Selenium Grid with Docker

Cybertek

What is AWS?

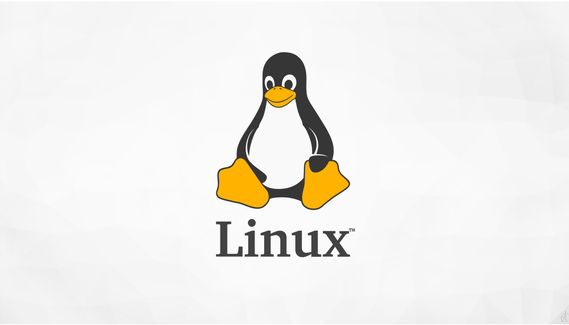
Amazon Web Services (AWS) is a secure cloud services platform, offering compute power, database storage, content delivery and other functionality to help businesses scale and grow.
In simple words AWS allows you to do the following things-

- Running web and application servers in the cloud to host dynamic websites.
- Securely store all your files on the cloud so you can access them from anywhere.
- Using managed databases like MySQL, PostgreSQL, Oracle or SQL Server to store information.
- Deliver static and dynamic files quickly around the world using a Content Delivery Network (CDN).
- Send bulk email to your customers.

Cybertek

EC2

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides secure, resizable compute capacity in the cloud. It is designed to make web-scale cloud computing easier for developers. Amazon EC2's simple web service interface allows you to obtain and configure capacity with minimal friction. It provides you with complete control of your computing resources and lets you run on Amazon's proven computing environment.



Cybertek

What is Linux?

From smartphones to cars, supercomputers and home appliances, home desktops to enterprise servers, the Linux operating system is everywhere.

Linux has been around since the mid-1990s and has since reached a user-base that spans the globe. Linux is actually everywhere: It's in your phones, your thermostats, in your cars, refrigerators, Roku devices, and televisions. It also runs most of the Internet, all of the world's top 500 supercomputers, and the world's stock exchanges.

But besides being the platform of choice to run desktops, servers, and embedded systems across the globe, Linux is one of the most reliable, secure and worry-free operating systems available.

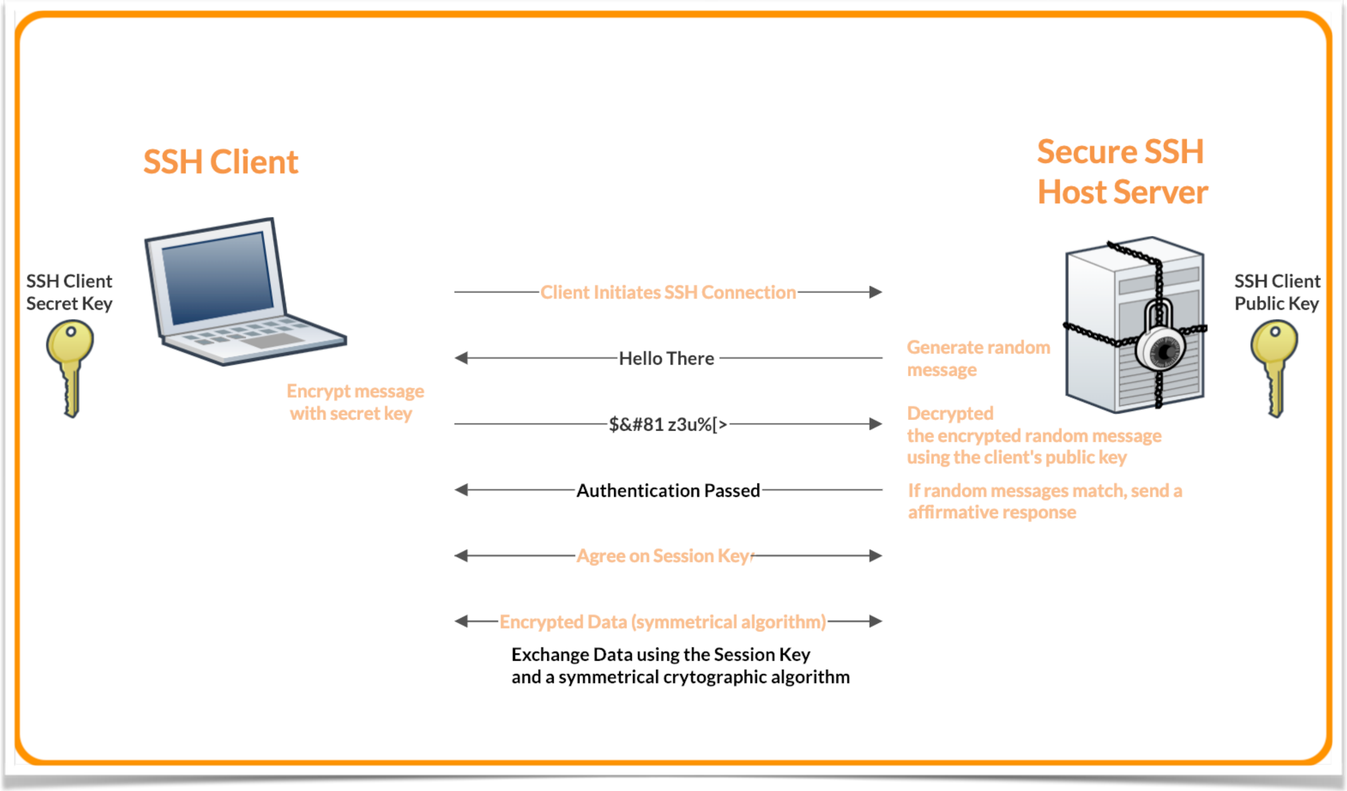
Just like Windows, iOS, and Mac OS, Linux is an operating system. In fact, one of the most popular platforms on the planet, Android, is powered by the Linux operating system. An operating system is software that manages all of the hardware resources associated with your desktop or laptop. To put it simply, the operating system manages the communication between your software and your hardware. Without the operating system (OS), the software wouldn't function.

Cybertek

SSH (Secure Shell)

An SSH client allows you to connect to a remote computer running an SSH server.

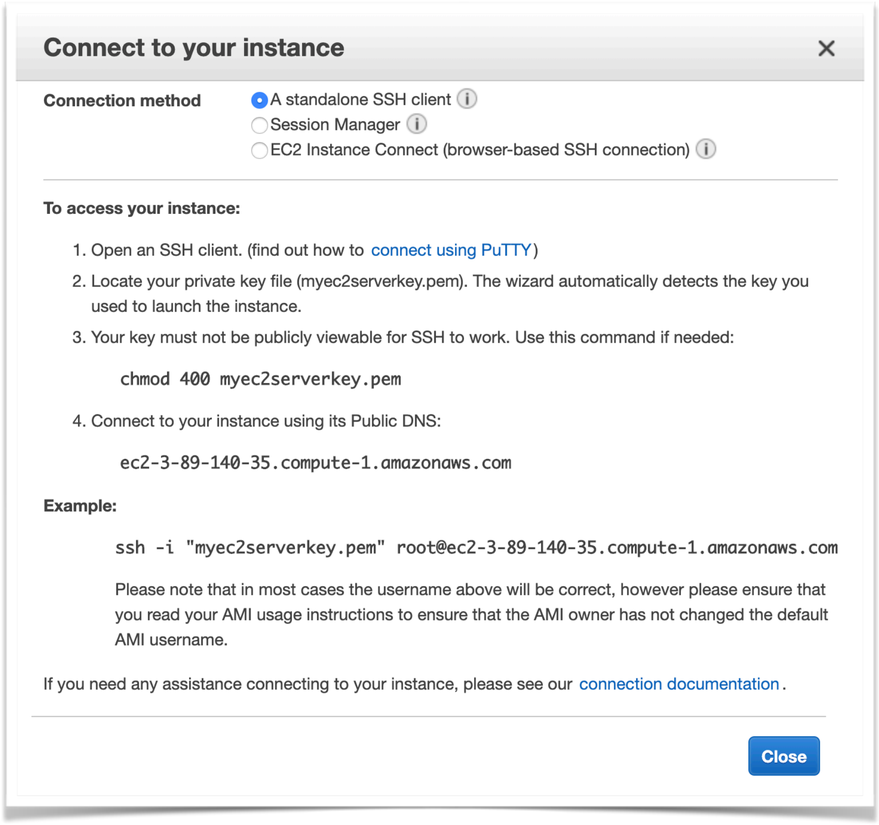
The Secure Shell (SSH) protocol is often used for remote terminal connections, allowing you to access a text-mode terminal on a remote computer as if you were sitting of it.



