mini project

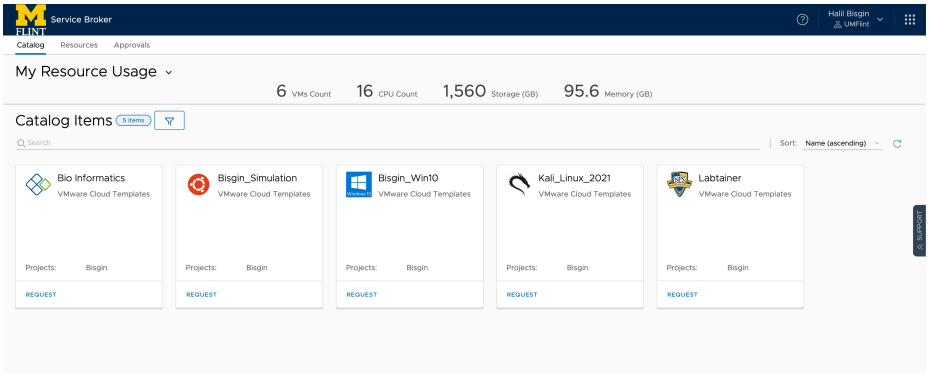
Advanced Data Mining: Algorithms and Applications

Due on Mar 30, 11:59PM

Important

- This is a follow-along assignment for which you can find the details at: `https://github.com/ML4CYB/DOH-Lab/blob/main/LAB%20DoH%20vs%20non-DoH%20Packet%20Capture%20Manual.pdf`
- This is the tar file you'll need inside the lab: `https://github.com/ML4CYB/DoH_Labtainer_tar_file`. No need to download here.

1. This assignment aims to show you how you can collect or produce synthetic data for your research and make use of it through a machine learning algorithm inside a "sandbox", i.e., docker containers.
2. Since the exercise is focused on cybersecurity data, you're not expected to answer those questions in the lab which are basically related to networks. However, reading through them will help you understand the nature of the data set and give an understanding of a real world problem.
3. Before you start working on your assignment, you need to have a Virtual Machine (VM) set up on which Labtainer will be hosted. To do so, please visit `https://vcloud.umflint.edu/` first where you will have the following catalog. There, you will click on REQUEST a Labtainer, which is the name of the VM.



4. Please refer to the next page to see how you can access your VM after you make your request. The same information can be found in the link above.
5. Once you have the VM, it will be an Ubuntu on which you can start and stop labtainer labs. However, for this specific lab, you'll have to go through a certain way described in the manual to start. Then the same document will tell you how to stop.
6. Meantime, you may need run `sudo apt-get install labtainer` to your VM on the command line. (Command line is nothing but the Terminal you can access on the toolbar which will appear on the left of your desktop when you start your Ubuntu)
7. In case you need to see some examples and learn more about Labtainers, you can see the manual attached to this assignment.

Deploying Virtual Machines to the UM Flint vCloud Environment

You will be using the UM Flint vCloud environment to deploy your virtual machine (VM). You will interact with two parts of a service broker to manage your deployments, the catalog and deployment sections.

Service Broker	
Catalog	**Deployments**
This section contains all of the VM templates your computer science instructors have given you access to. If you are enrolled in multiple classes that use this system you will see templates for all of your classes here.	This section contains all of your active and pending deployments. You can view information about and interact with your deployed virtual machines here.

VPN Required

Remote access to the vCloud environment is only available over the UM Flint VPN. On campus systems can directly connect. Configuring a VPN connection on your system is outside the scope of this document, more information can be found here: https://support.umflint.edu/its/article/using-vpn-resources-at-um-flint-2

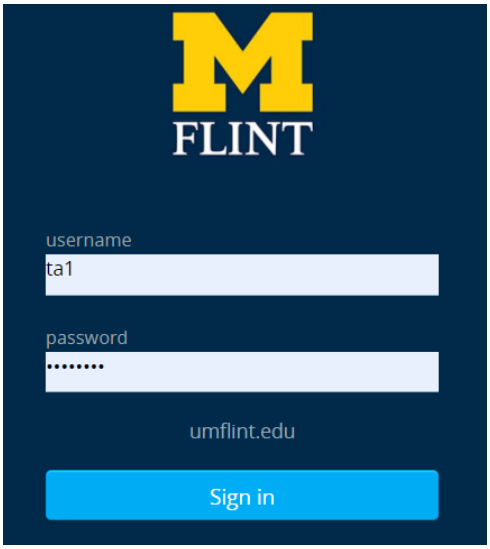
If you are working remotely, establish a VPN connection the university to continue with the rest of this guide.

