install Apache HTTP server using : **sudo yum install httpd**
   - Start Apache : **sudo service httpd start**
3. Configure Security Group
   - Goback to **Security Groups** under the Network & Security section.
   - Choose the security group that you created.
   - Click the button **Actions** $\rightarrow$ **Edit inbound rules**
   - Click on **Add Rule**. Choose HTTP as a **type**, select the **source** to be anywhere.
   - Click on **Save**.
   - Refresh the page on which you pasted the IP address of your instance
4. It works but as you see, there is not homepage. Create a HTML home page. Use fileZilla to copy your page into your instance. It should go into var/www/html/.
5. Check your VMs logs (through the management console). In the Action menu, choose Instance Setting $\rightarrow$ Get System Log. The log contains all log messages that would be displayed on the monitor of your machine if you were running it on premises.

   This is a good tool for debugging a virtual machine by watching out for any log messages stating that an error occured during startup for instance.

Task 4 : Load Balancer

Create a load balancer (ELB) for your infrastructure. Details how you proceeded with screenshots.

Task 5 : Estimate the price of your infrastructure

Part of evaluating AWS is estimating cost. To proceed, use the AWS Simple Monthly Calculator.

Before going to the next part, switch off your instance.

Amazon Simple Storage Service (S3)

This section introduces you to Amazon Simple Storage Service (S3) using the AWS Management Console. Amazon S3 allows to store and retrieve any amount of data at any timen from anywhere on the web.

By the end of this section, you will able to :

- Create a bucket on Amazon S3
- Add an object to your bucket
- Manage access permissions on an object
- Create a bucket policy
- Use bucket versioning

In 5 lines, explain the pricing policy of Amazon S3 (standard).
What is the difference between Amazon EBS and Amazon S3 ?

Task 1 : create a bucket

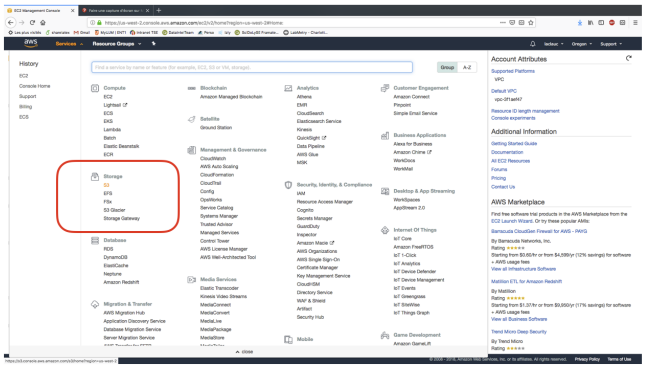
Every object in Amazon S3 is stored in a bucket.

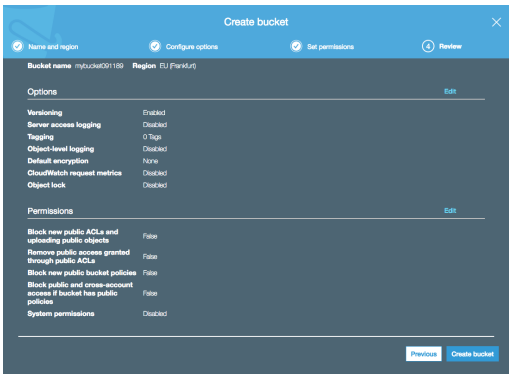
1. In the **AWS Management Console**, on the **Services** menu click **S3** (in Storage)
2. Click on **Create bucket**, then configure your new bucket as follows :
   - **Bucket name** : mybucketNumber
   - Replace **NUMBER** with a random number
   - Leave **Region** at its default value

   Selecting a particular region allows you to optimize latency, minimize costs, or address regulatory requirements. Objects stored in a region never leave that region unless you explicitely transfer them to another region.

   What is the role of the **Copy settings from an existing bucket** option ?
3. Click **Next** then configure
   - **Versioning** : select *Keep all versions of an object in the same bucket*
   - Click next

   What is the role of the **Versioning** option ?
4. At the **Set permissions** screen :
   - De-select : *Block new public ACLs and uploading public objects*
   - De-select : *Remove public access granted through public ACL's*
   - De-select : *Block new public bucket policies*
   - De-select : *Block public and cross-account access if bucket has public policies*
   - Click **Next**
5. Click **Create bucket**





Task 2 : upload an object to the bucket

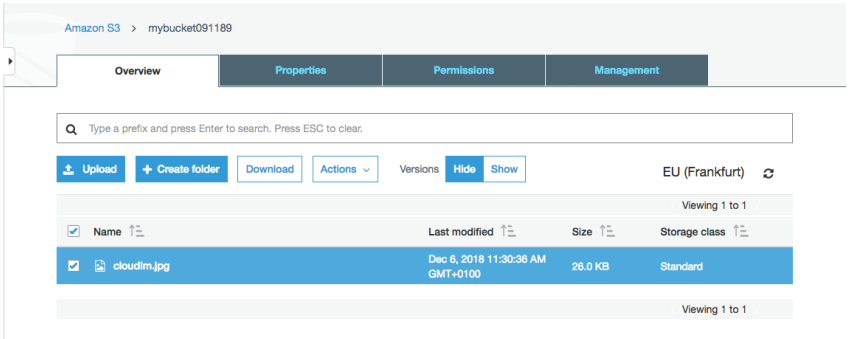
Now that you have created a bucket, you are ready to store objects. An **object** can be any kind of files : a text file, a photo, a video, a zip file, etc. When you add an object to Amazon S3, you have the options of including **metadata** with the object and setting **permissions** to control access to the object.