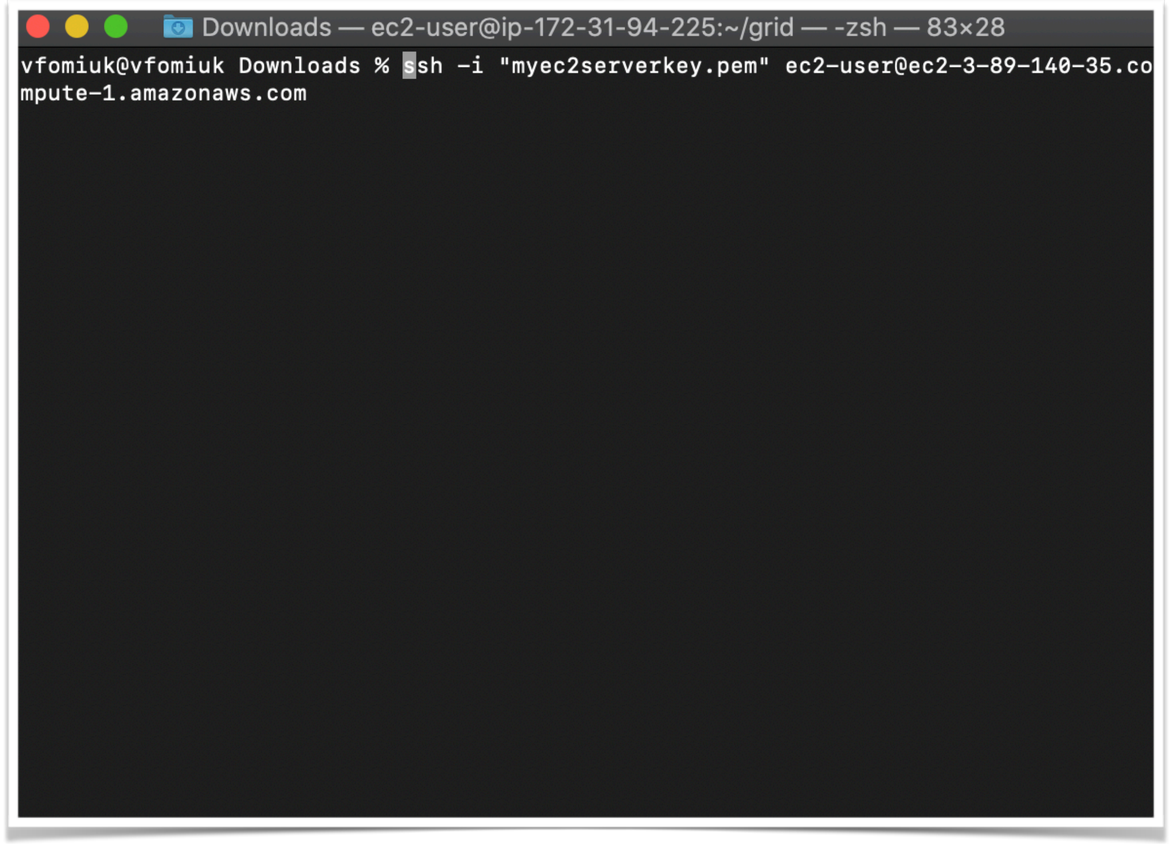
Cybertek

Connection

Execute command from step 3



Change root to ec2-user



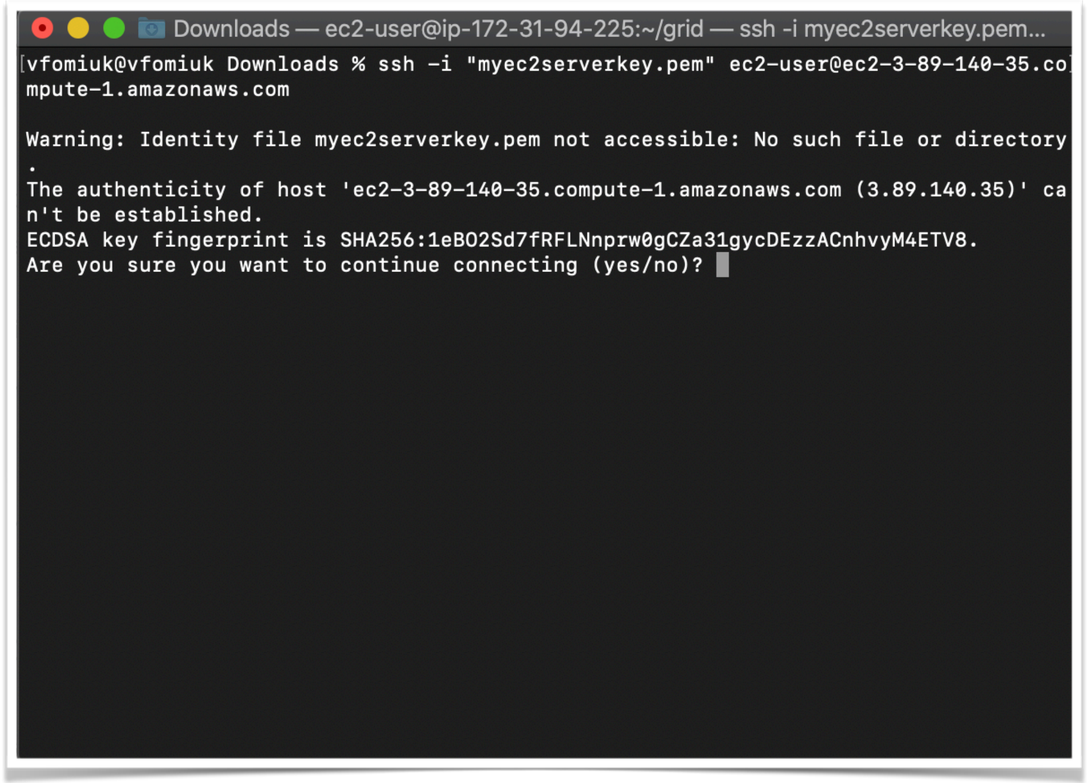
Copy command from example

Hit enter/return

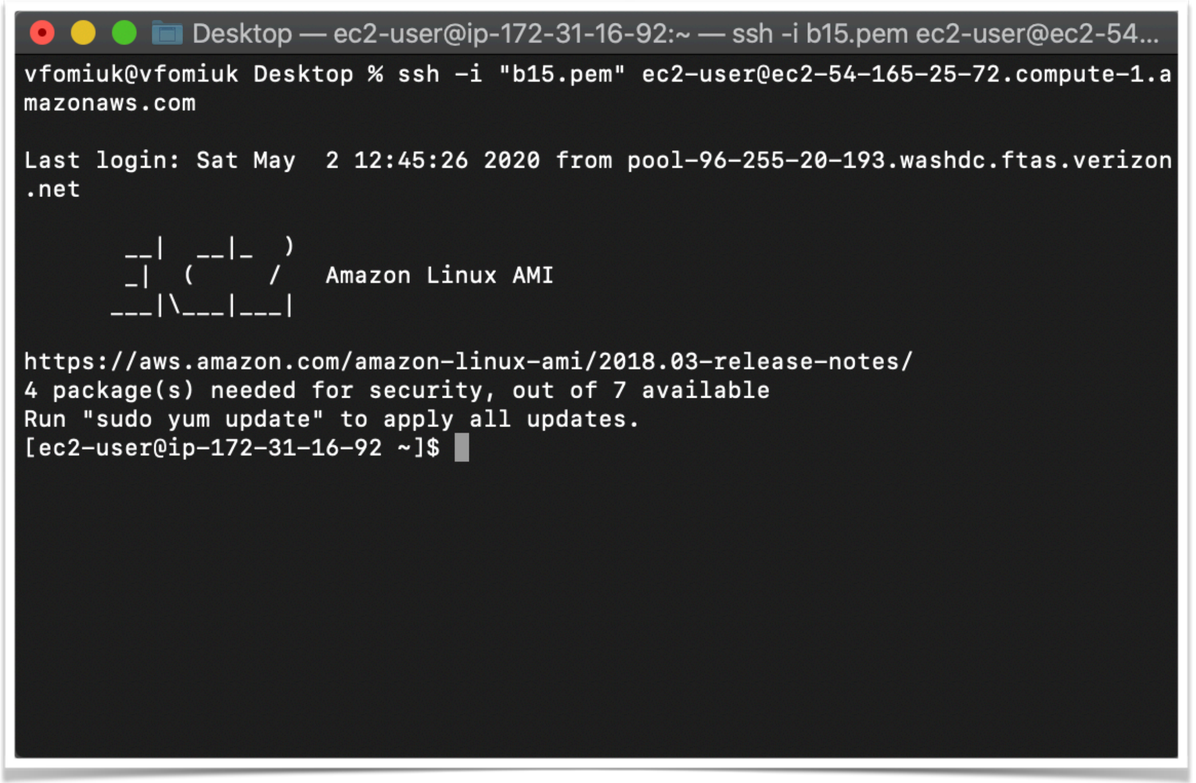
Cybertek

Connection

Enter yes and hit enter/return



Connected!



Make sure that .pem file is in the same directory

Cybertek

Docker

Docker is a tool designed to make it easier to create, deploy, and run applications by using containers. Containers allow a developer to package up an application with all of the parts it needs, such as libraries and other dependencies, and deploy it as one package. By doing so, thanks to the container, the developer can rest assured that the application will run on any other Linux machine regardless of any customized settings that machine might have that could differ from the machine used for writing and testing the code.