Login

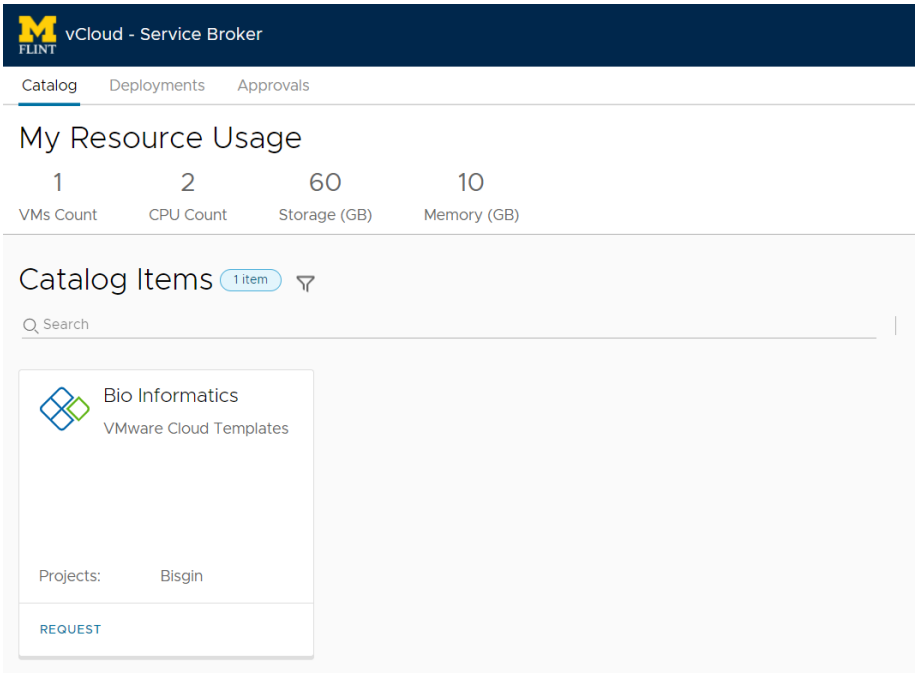
Go to: https://vcloud.umflint.edu/catalog/#/library

If this is the first time you are accessing the portal be sure the “umflint.edu” domain is selected and click the “Next” button.

Sign in to the portal using your unique name (the first part of your email address without the @umich.edu part) and your password. Be sure the domain is listed as “umflint.edu” under the password box.

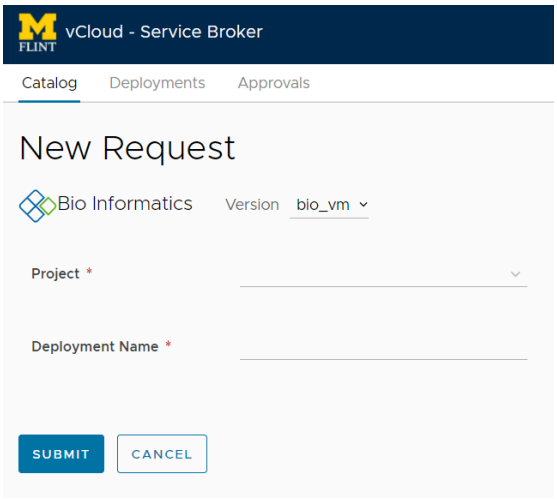


You will be taken to the “Catalog” section of the service broker. Here you will see a listing of the templates you have access to. The instructor that issued the template is listed next to the template “Projects” field.