Let's start uploading an object to your S3 bucket.

1. Download the cloudlm.jpg file to your computer.
2. In the **S3 Management console**, click on the bucket that you have created in Task 1.
3. Click on **Upload**

   This launches an upload wizard that will assist you in uploading files. Using this wizard you can upload files, either by selecting them from a file chooser or by dragging them to the S3 window.
4. Ath the **(1) Select files** dialog box, click on **Add files** then configure :
   - Browse to and select the cloudlm.jpg file that you downloaded
   - Click upload

You can watch the progress of the upload from within the Transfer panel at the bottom of the screen. Since this is a very small file, you might not see the transfer. Once your file has been uploaded, it will be displayed in the bucket.



Task 3 : make your object public

Now, you are going to configure permissions on your object so that it is publicly accessible.

First, let us attempt to access the object to confirm that it is private by default.

1. Select the cloudlm.jpg file in your bucket. A small window appear and it contains three section : **Overview**, **Properties** and **Permissions**.

2. Copy the **S3 Link** displayed at the bottom of the **Overview section**.

   The link should look similar to this: *https://s3.eu-central-1.amazonaws.com/mybucket091189/cloudIm.jpg*

3. In a new browser tab, paste the link into the address field, then press enter.

   What do you see on your screen ? Why ?

4. Keep this browser tab open, but returns to the web browser tab with the S3 Management Console.
5. Now, click the **Permissions** section, then configure
   - Under the **Public Access** section, select *Everyone*.
   - Select *Read object*
   - Click on **Save**
6. Return to the browser tab with the S3 Link, and refresh the page.

   What do you see on your screen

In this task, you granted read access only to a specific object. If you wish to grand access to an entire bucket, you would use a **Bucket Policy**.

Task 4 : create a Bucket Policy

A **Bucket Policy** is a set of permissions associated with an Amazon S3 bucket. It can be used to control access to a whole bucket or to specific directories within a bucket.

You will now upload a new file.

1. Follow step 1-2 (upload process) from Task 3 with a new image.
2. Copy the **S3 Link** associated to this new image in a new web browser tab.

   What do you see on your screen ? Why ?

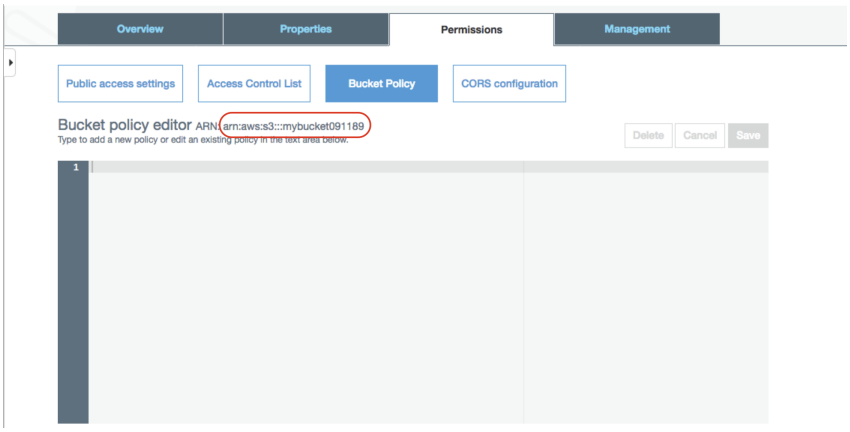
3. Back in your S3 Management Console, click the name of your bucket at the top of the window.

   You should see a list of the objects in your bucket.

4. Click on the **Permissions** tab
5. In the **Permissions** tab, click on **Bucket Policy**.

   A blank **Bucket policy editor** is displayed. Bucket policies can be created manually, or they can be created with the assistance of the **AWS Policy generator**.

6. Copy the **ARN** (Amazon Resource Name) of your bucket to the clipboard. It is displayed at the top of the policy editor.



What is an ARN ?

7. Click on the **Policy generator** link at the bottom of the page. A new browser tab will open with the AWS Policy Generator. Configure the following :
   - **Select type of policy** : *S3 Bucket Policy*
   - **Principal** : ***

     This means that anyone will be able to perform the actions in the policy
   - **Actions** : *GetObject*

     The *GetObject* action grants permission of objects to be retrieved from Amazon S3.
   - **Amazon Resource Name (ARN)** : paste the ARN that you previously copied. At the end of the ARN, append : /* . The ARN should look similar to **arn:aws:s3:::lab-xxx/***

     Adding /* to the end of the bucket name allows the policy to apply to all objects within the bucket.
8. Click on **Add Statement**
9. Click on **Generate Policy**

   Your bucket policy is now displayed.
10. Copy the policy to your clipboard
11. Paste the bucket policy into the **Bucket Policy editor**.
12. Click **Save**
13. Go back in the tab where you pasted the link for this new image and refresh the page. You should see the picture.