Cybertek

Docker

Cybertek

Docker

This spawned a Shipping Container Ecosystem!

- 90% of all cargo now shipped in a standard container
- Order of magnitude reduction in cost and time to load and unload ships
- Massive reduction in losses due to theft or damage
- Huge reduction in freight cost as percent of final goods (from >25% to <3%)

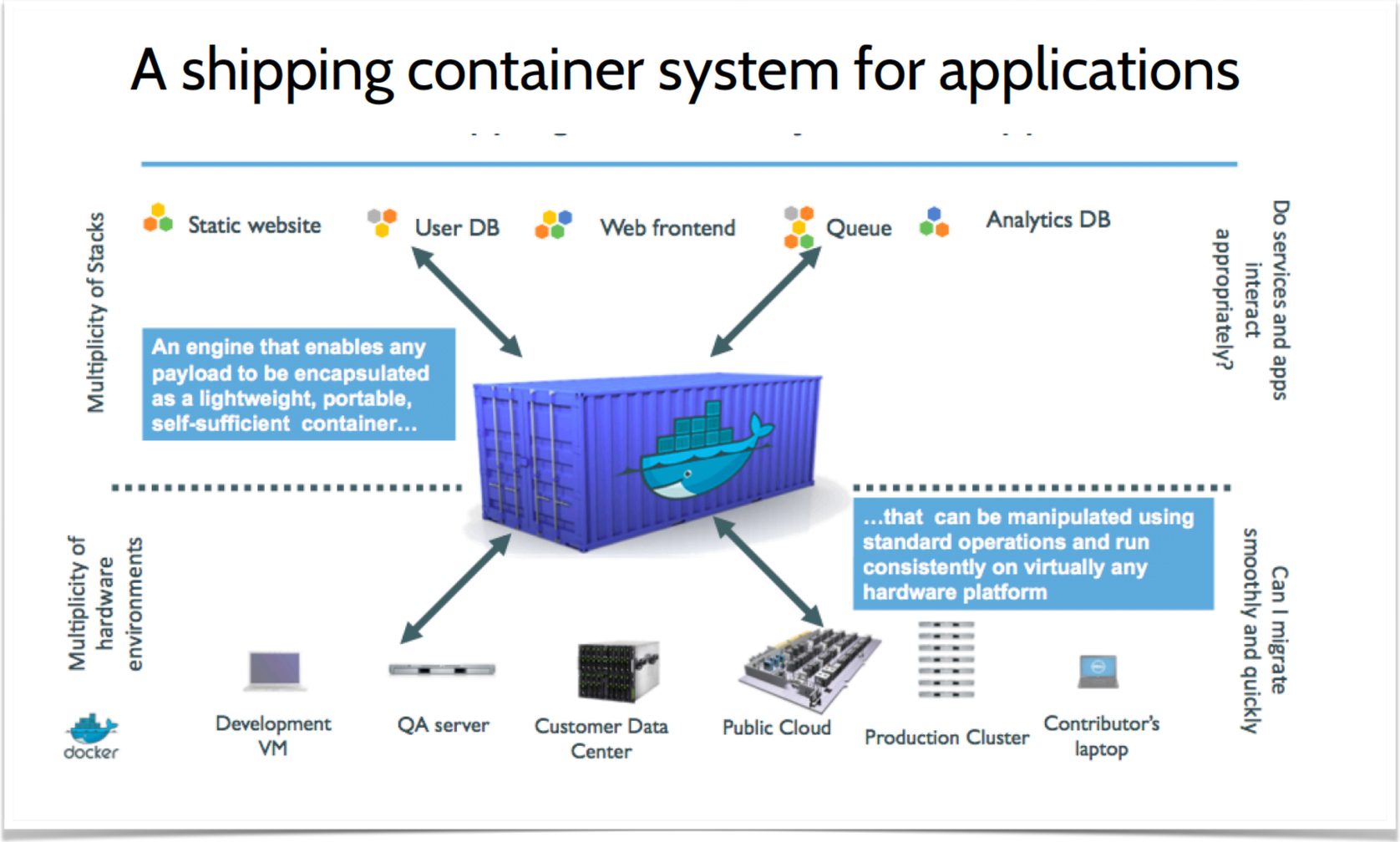
→ massive globalization

- 5000 ships deliver 200M containers per year

docker

Cybertek

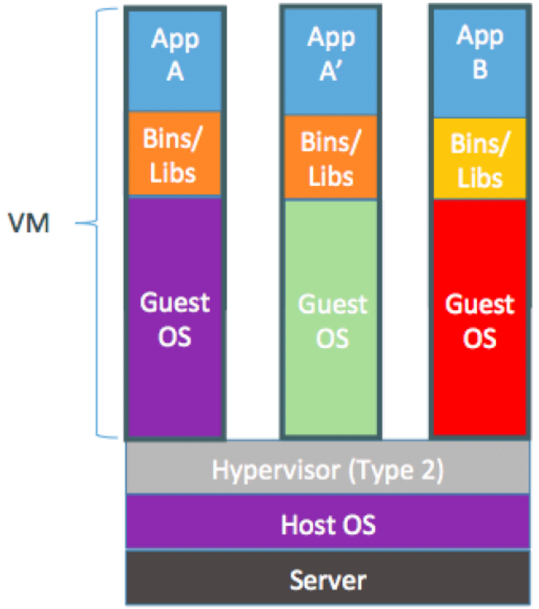
Docker



Cybertek

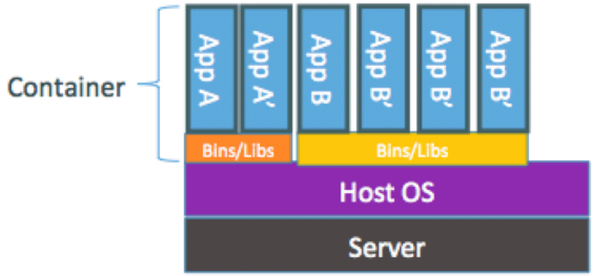
Docker

Step 1: containers as lightweight VMs



Containers are isolated,
but share OS kernel and, where appropriate, bins/libraries

