**Project Name:** Auto Scaling

**Create Auto scaling virtual Machine & test the load, configure monitor, alerts, logs.**

Creating an auto-scaling VM (Virtual Machine), load testing, configuring monitoring, setting up alerts, and managing logs involves several steps and tools.

What is Auto Scaling in Azure?

* Autoscale allows you to automatically scale your applications or resources based on demand.
* Use Autoscale to provision enough resources to support the demand on your application without over provisioning and incurring unnecessary costs.

What is Virtual Machine?

* A virtual machine (VM) is a digital version of a physical computer.
* Virtual machine software can run programs and operating systems, store data, connect to networks, and do other computing functions, and requires maintenance such as updates and system monitoring.

What is Test the Load?

* With Azure Load Testing, you can run performance tests against web applications, mobile applications, web APIs, microservices, database connections .

What is Monitor?

* Azure Monitor autoscaling allows you to scale the number of running instances in or out, based on telemetry data or metrics.
* Scaling can be based on any metric, even metrics from a different resource.
* For example, scale a Virtual Machine Scale Set based on the amount of traffic on a firewall.

What is Alerts?

* Alerts help you detect and address issues before users notice them by proactively notifying you when Azure Monitor data indicates there might be a problem with your infrastructure or application.

What is Logs?

* Autoscale has two log categories and a set of metrics that can be enabled via the **Diagnostics settings**tab on the **Autoscale setting** page.

The two categories are:

* [Autoscale Evaluations](https://learn.microsoft.com/en-us/azure/azure-monitor/reference/tables/autoscaleevaluationslog) contain log data relating to rule evaluation.
* [Autoscale Scale Actions](https://learn.microsoft.com/en-us/azure/azure-monitor/reference/tables/autoscalescaleactionslog) log data relating to each scale event.

Why We using Auto Scaling in Azure?

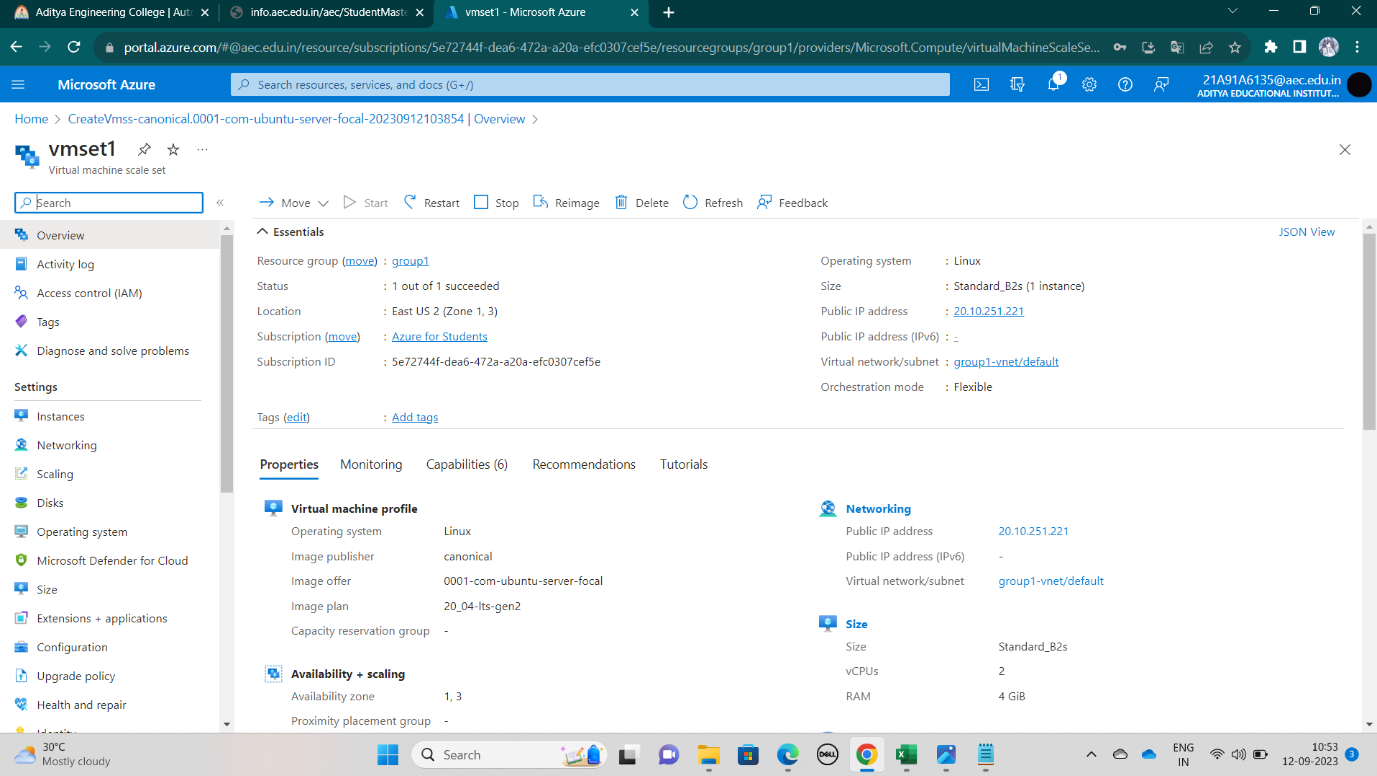
* Autoscaling is a cloud computing feature that enables organizations to scale cloud services such as server capacities or virtual machines up or down automatically, based on defined situations such as traffic or utilization levels.

