

Introduction to Linux Operating System

Harvesting, Storing, and Retrieving Data

What is an Operating System (OS)

The most crucial program that runs on a computer is the operating system. It controls the memory and operations of the computer, as well as all of its software and hardware. It also allows you to communicate with the computer even if you don't understand its language. A computer is useless without an operating system.

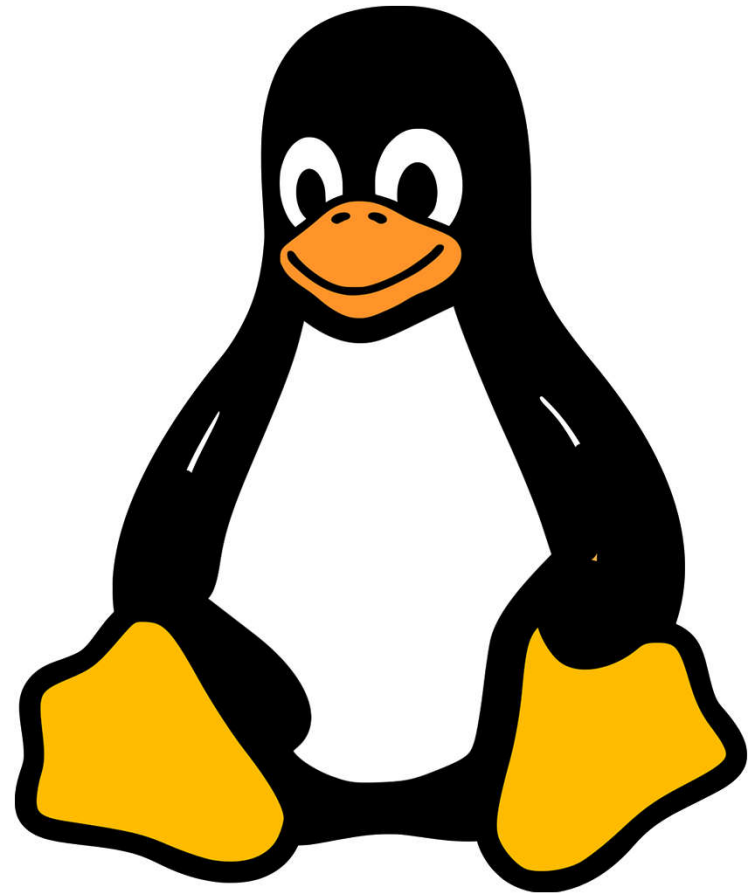
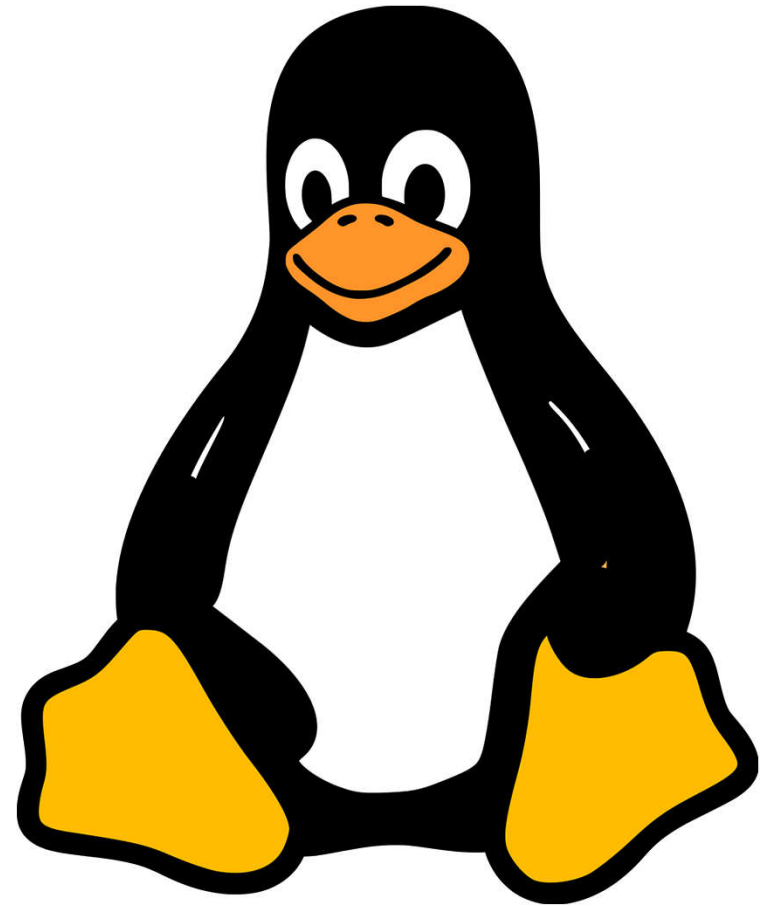
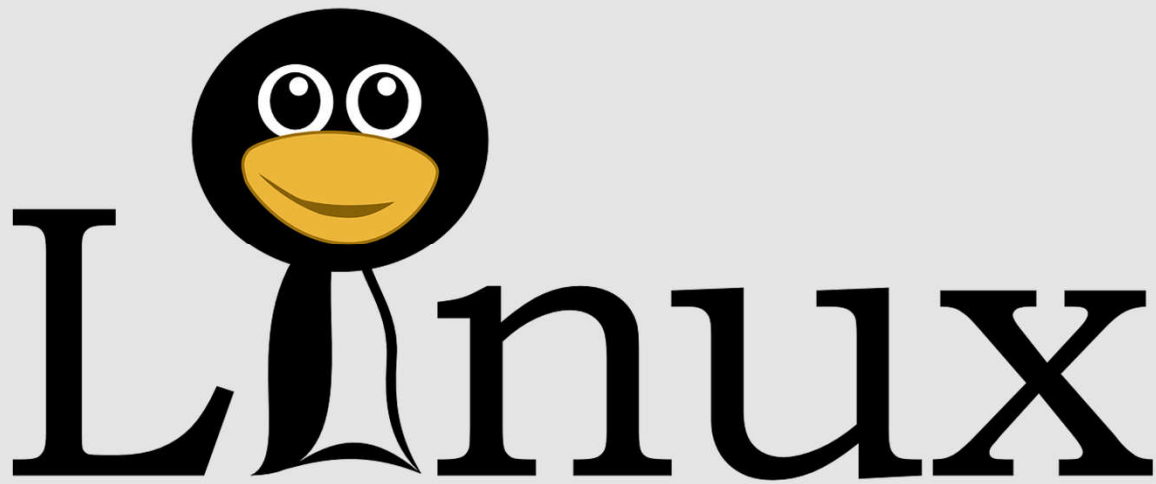