...result is significantly faster deployment, much less overhead, easier migration, faster restart



docker

Cybertek

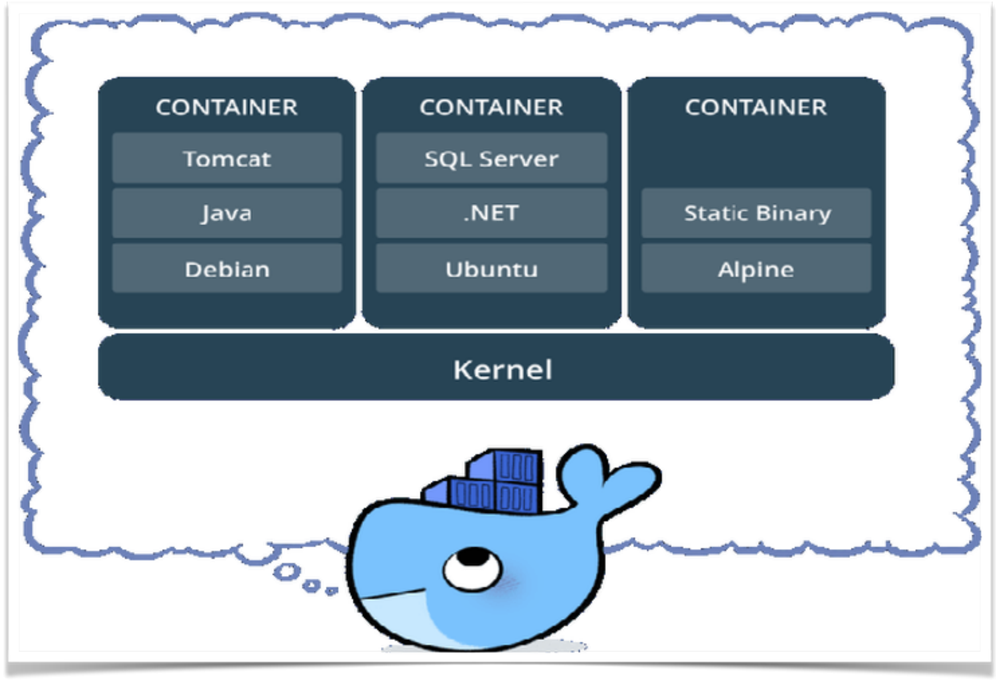
What’s the Diff: VMs vs Containers

VMs	Containers
Heavyweight	Lightweight
Limited performance	Native performance
Each VM runs in its own OS	All containers share the host OS
Hardware-level virtualization	OS virtualization
Startup time in minutes	Startup time in milliseconds
Allocates required memory	Requires less memory space
Fully isolated and hence more secure	Process-level isolation, possibly less secure

Cybertek

Docker container

Container is built from an **image**.
A container consists of an operating system, user-added **files**, and **meta-data**.
That image tells Docker what the container **holds**, what process to **run** when the container is launched, and a variety of other **configuration** data.
The Docker image is **read-only**.
When Docker runs a container from an image, it adds a read-write layer on top of the image (using a union file system as we saw earlier) in which your application can then run.



Cybertek

Docker compose

Compose is a tool for defining and running multi-container Docker applications. With Compose, you use a YAML file to configure your application's services.

Then, with a single command, you create and start all the services from your configuration. To learn more about all the features of Compose, see the list of features.

Compose works in all environments: production, staging, development, testing, as well as CI workflows. You can learn more about each case in Common Use Cases.

```
version: '2.0'
services:
  web:
    build: .
    ports:
    - "5000:5000"
    volumes:
    - .:/code
    - logvolume01:/var/log
    links:
    - redis
  redis:
    image: redis
volumes:
  logvolume01: {}
```


Cybertek

Some docker commands

```
docker build -t friendlyname .  # Create image using this directory's Dockerfile
docker run -p 4000:80 friendlyname  # Run "friendlyname" mapping port 4000 to 80
docker run -d -p 4000:80 friendlyname         # Same thing, but in detached mode
docker ps                                 # See a list of all running containers
docker stop <hash>                      # Gracefully stop the specified container
docker ps -a           # See a list of all containers, even the ones not running
docker kill <hash>                    # Force shutdown of the specified container
docker rm <hash>                # Remove the specified container from this machine
docker rm $(docker ps -a -q)           # Remove all containers from this machine
docker images -a                               # Show all images on this machine
docker rmi <imagename>            # Remove the specified image from this machine
docker rmi $(docker images -q)             # Remove all images from this machine
docker login             # Log in this CLI session using your Docker credentials
docker tag <image> username/repository:tag  # Tag <image> for upload to registry
docker push username/repository:tag            # Upload tagged image to registry
docker run username/repository:tag                   # Run image from a registry
```

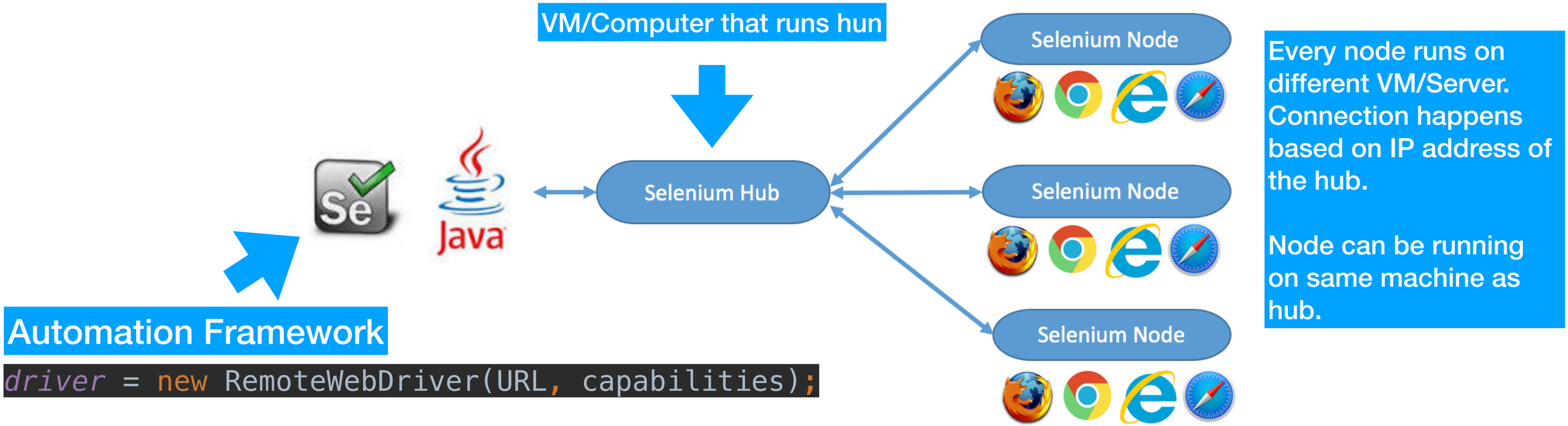

Cybertek

Selenium Grid

Cybertek

What is Selenium Grid?

The ***Selenium Grid*** is a testing tool which allows us to run our tests on different machines against different browsers. It is a part of the Selenium Suite which specialize in running multiple tests across different browsers, operating system and machines. You can connect to it with Selenium Remote by specifying the browser, browser version, and operating system you want. You specify these values throu Selenium Remote's Capabilities. There are two main elements to ***Selenium Grid*** — a ***hub*** and ***node***.



Every node runs on different VM/Server. Connection happens based on IP address of the hub.

Node can be running on same machine as hub.

Cybertek

The HUB

- The hub is the central point where you load your tests into.
- There should only be one hub in a grid.
- The hub is launched only on a single machine, say, a computer whose O.S is Windows 10 and whose browser is Chrome.
- The machine containing the hub is where the tests will be run, but you will see the browser being automated on the node.
- The hub can be parametrized with json file.