Step by Step Processes:

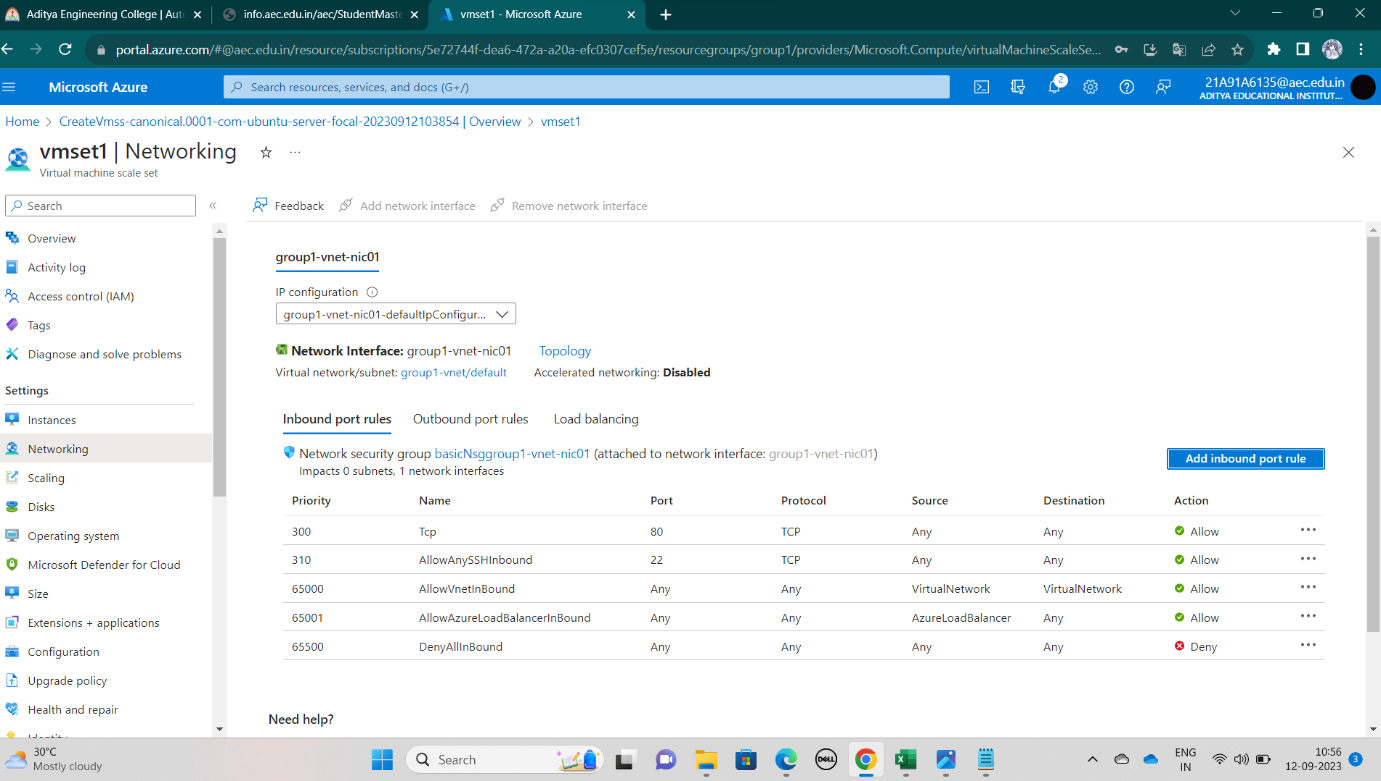
Step1:

* Open and Sign in to Azure Portal
* Then go to the Search bar and then search virtual machine scale sets
* After that create on scale set

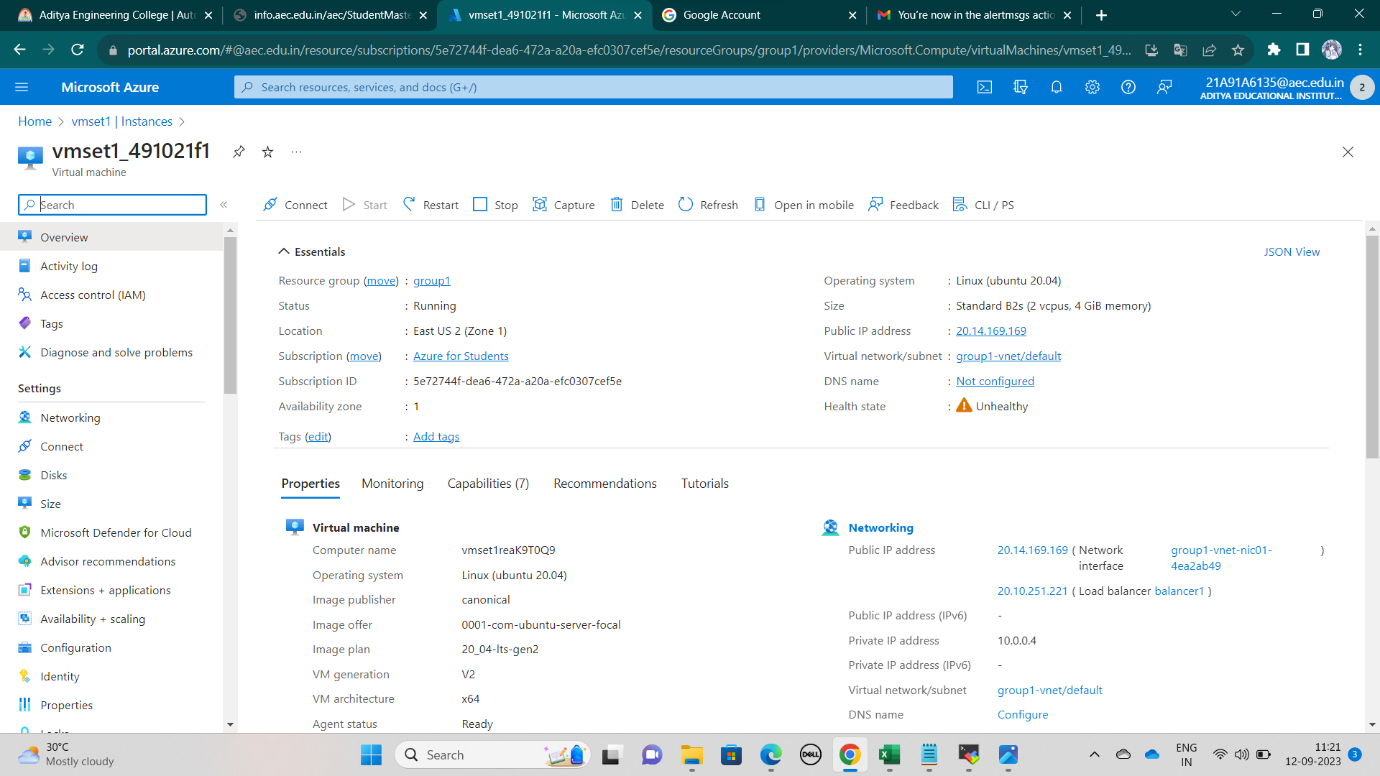
Create a Virtual Machine Scale Set as Shown Below.