Image by [OpenClipart-Vectors](#) from [Pixabay](#)

What is Linux

Linux (pronounced LINN-ux) is a group of open-source operating systems that anybody can modify and distribute all over the world. Unlike proprietary software such as Windows, which can only be modified by the corporation that owns it, open-source software can be modified by anybody. Linux has the advantages of being free and having a wide variety of distributions (or versions) to select from.

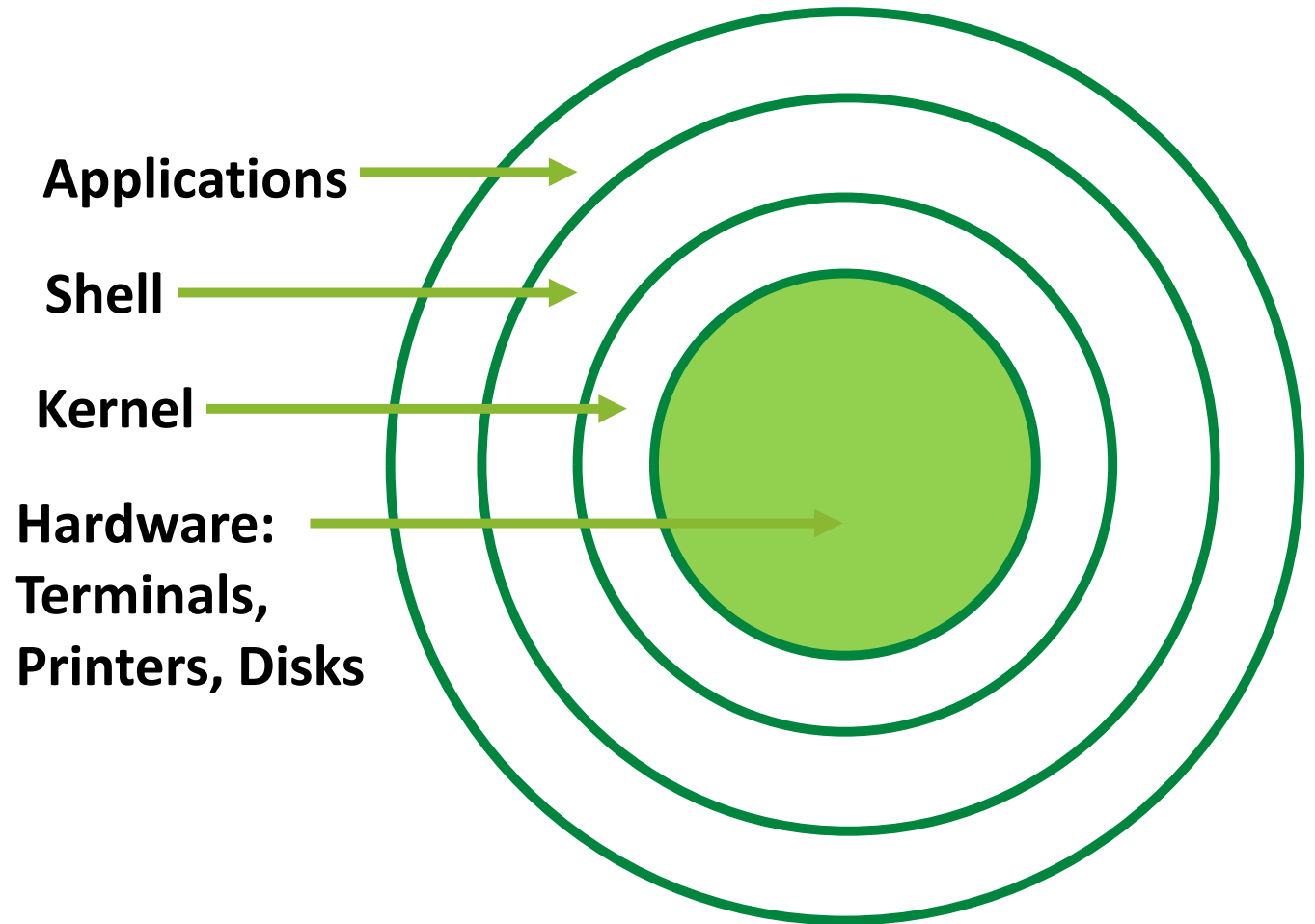
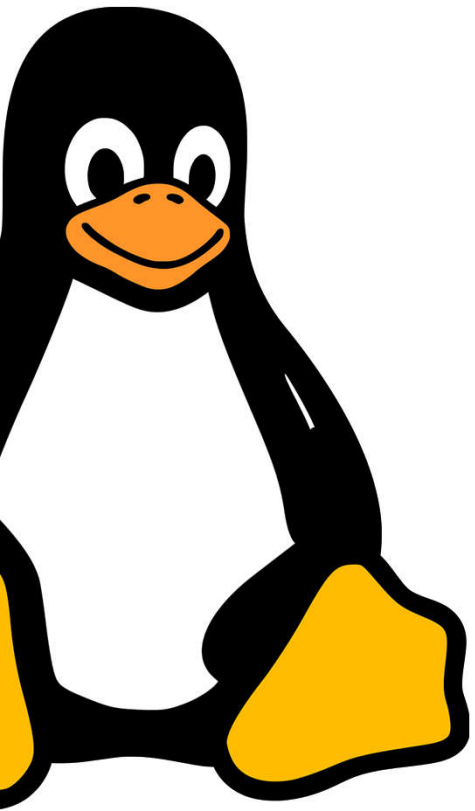