Cybertek

The NODE

- Nodes are the Selenium instances that will execute the tests that you loaded on the hub.
- There can be one or more nodes in a grid.
- Nodes can be launched on multiple machines with different platforms and browsers.
- Nodes can be located on different machines.
- Nodes can be parametrized with json file.

Cybertek

How Selenium-Grid Works–With a Hub and Nodes

A grid consists of a single hub, and one or more nodes. Both are started using the selenium-server.jar executable.

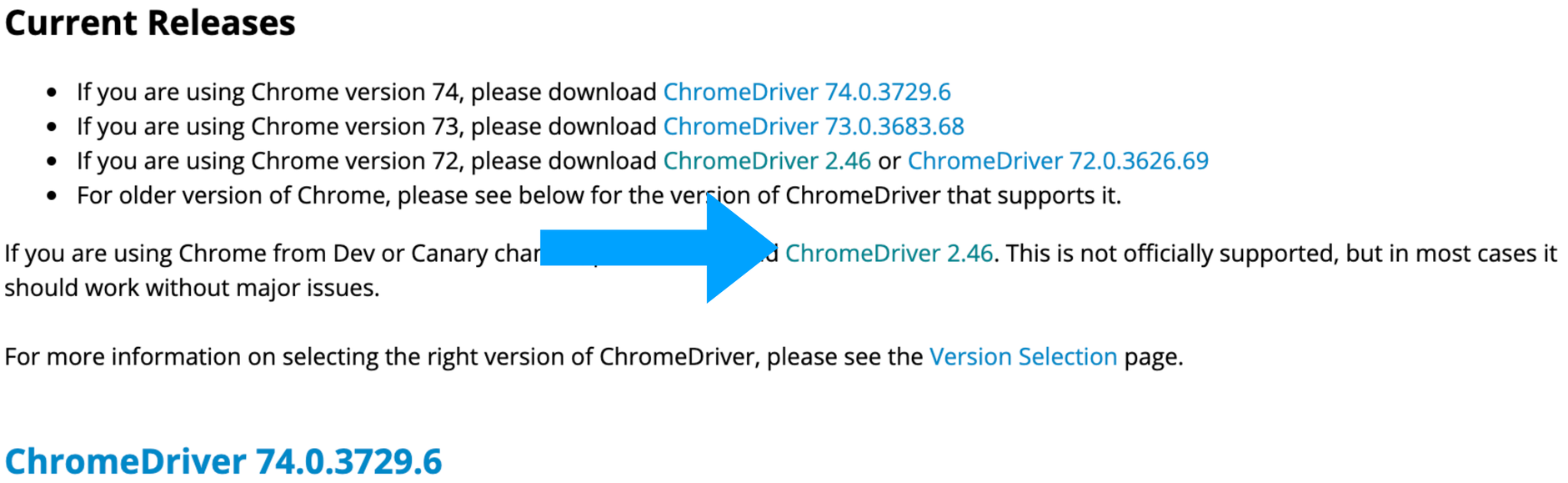
The hub receives a test to be executed along with information on which browse and ‘platform’ (i.e. WINDOWS, LINUX, etc) where the test should be run. It ‘knows’ the configuration of each node that has been ‘registered’ to the hub. Using this information it selects an available node that has the requested browser-platform combination.

Once a node has been selected, Selenium commands initiated by the test are sent to the hub, which passes them to the node assigned to that test. The node runs the browser, and executes the Selenium commands within that browser against the application under test.

Cybertek

How to Set Up Selenium Grid

- Step 1: Go to https://sites.google.com/a/chromium.org/chromedriver/downloads
- Step 2: download chrome driver



Current Releases

- If you are using Chrome version 74, please download ChromeDriver 74.0.3729.6
- If you are using Chrome version 73, please download ChromeDriver 73.0.3683.68
- If you are using Chrome version 72, please download ChromeDriver 2.46 or ChromeDriver 72.0.3626.69
- For older version of Chrome, please see below for the version of ChromeDriver that supports it.

If you are using Chrome from Dev or Canary char d ChromeDriver 2.46. This is not officially supported, but in most cases it should work without major issues.

For more information on selecting the right version of ChromeDriver, please see the Version Selection page.

ChromeDriver 74.0.3729.6

Cybertek

Step 3: Run the Hub

Open a command prompt and navigate to the directory where you copied the selenium-server-standalone file.

Execute the following command:

java -jar selenium-server-standalone-<version>.jar -role hub

The hub will automatically start-up using port 4444 by default. To change the default port, you can add the optional parameter -port when you run the command. You can view the status of the hub by opening a browser window and navigating to: http://localhost:4444/grid/console

Cybertek

Step 4: Run the node

Open a command prompt and navigate to the directory where you copied the selenium-server-standalone file. Type the following command (for chrome):

java -Dwebdriver.chrome.driver="chromedriver" jar selenium-server-standalone-<version>.jar -role node -hub http://localhost:4444/grid/register

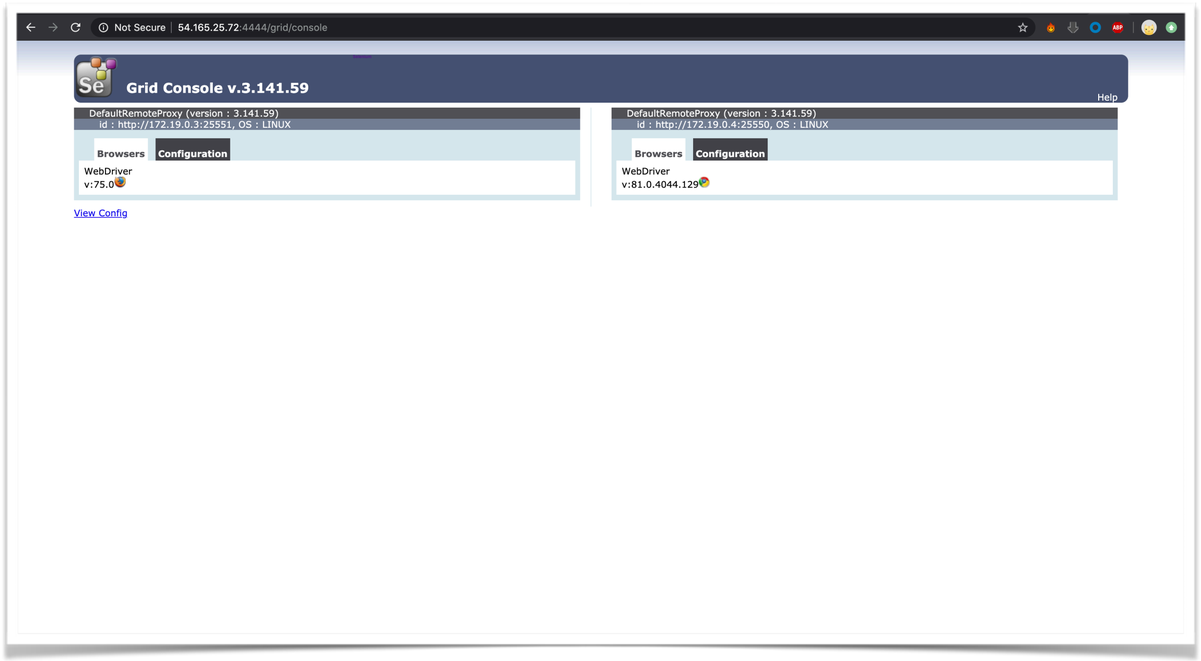
Note: The port defaults to 5555 if not specified whenever the "-role" option is provided and is not hub.