Step2: open the virtual machine scale set we created, then go to networking click on the networking and add the inbound port rule then check it once it is added or not.



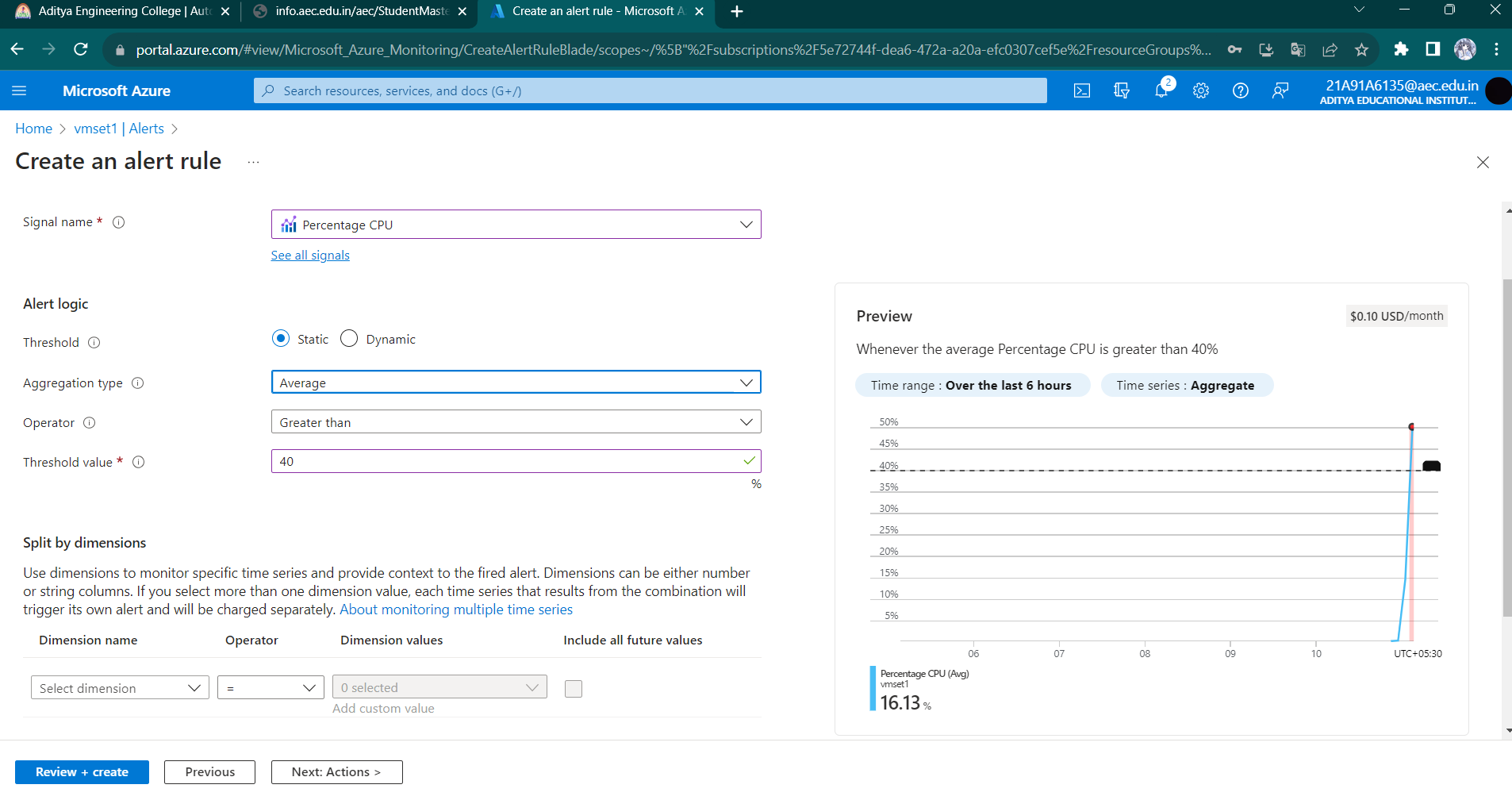
Step3: In the step 3 we can see the virtual machine also known as Instances is created successfully.