Deploy

To deploy a desired template click on the “REQUEST” link on the bottom of the template.

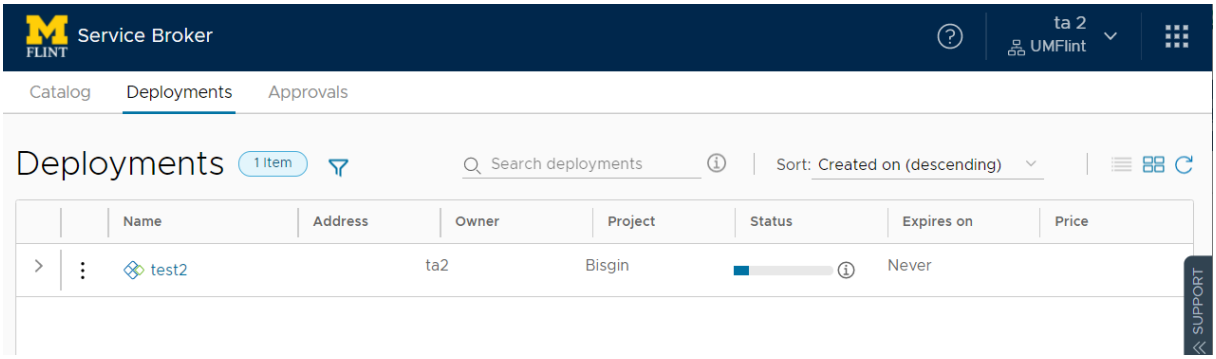


The latest template version will be selected for you from the drop down “Version” menu.

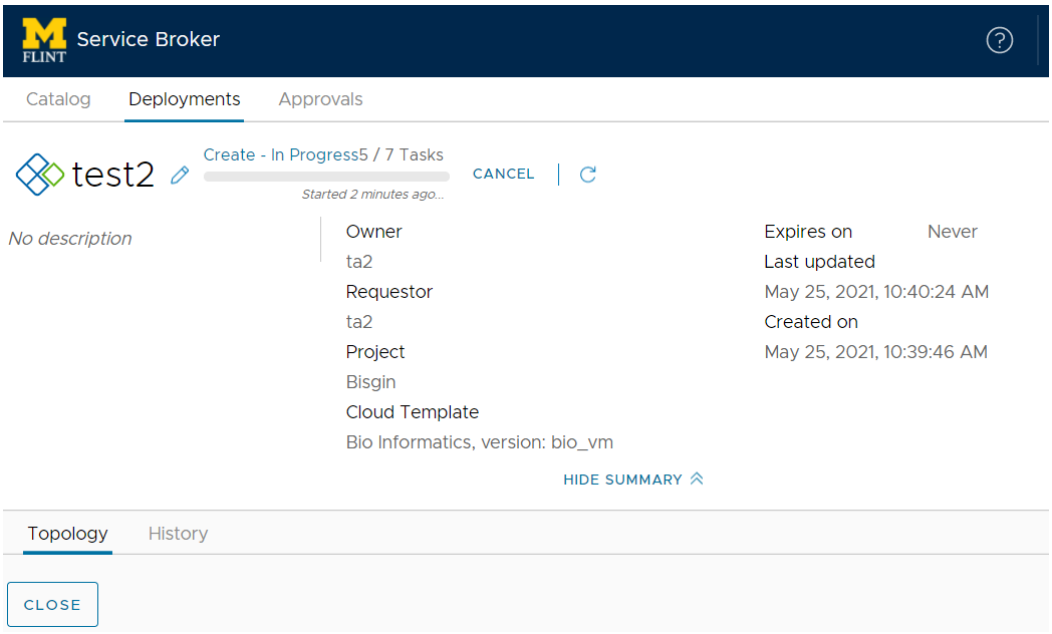
Select the instructor name under the “Project” drop down menu.

Give your deployment a descriptive name in the “Deployment Name” field. Try to include your course number, unique name (user ID), or other information to help your instructor identify the purpose of your deployment.