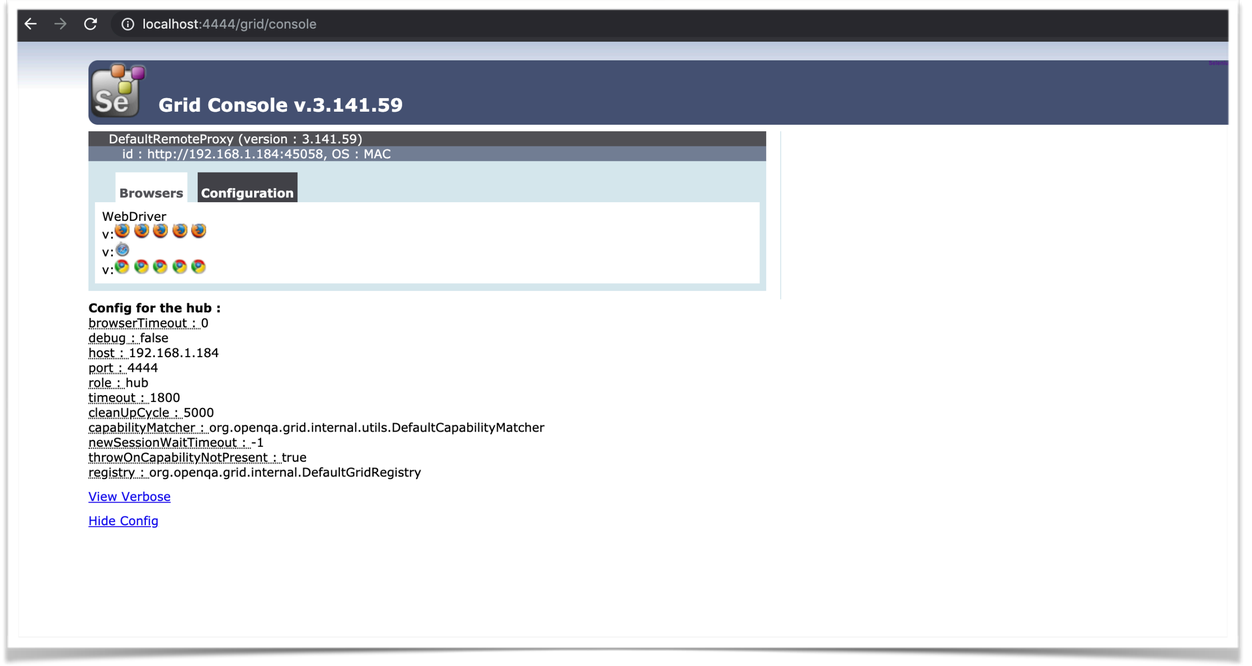
For backwards compatibility "wd" and "rc" roles are still a valid subset of the "node" role. But those roles limit the types of remote connections to their corresponding API, while "node" allows both RC and WebDriver remote connections.

Cybertek

Step 5: Check Grid console

Go to localhost:4444/grid/console in order to view your selenium grid console
Or to http://serverip:4444/grid/console





Cybertek

Step 6: Configure Driver class

Create RemoteWebDriver instance.
Provide hub URL address and DesiredCapabilities in order to specify node preferences.

```
case "remote-chrome":
    ChromeOptions chromeOptions = new ChromeOptions();
    try {
        URL url = new URL( spec: "http://3.89.140.35:4444/wd/hub");
        driverPool.set(new RemoteWebDriver(url, chromeOptions));
    } catch (MalformedURLException e) {
        e.printStackTrace();
    }
    break;
case "remote-firefox":
    FirefoxOptions firefoxOptions = new FirefoxOptions();
    try {
        URL url = new URL( spec: "http://3.92.180.21:4444/wd/hub");
        driverPool.set(new RemoteWebDriver(url, firefoxOptions));
    } catch (MalformedURLException e) {
        e.printStackTrace();
    }
    break;
```


Cybertek

Configure the nodes by command line

By default, starting the node allows for concurrent use of 11 browsers... : 5 Firefox, 5 Chrome, 1 Internet Explorer. The maximum number of concurrent tests is set to 5 by default.

To change this and other browser settings, you can pass in parameters to each -browser switch (each switch represents a node based on your parameters). If you use the -browser parameter, the default browsers will be ignored and only what you specify command line will be used.

-browser browserName=firefox,version=3.6,maxInstances=5,platform=LINUX

Cybertek

Configure the nodes by JSON

java -Dwebdriver.chrome.driver="chromedriver" jar selenium-server-standalone-<version>.jar -role node -hub http://localhost:4444/grid/register -nodeConfig nodeconfig.json

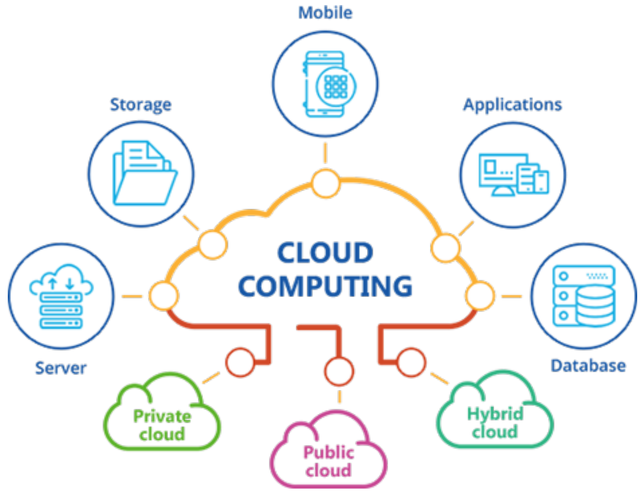
Cybertek

Terminologies used in Grid Configuration

role = When launching a node, it will forward the parameter to server on the node like, -role node.
host = Though it not needed and determined automatically. For some network configuration, network with VPN, specifying the host might be necessary.
port = the port the remote/hub will listen on. Default to 4444.
throwIn Capability Not Present = Default value is true; if no proxy is currently registered hub will reject the test request. On contrast, the request queued until a node supporting the capability which is added to the grid.
new Session Wait_Timeout = Default to no timeout (-1)ms after which a new test waiting for a node to become available will timeout else test will throw an exception before starting a browser.
hubConfig = It defines the hub properties in JSON format.
nodeConfig = It defines the node properties in JSON format.
cleanupCycle = It will check the timeout thread in ms.
node Timeout = The timeout in seconds prior to hub automatically ends the test so that browser will be released for another test to run.
browser Timeout= The browser gets hang after defined timeout in ms.
hub = It used to post the registration request using url – <http://localhost:4444/grid/register>
proxy =This will be used to represent the node. By default org.openqa.grid.selenium.proxy.DefaultRemoteProxy.
maxSession = Maximum number of sessions to run multiple tests at same time independently in different browsers.
register Cycle = It defines how often registered node will try to register itself again in ms,however, without restarting the node command file.
nodePolling = How often the hub checks if the node is still alive in ms.

Cloud computing - using remote data centers to store data, use remote machines for servers. Virtualization of resources, not relying on physical machines.
reasons to use cloud: easy get started, easy/cheap to maintain

who are main players in the market?
Amazon AWS
Microsoft Azure
Google Cloud
IBM cloud platform



Cloud computing has two meanings. The most common refers to running workloads remotely over the internet in a commercial provider's data center, also known as the "public cloud" model. Popular public cloud offerings—such as Amazon Web Services (AWS), Salesforce's CRM system, and Microsoft Azure—all exemplify this familiar notion of cloud computing. Today, most businesses take a multicloud approach, which simply means they use more than one public cloud service.

The second meaning of cloud computing describes how it works: a virtualized pool of resources, from raw compute power to application functionality, available on demand. When customers procure cloud services, the provider fulfills those requests using advanced automation rather than manual provisioning. The key advantage is agility: the ability to apply abstracted compute, storage, and network resources to workloads as needed and tap into an abundance of prebuilt services.

Cybertek

**AWS is cloud service provided by amazon
the have services such as virtual computing, data storage
with enhanced security features.**

Cloud computing has become an integral part of businesses across all industries. AWS is the most popular form. It improves efficiency and provides relief for any number of business practices. Companies using AWS have servers available instantly, and AWS provides various workloads, increased storage options, and enhanced security measures.