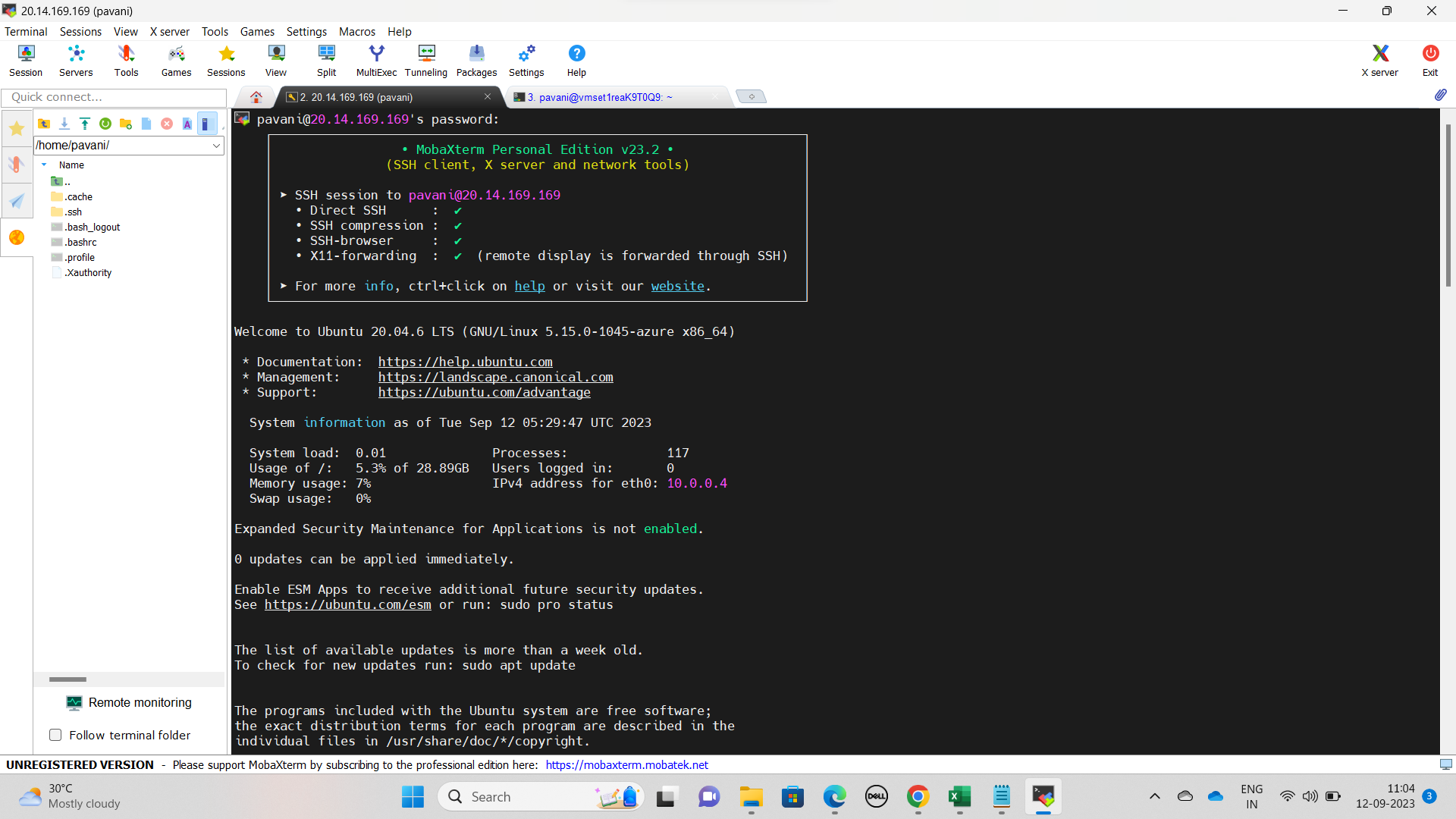
Step 4: Now create an action group by clicking alerts in the left navigation.

Then open alerts fill details like your email, mobile number must be given then only you’ll receive a notification as an alert whenever it is important.

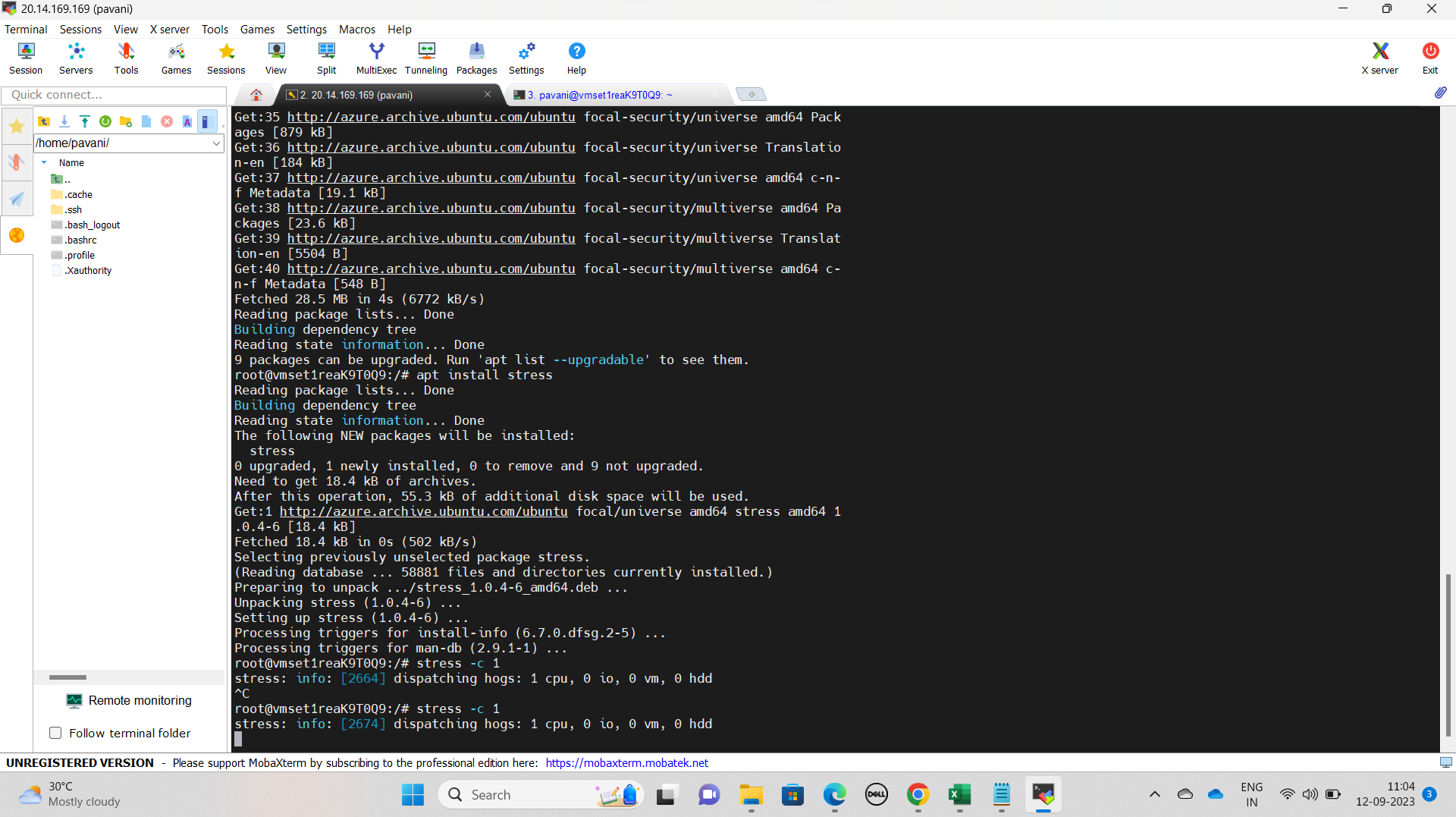
As shown below create an alert rule.



