Mansoura University
Faculty of Computers and Information Sciences
Department of Computer Science
First Semester- 2020-2021

[CS412P] Distributed Systems

Grade : Fourth grade

By : Zeinab Awad

WHAT IS A CLUSTER?

- A cluster is a type of parallel or distributed processing system that consists of a collection of interconnected stand-alone computers working together as a single, integrated computing resource.
- The main benefits of clusters are:
- Availability
- Performance
- Scalability

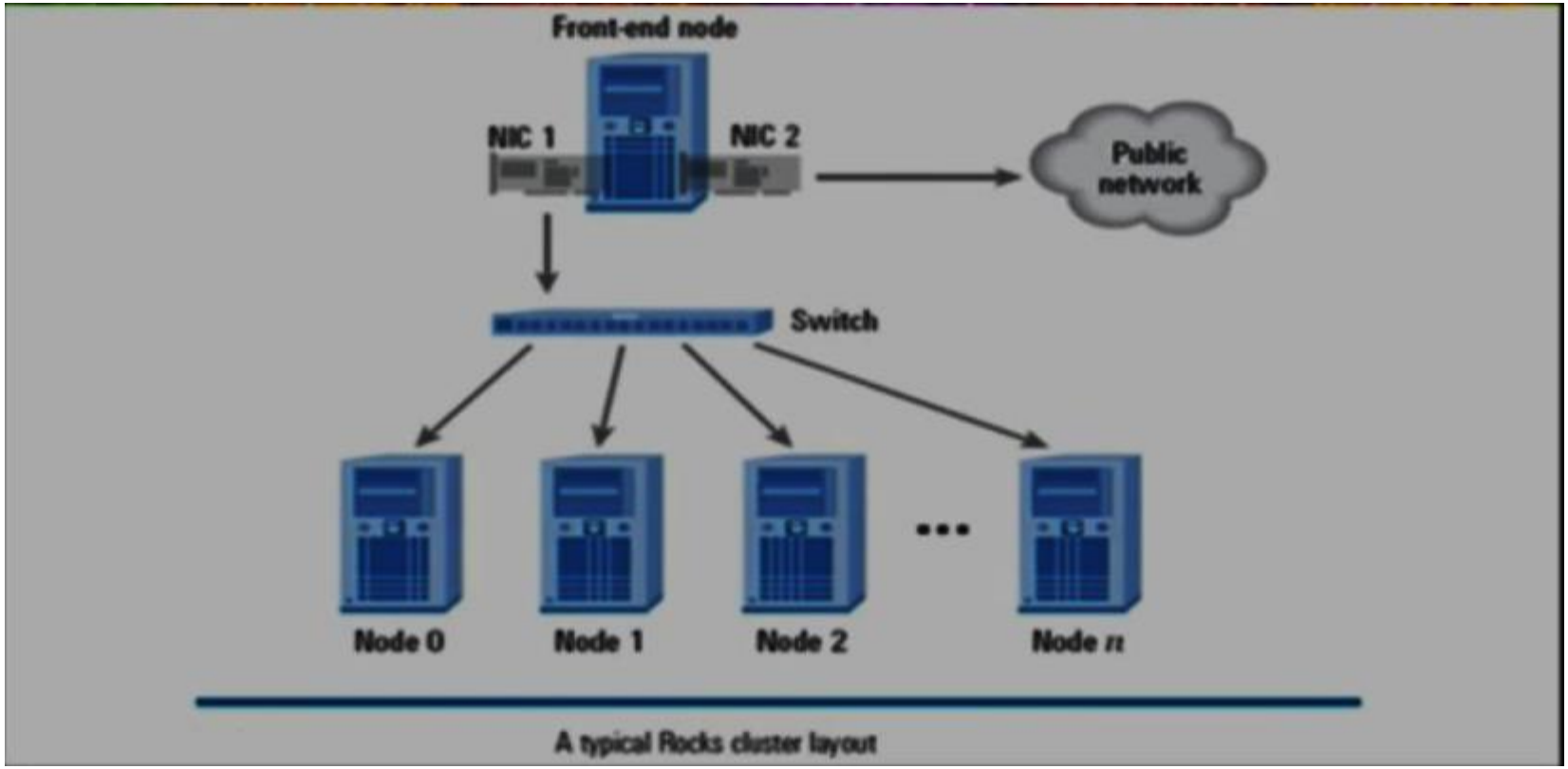
WHY USE A CLUSTER ?

- A cluster is a way of providing highly-available applications. It contains two or more machines, each hosts a replica of an application or service. For example, a web application.
- All machines share a common IP address called service IP or virtual IP (VIP). Any traffic received by this IP is routed to one or more of the backend machines depending on the architecture.
- Machines in a cluster may also share storage; so that data is kept in a consistent state.
- By using a cluster, you ensure that the service is always up and running; because if one machine is down, the other machine(s) will continue to provide the service.

HOW DOES IT WORK?

- A user submits a job to the head node.
- The job identifies the application to run on the cluster.
- The job scheduler on the head node assigns each task defined by the job to a node and then starts each application instance on the assigned node.
- Results from each of the application instances are returned to the client via files or databases.