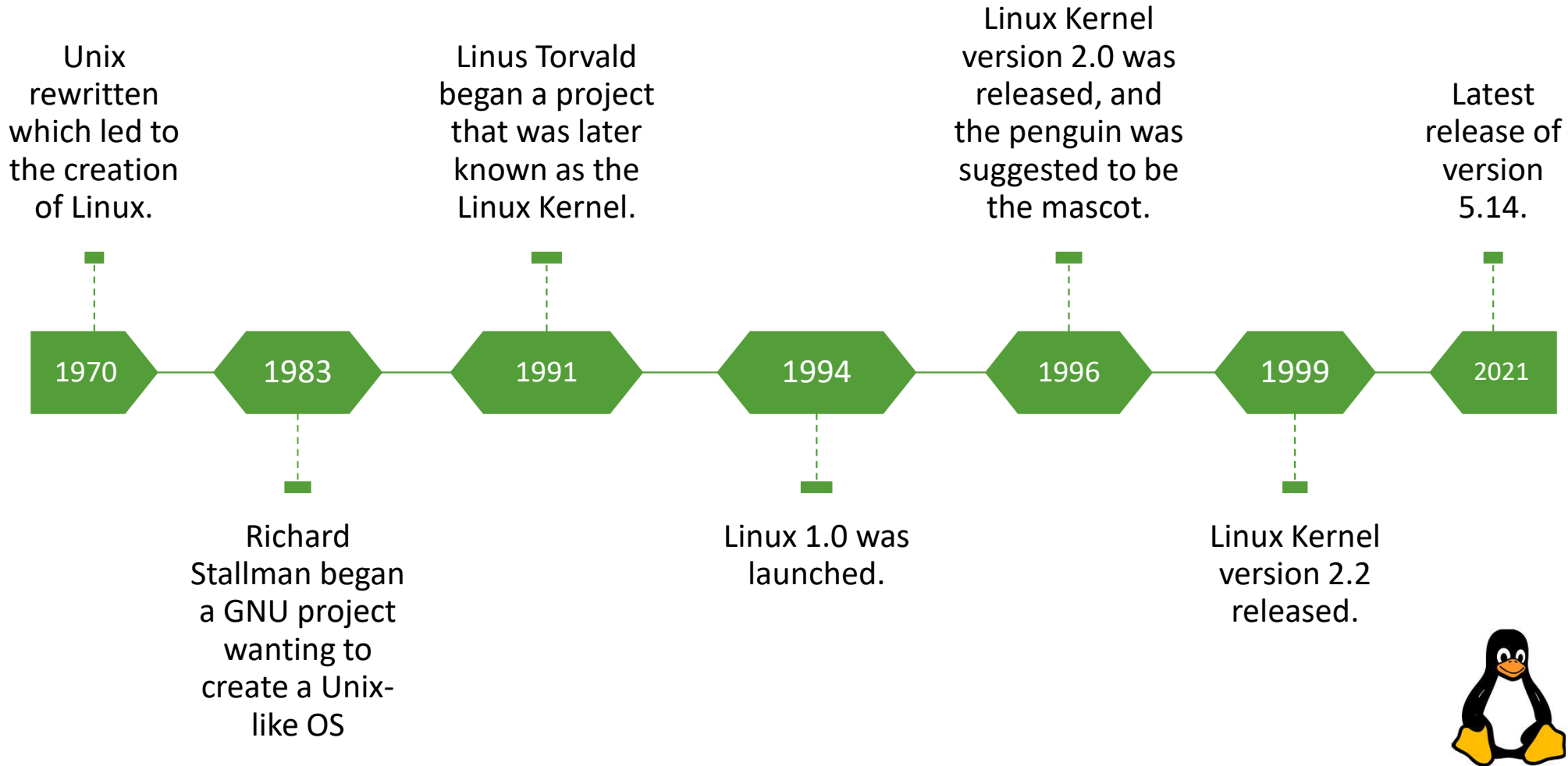
Why Should I Care about
Linux

- High Security
- High Stability
- Ease of Maintenance
- Runs on any Hardware
- Free
- Open-source
- Easy to Use
- Customization
- Support

Simplified Architecture of Linux

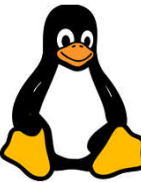


Brief History of Linux



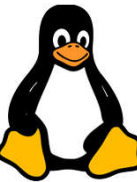
Role and Function of Linux

- Application Platform
- Hardware Moderator
- Data Storage
- Security
- Connectivity



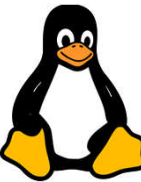
Bash – the Linux shell

- BASH is a programming/scripting language
- BASH shell is the Linux equivalent of the Windows cmd
- BASH is a command processor that typically runs in a text window, where the user types commands that cause actions
- BASH runs script (Python, Perl, etc)
- The Golden Rule of BASH – when you do not know what a command does:
 - Man stands for manual
 - man ls
 - man cd
 - nan grep