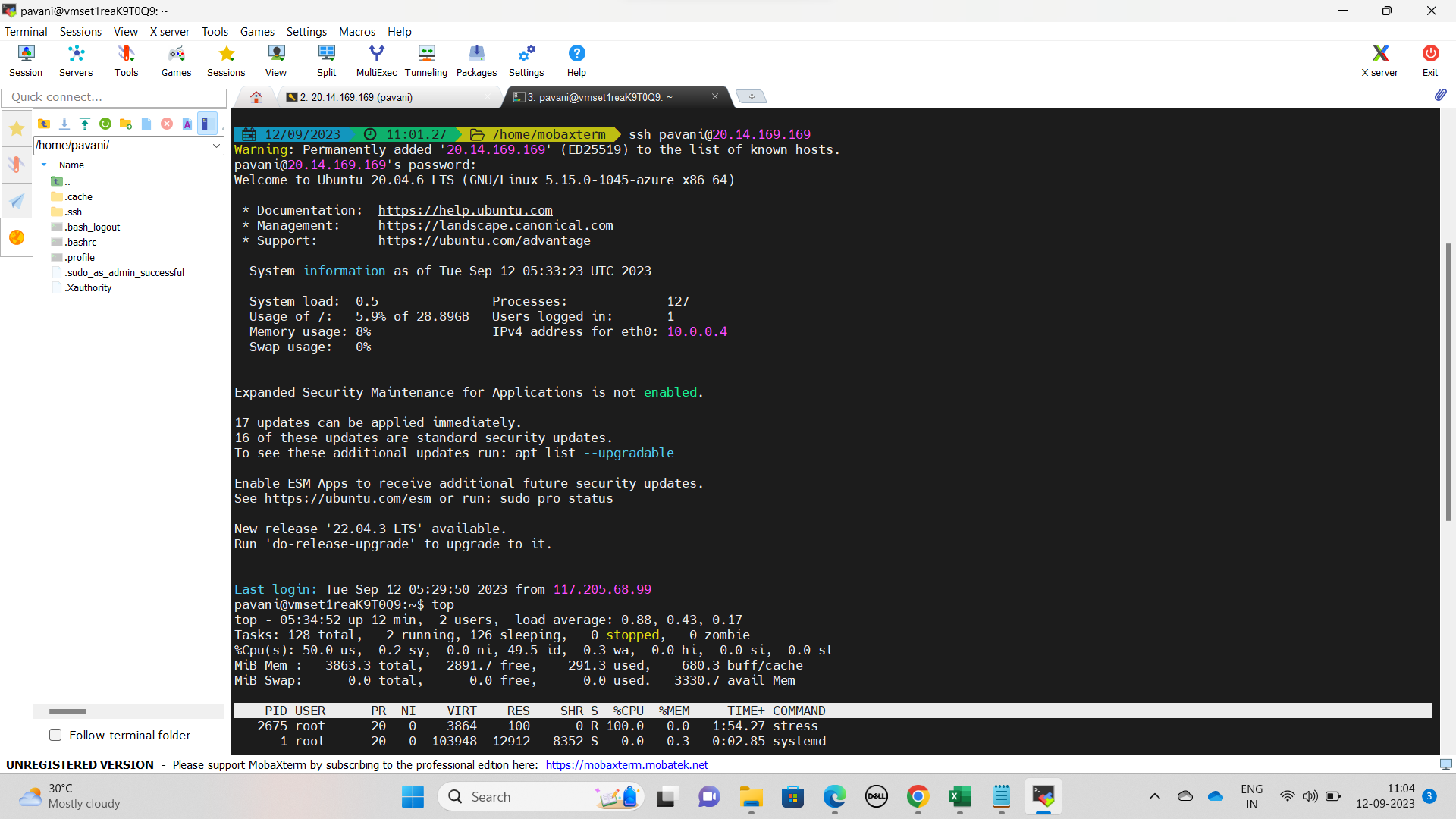
Step 5: Now using public IP we can open and start the server.