Click “Submit”.



In the deployment section you will see the new deployment listed. You can click on the deployment name to see the progress of the deployment.



The topology section on the bottom shows a map of what is being created and deployed as part of your request.

Be patient. It may take 15 minutes or more for the system to clone the VM and deploy it to your workspace. This time may differ based on the template size, complexity, and other deployment requests in the queue.

Accessing the VM

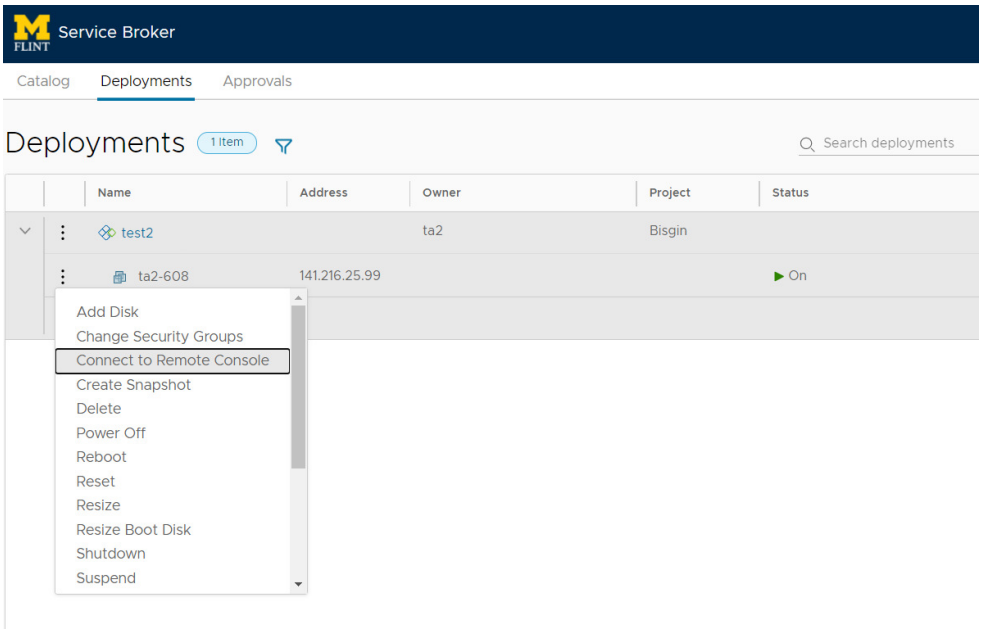
Once a deployment has finished, there are two primary ways you can interact with your VM. Once access is configured, you will probably interact with your VM most often through the use of an integrated or third party connection tool. Depending on the operating system and how your instructor has configured your template, the tool may be SSH for a Linux system, RDP for a windows system, or a TeamViewer session. Your instructor may have more information on the best method for connecting to your particular VM.

The second way you can interact with your VM is using the console from the deployment menu. The console is a representation of what you would see sitting in front of the screen if the VM was a physical machine.

Most VMs will power on once they have deployed. In the “Deployments” section expand a deployment using the arrow to the left of the deployment you want to interact with, you will see the network connection and the VM listed in the deployment.

Notice that the VM will have the IP address listed next to the VM name as 141.216.XX.XXX. This is the IP address you can use for RDP or SSH access to the VM (if enabled). If you are using a third party tool you may need to access the VM console first to get the session's configuration information.

Left click the three dots to the left of the VM to pop up a menu, from here you can choose "Connect to Remote Console". This will open a new browser tab where you can interact with the VM console.



On the console's browser page, buttons should be available in the upper right corner to pass special key combinations to the VM.

Your instructor will provided you with credentials to log onto your VM. Now would be a good time to log on to the VM and change the default credentials to something more secure.

Tips

If you forget the direct URLs for accessing the catalog or your deployment you can simply use https://vcloud.umflint.edu/ Although, you will need to click through a couple of extra prompts to log into the system and choose the service broker.

A VPN connection is required if you are using SSH or RDP to connect to your VM.