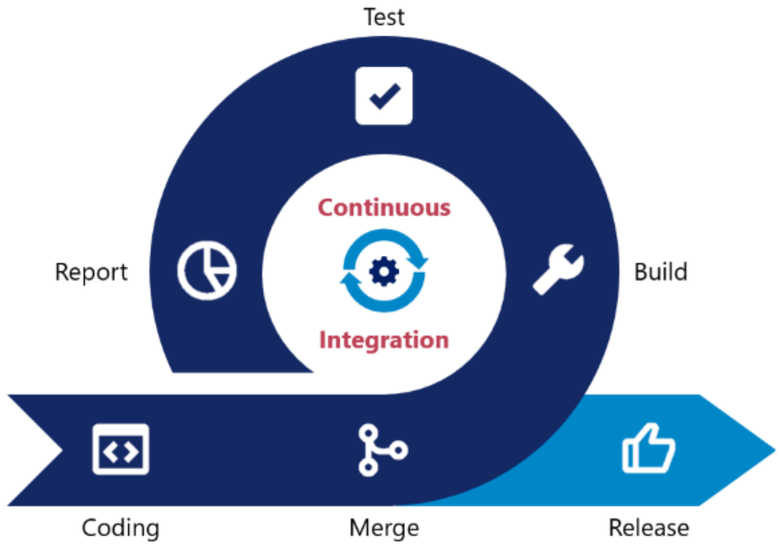
EC2 stands for Elastic Cloud Compute

It is one of popular services provided by Amazon
using EC2 we can host applications, run programs on them. They Virtual Machines
that we very easy to set up, scale, bring down.
Ec2 can also be created based on images known as AMI which makes them very easy to set up.
When EC2 created based on images, they will come preloaded by the applications, configurations from the original AMI
In my project out Jenkins is hosted on a EC2 server.

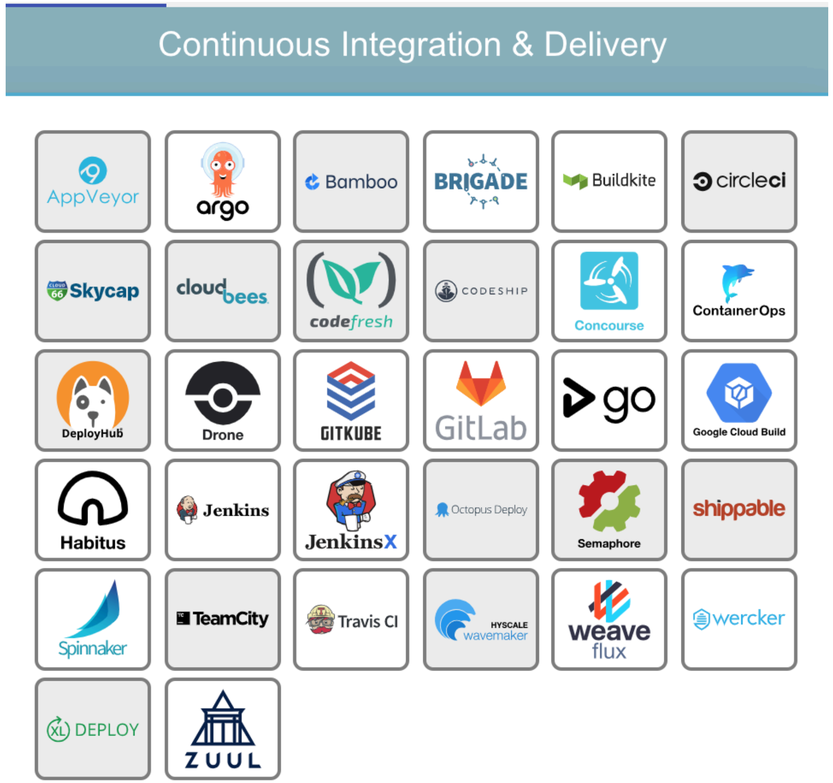
Amazon EC2 (Elastic Compute Cloud), one of Amazon Web Services' most well-known services, offers businesses the ability to run applications on the public cloud. Developers can create instances of virtual machines and easily configure the capacity scaling of instances using the EC2 web interface.

EC2 also allows users to build apps to automate scaling according to changing needs and peak periods, and makes it simple to deploy virtual servers and manage storage, lessening the need to invest in hardware and helping streamline development processes.

EC2 setup involves creating an Amazon Machine Image (AMI), which includes an operating system, apps, and configurations. That AMI is loaded to the Amazon Simple Storage Service (S3), and it's registered with EC2, at which point users can launch virtual machines as needed.



Continuous integration it is a process/ set of processes purpose of which is to
integrate the new code/changes to the application and deploy
For automating the process several tools are available such as Jenkins, Bamboo, TravisCI…



Cybertek

Jenkins

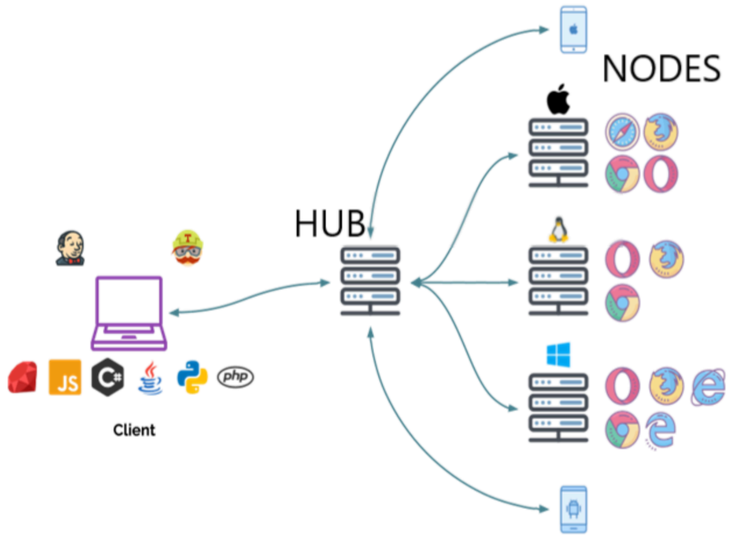
Continuous Integration tool
Written in java. Open source and free to use
It has big community support
Boasts many plugins for different purposes. Using those wide range of plugins Jenkins can be used for automating deployments, building application etc.
In can installed and run from almost anywhere. It can in an VM, in personal machine (don’t do it), docker containers.

Cybertek

Docker containers allows creating virtual environments that make is easy to build, deploy and host applications.
when building docker containers we can create it with all the required applications, dependencies and ship them as one package. It is an isolated environment

In my framework, our selenium grid runs on a docker.
We use the selenium official docker image to create the docker container

Cybertek



Selenium Grid is a smart proxy server that allows Selenium tests to route commands to **remote** web browser instances. Its aim is to provide an easy way to run tests in parallel on multiple machines.

Hub conducts a concurrent execution of tests on multiple machines, managing different browsers centrally, instead of conducting different tests for each of them. Selenium Grid makes cross browser testing easy as a single test can be carried on multiple machines and browsers, all together, making it easy to analyze and compare the results.

The two major components of Selenium Grid are:

Hub is a server that accepts the access requests from the WebDriver client, routing the JSON test commands to the remote drives on nodes. It takes instructions from the client and executes them remotely on the various nodes in parallel

Node is a remote device that consists of a native OS and a remote WebDriver. It receives requests from the hub in the form of JSON test commands and executes them using WebDriver

Using grid does not make the framework run tests in parallel automatically. It also does not make it cross browser automatically. These two must be configured in the framework code level. In the code we make the tests run in parallel. Grid just provides browsers fort those parallel tests

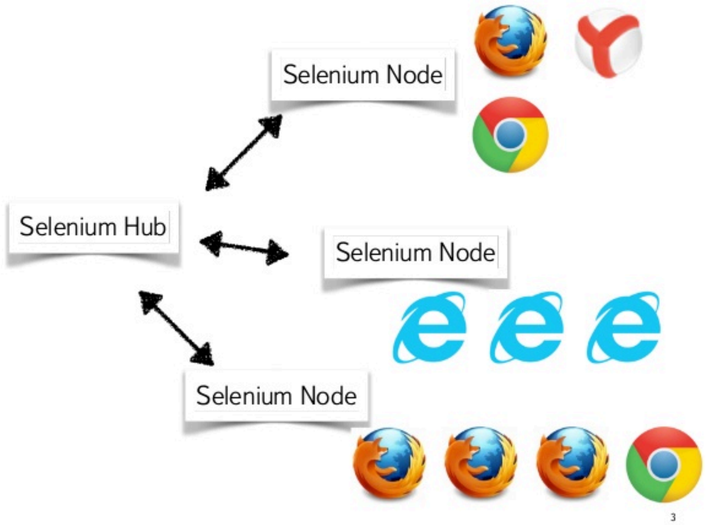
Cybertek

Parallel testing and grid

Parallelization is achieved differently based on what tools we are using in our framework. In my automation framework we use Cucumber with Junit. Cucumber has built in support for parallelization since version 4.0. This is done in the pom file using plugins.