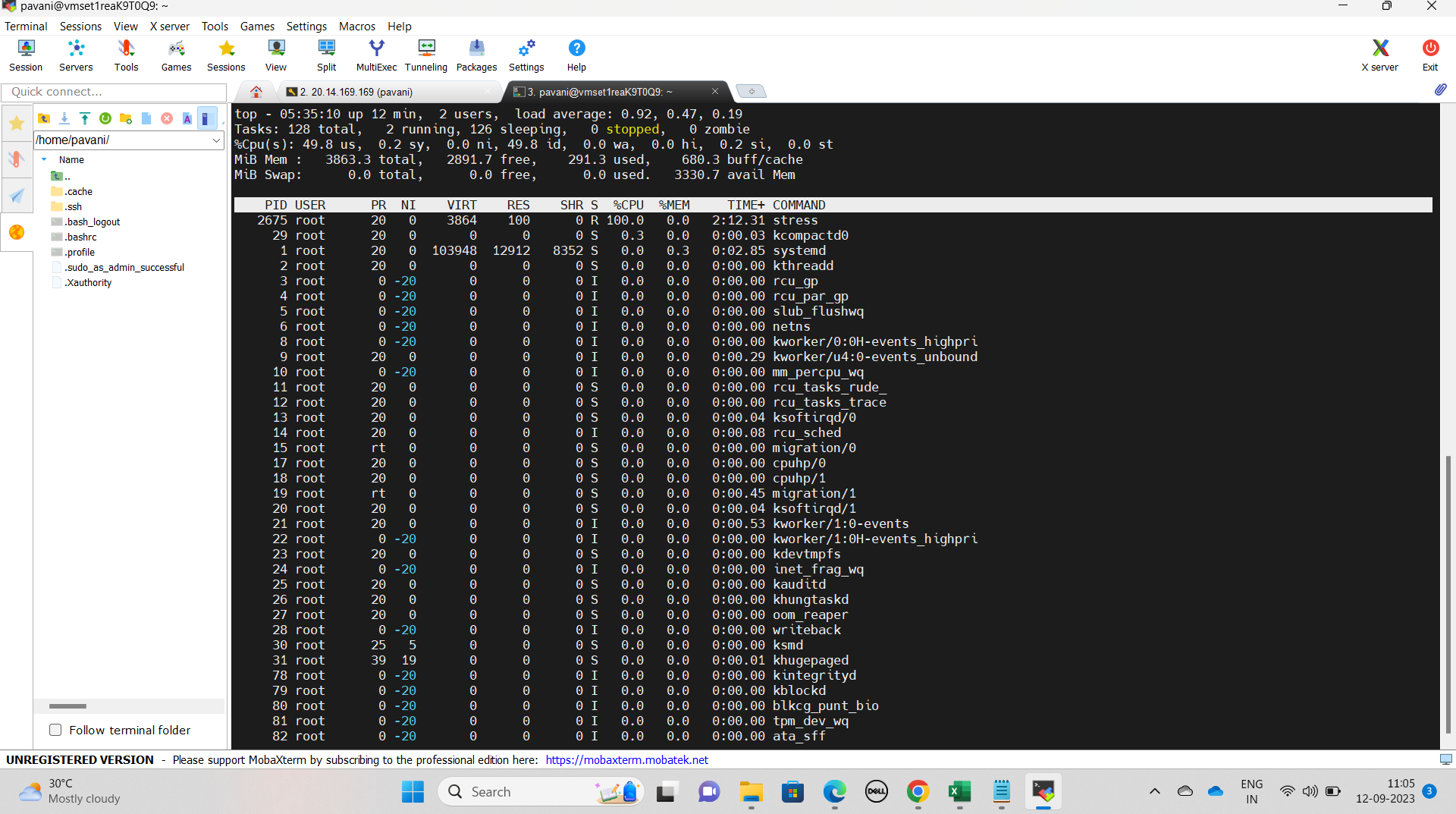


Here we can see that Stress is installed in order to increase the load.