A TYPICAL CLUSTER LAYOUT



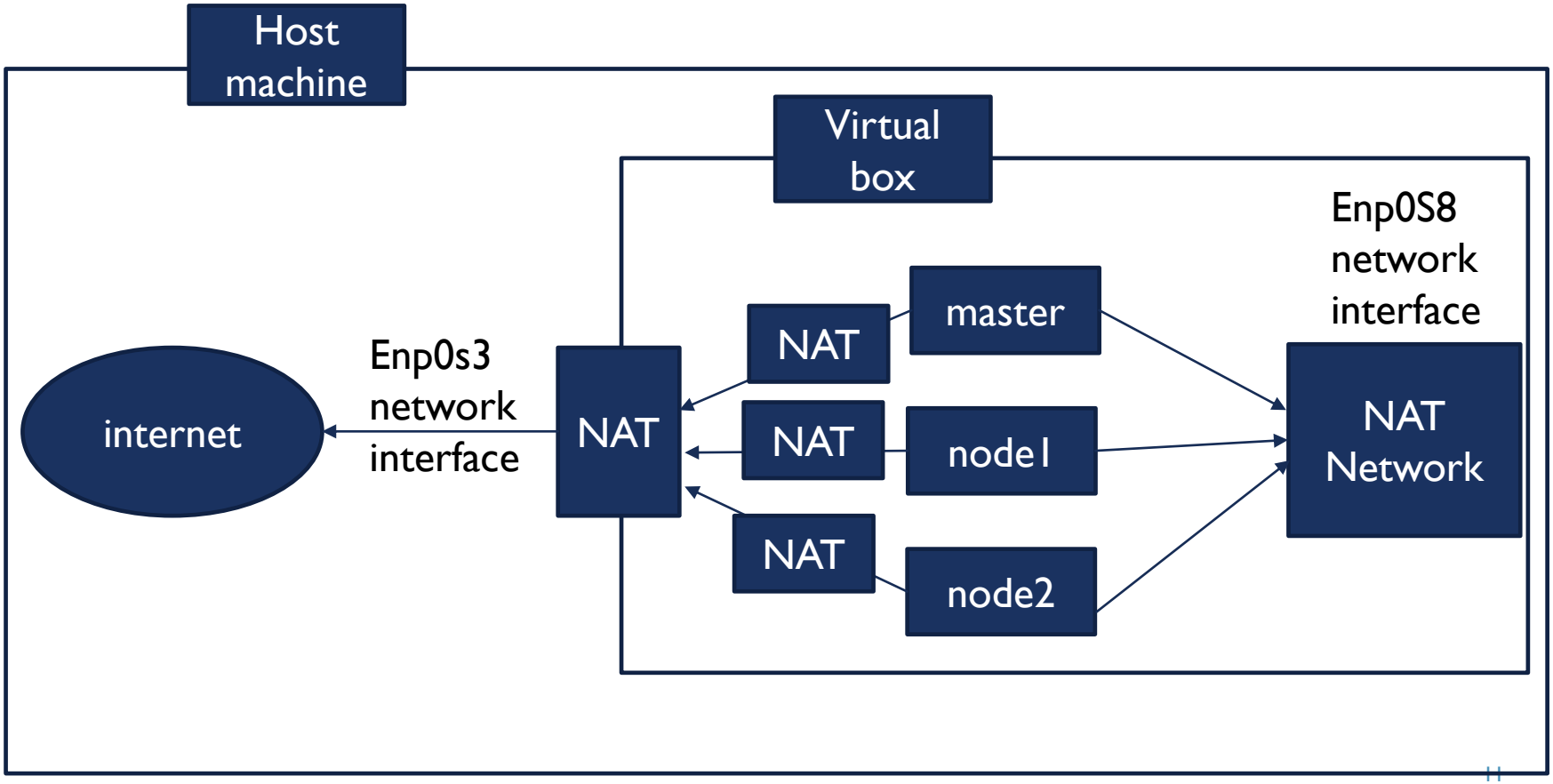
CREATING A THREE NODES CLUSTER(SETUP)

- In that cluster all nodes can access each others and also have access to the internet.
- All interaction will be made through the master machine
- Which serve as an entry point. The nodes should ssh to master machine in order to interact with other nodes.
- The master is the main work machine, and it has the desktop GUI and all other application like editors, compilers, word processors (desktop environment).
- The other 2 nodes are the worker machines and has a terminal-based OS.
- Create three virtual machines using Oracle VM VirtualBox Manager ,one for the master node named master and the other two for node1 and node2 respectively.

NETWORK SETUP FOR VIRTUAL MACHINES CLUSTER

- Install ubuntu desktop 18.04 virtual disk image on the master virtual machine.
- install ubuntu server 18.04 virtual disk image on the other two nodes (light weight and does not have GUI)
- Every machine has an ip address that's uniquely identify that machines .
- The ips assigned dynamically through a virtual machines by the dhcp server
- If the virtual machines restarted, then the connection with other machines would be losted .
- To avoid such condition a static ip address should be assigned to all the virtual machines .

NETWORK SETUP FOR VIRTUAL MACHINES CLUSTER



CLUSTER BUILDING

Thanks