To make Cucumber Junit test in parallel we need to run test using maven surefire or maven failsafe plugin. In the configuration make parallel equal to methods to run feature files in parallel. And control how many parallel instances you want to run, use the threadCount option.

```
<plugin>
    <groupId>org.apache.maven.plugins</groupId>
    <artifactId>maven-surefire-plugin</artifactId>
    <version>3.0.0-M4</version>
    <configuration>
        <includes>
            <include>**/CukesRunner.java</include>
        </includes>
        <testFailureIgnore>true</testFailureIgnore>
        <runOrder>alphabetical</runOrder>
        <parallel>methods</parallel>
        <threadCount>3</threadCount>
    </configuration>
</plugin>
```



Cybertek

Cross browser testing

Cross browser testing is also achieved in the code. Grid only gives you the browser my code asks for. SO framework must support cross browser testing

Pass browser from properties file

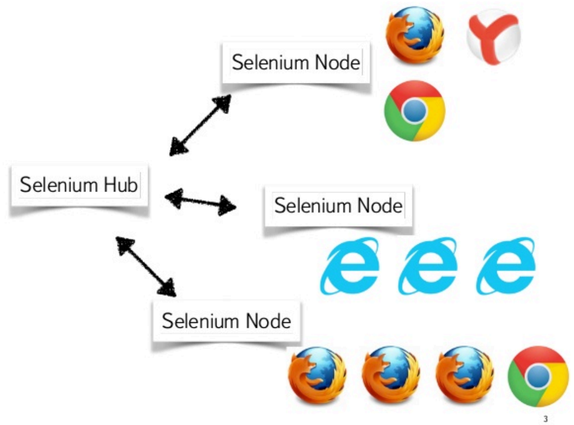
configuration.properties

```
browser=remote-chrome
url=http://library2.cybertekschool.com/
env=qa2
```

Pass browser from terminal

```
if (System.getProperty("BROWSER") == null) {
    browser = ConfigurationReader.getProperty("browser");
} else {
    browser = System.getProperty("BROWSER");
}
```

```
mvn test -DBROWSER=firefox
```



Cybertek

Setting up code for grid

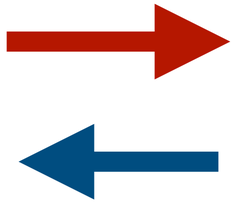
```
case "remote-chrome":
    try {
        URL url = new URL( spec: "http://54.161.125.141:4444/wd/hub");
        ChromeOptions chromeOptions = new ChromeOptions();
        WebDriver driver = new RemoteWebDriver(url, chromeOptions);
    } catch (MalformedURLException e) {
        e.printStackTrace();
    }
    break;
case "remote-firefox":
    try {
        URL url = new URL( spec: "http://54.161.125.141:4444/wd/hub");
        FirefoxOptions firefoxOptions = new FirefoxOptions();
        WebDriver driver = new RemoteWebDriver(url, firefoxOptions);
    } catch (MalformedURLException e) {
        e.printStackTrace();
    }
```

In selenium class RemoteWebDriver is used to create a browser in the selenium grid.
RemoteWebDriver needs to know :
1. The grid url. For this use the URL class which we create by passing the ip address and the port number of the HUB
2. What browser we want. Based on what browser we need, we create the chromeoptions or the firefoxoptions objects

We create object of RemoteWebDriver by passing the URL and chromeoptions/ firefoxoptions as arguments

Jenkins

Runs smoke tests.
Get the code from git



GitHub

Code is stored here

Jenkins runs the test by passing maven command

```
mvn test -Dcucumber.filter.tags=@regression
```

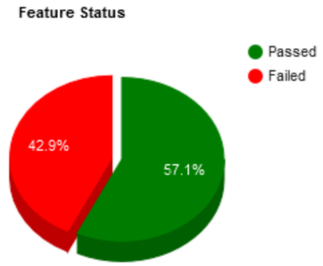
Smoke tests start running on grid due to set up in code

```
URL url = new URL( spec: "http://54.161.125.141:4444/wd/hub");
ChromeOptions chromeOptions = new ChromeOptions();
WebDriver driver = new RemoteWebDriver(url, chromeOptions);
```

Url points the grid hub.
The grid is running in docker

Amazon
EC2

Docker is located in AWS EC2 (Linux)



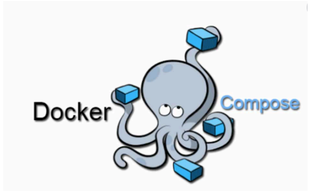
Jenkins generates HTML reports and sends results to the team: email or slack

Cybertek

BRAKE 4. 00

How to create automated tests job in Jenkins?

1. Source code management —> this is where we point to the git repository
2. Trigger —> this is where we say how often we want to run
3. Build —> this is where specify who to run the project. My framework Cucumber BDD framework that is built with Maven. So I run my test as a maven by selecting Build option Invoke top level maven targets. For the command In the command:
   test -Dcucumber.filter.tags=@smoke -Dbrowser=remote-chrome -Denv=qa1
4. Add post build actions for reporting and sending results. In Jenkins I have 1 html report and 2 ways sending results.
   4.1 Cucumber HTML Report —> I add post build step Cucumber reports. Cucumber reports a plugin in Jenkins that generates HTML reports for cucumber test with test steps and screenshots
   4.2 Email notification —> in Jenkins job I have configured another post build action Editable Email Notification which sends email to the team about the build results.
   4.3 Slack notification —> we also use notifications. My team prefers slack notifications. I added post build action Slack notifications. This is also another plugin in Jenkins.




```
case "remote-chrome":
    try {
        URL url = new URL( spec: "http://54.161.125.141:4444/wd/hub");
        ChromeOptions chromeOptions = new ChromeOptions();
        WebDriver driver = new RemoteWebDriver(url, chromeOptions);
    } catch (MalformedURLException e) {
        e.printStackTrace();
    }
    break;
case "remote-firefox":
    try {
        URL url = new URL( spec: "http://54.161.125.141:4444/wd/hub");
        FirefoxOptions firefoxOptions = new FirefoxOptions();
        WebDriver driver = new RemoteWebDriver(url, firefoxOptions);
    } catch (MalformedURLException e) {
        e.printStackTrace();
    }
```

container —> application that allows packaging all the programs, dependencies required to build, run or deploy certain software

docker —> program using which we can create a container, run, stop basically manages a container.

docker-compose —> program using which we can create, run, stop basically manages many containers.

for creating docker containers with docker compose, we use docker compose yml file. This file contains configuration settings for the container that we want to create

image —>description, blue print, base template of how a container should be created

docker-compose up —> starts docker containers based on the given yaml file

docker-compose down —> stops the contains

docker-compose restart —> restarts the contains

docker-compose ps —> get status of the containers

Where is the grid: In my project we create the selenium grid using the selenium grid docker compose file.

How do you start the grid: In EC2 server, I have the docker compose file, it has configuration for the grid and also nodes. I ssh into the EC2 (connect to EC2 using ssh). Change directory to the location of the doker compose file. I start the grid by running the **docker-compose up** command. I can use docker-compose scale chrome=5 firefox=2 to scale the number of available browsers.

#

How do you configure the code for grid: Then in the selenium code point to the this docker by passing the ip address of the EC2 and the port number 4444

Cybertek