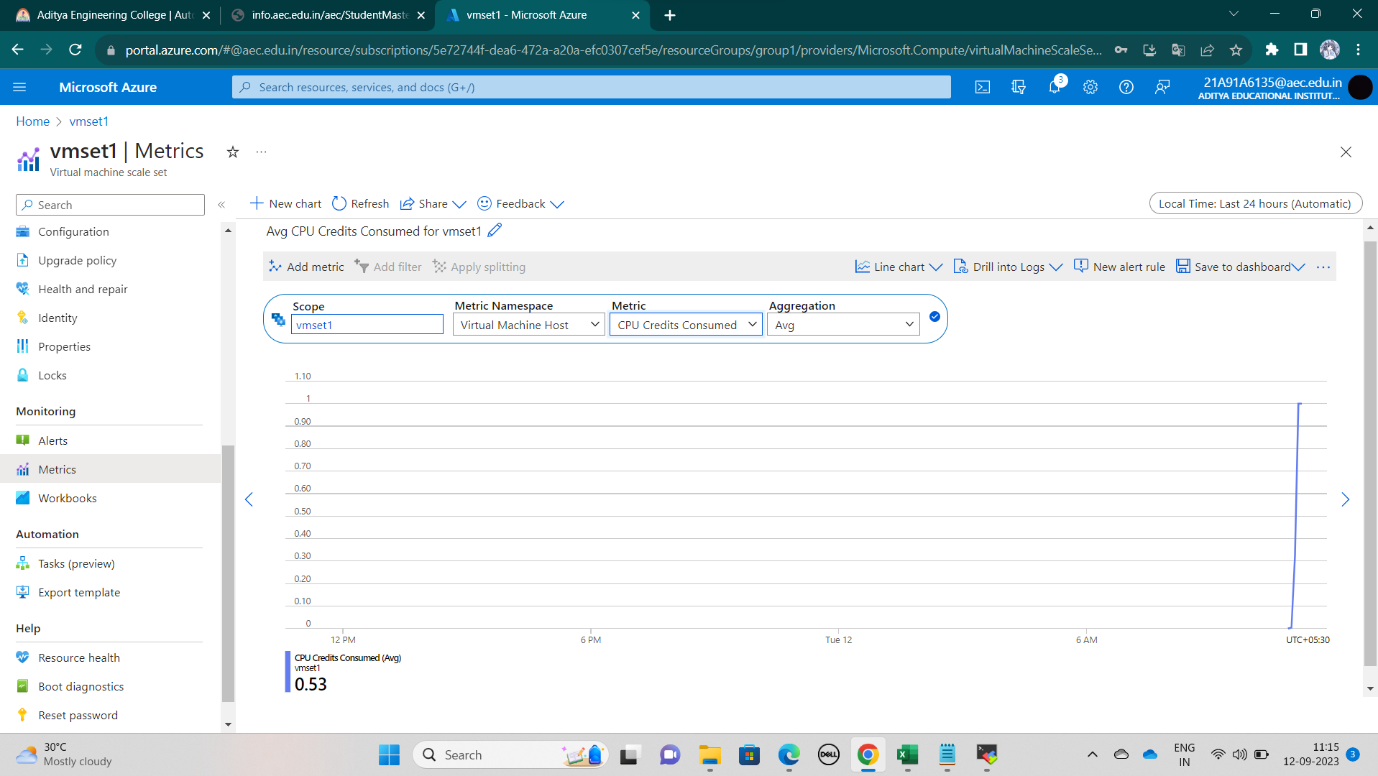


In another window observer the CPU load using top command

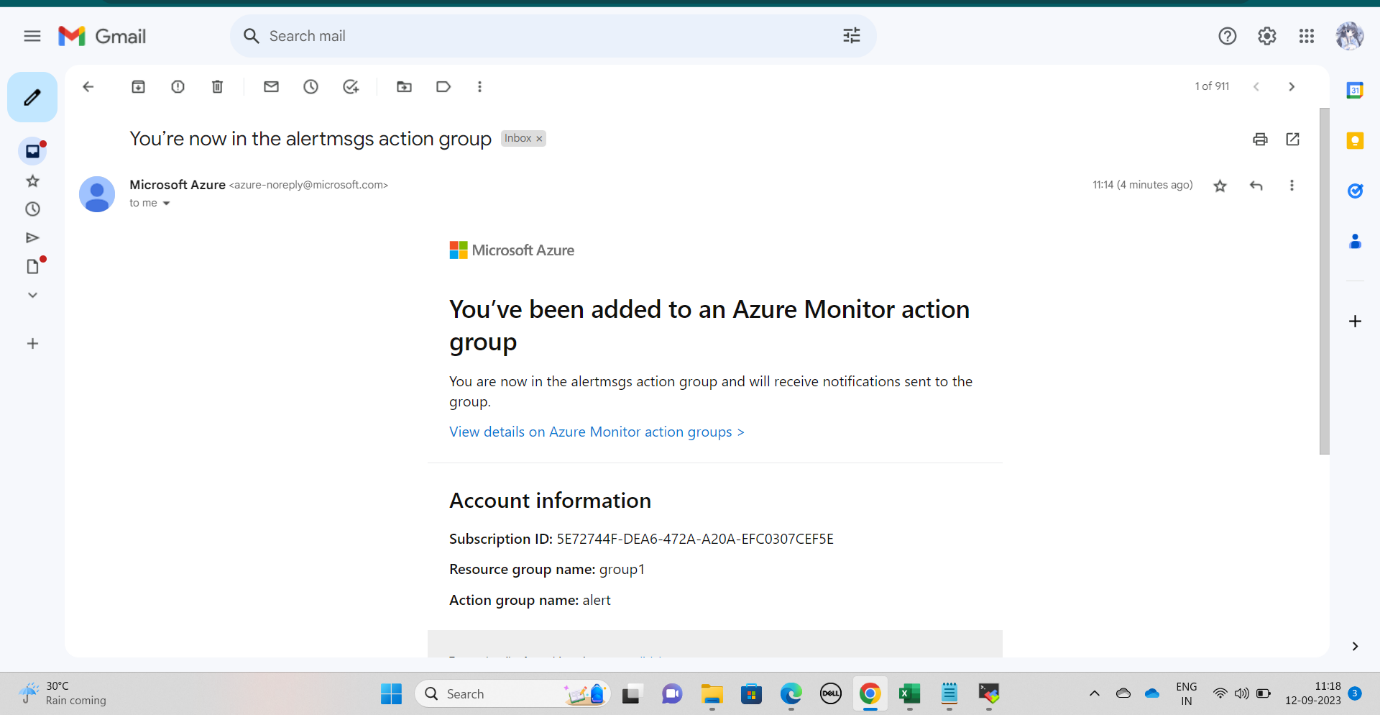


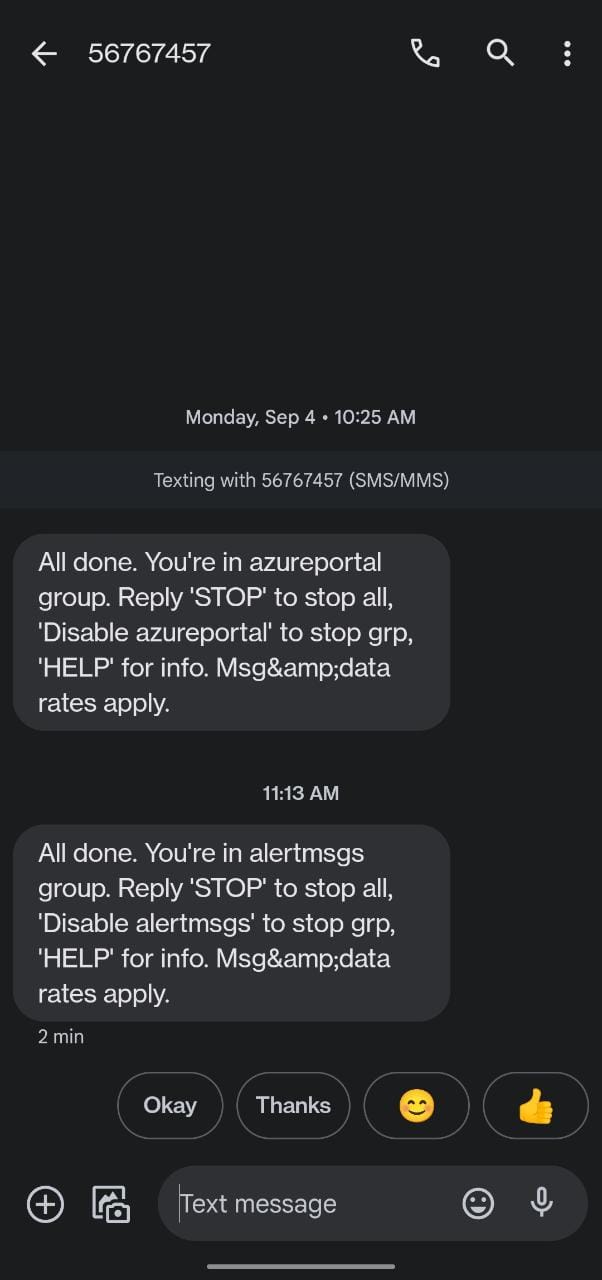


Step 6: We can see the usage of the CPU by clicking on metrics

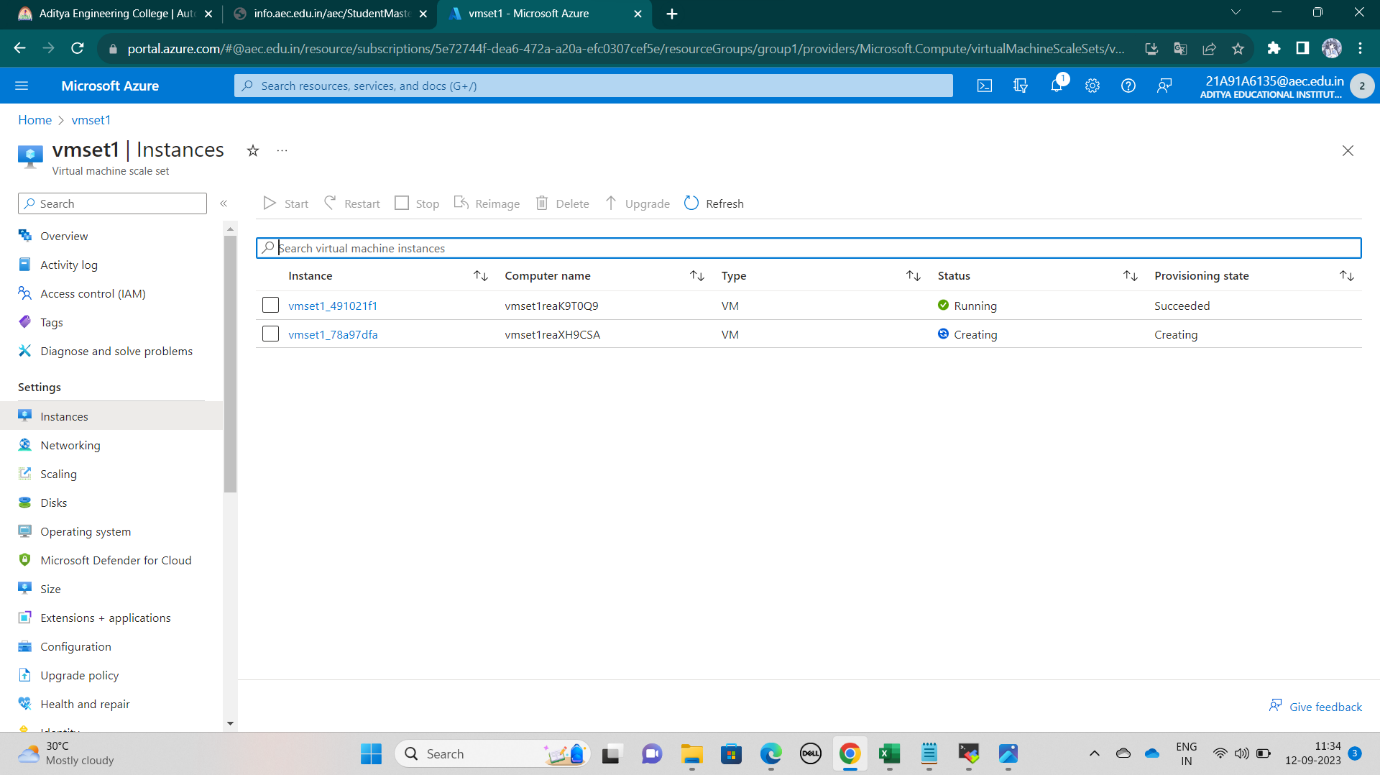


After the load got increased we’ll get an alert message to our email and mobile number.





When load is increased automatically it will create another instance we can see how another instance is created in given below.



Step 7: Now go to Activity log and check the logs.