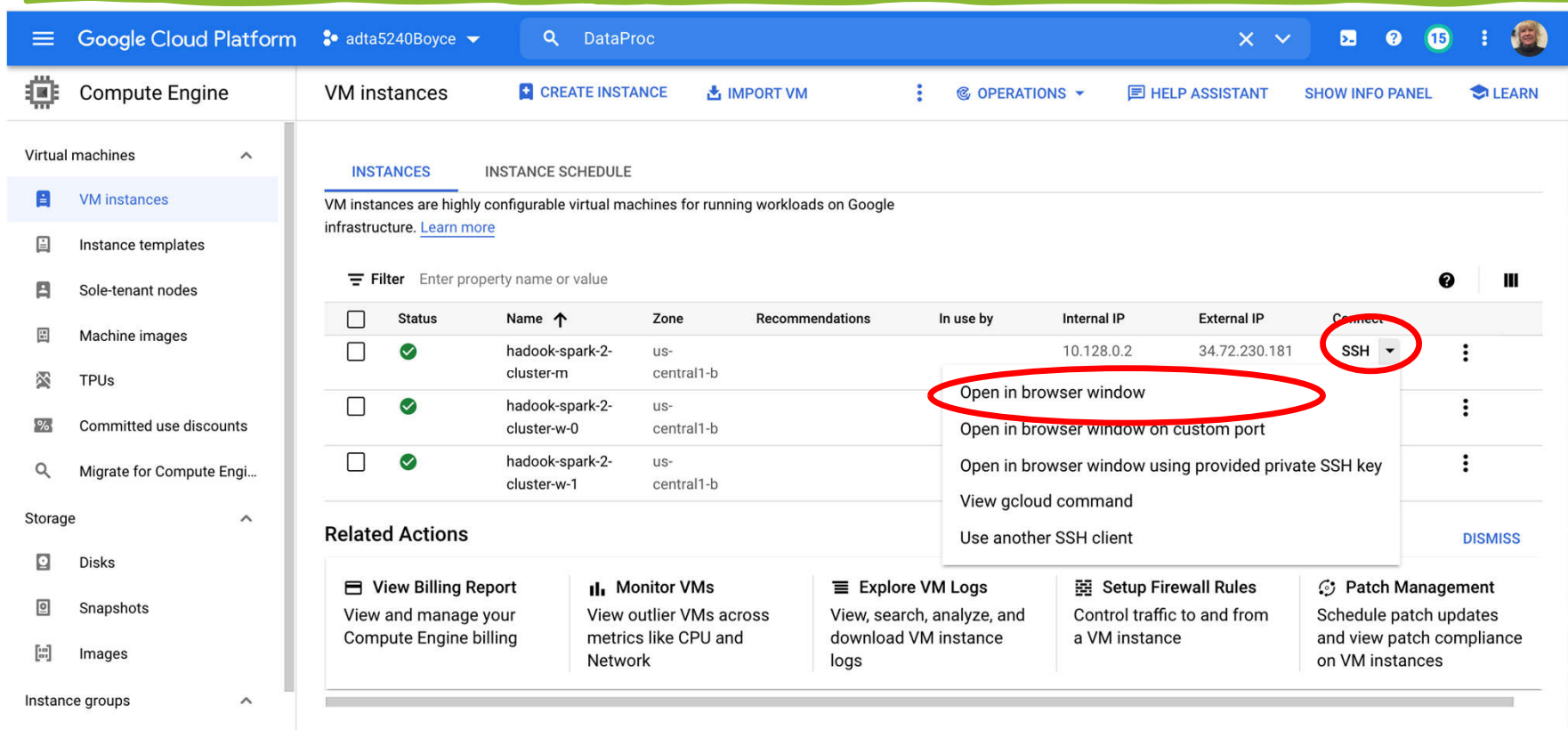


Exploring Remote Virtual Machine in the Cloud

- The remote virtual machine in the cloud has Linux as its installed OS.
- Linux Operating System: A multi-user computing system
- A multi-user operating system allows many different users to take advantage of the computer's resources simultaneously.
- The multi-user operating system must make sure:
 - The requirements of various users are balanced.
 - Each of the programs they are using has sufficient and separate resources so that a problem caused by one user does not affect the entire community of users.



Exploring Remote Virtual Machine in the Cloud



The screenshot shows the Google Cloud Platform interface for VM instances. The left sidebar contains navigation links for Virtual machines, Instance templates, Sole-tenant nodes, Machine images, TPUs, Committed use discounts, and Migrate for Compute Engi... The main content area displays a table of VM instances. A context menu is open over the 'SSH' button for the first instance, 'hadoop-spark-2-cluster-m'. The menu options are: 'Open in browser window' (highlighted with a red circle), 'Open in browser window on custom port', 'Open in browser window using provided private SSH key', 'View gcloud command', and 'Use another SSH client'. Below the table, there are 'Related Actions' such as 'View Billing Report', 'Monitor VMs', 'Explore VM Logs', 'Setup Firewall Rules', and 'Patch Management'.

Google Cloud Platform | adta5240Boyce | DataProc

Compute Engine | VM instances | CREATE INSTANCE | IMPORT VM | OPERATIONS | HELP ASSISTANT | SHOW INFO PANEL | LEARN

Virtual machines

- VM instances
- Instance templates
- Sole-tenant nodes
- Machine images
- TPUs
- Committed use discounts
- Migrate for Compute Engi...

Storage

- Disks
- Snapshots
- Images

Instance groups

INSTANCES | INSTANCE SCHEDULE

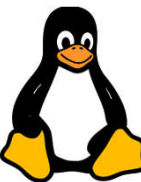
VM instances are highly configurable virtual machines for running workloads on Google infrastructure. [Learn more](#)

Filter Enter property name or value

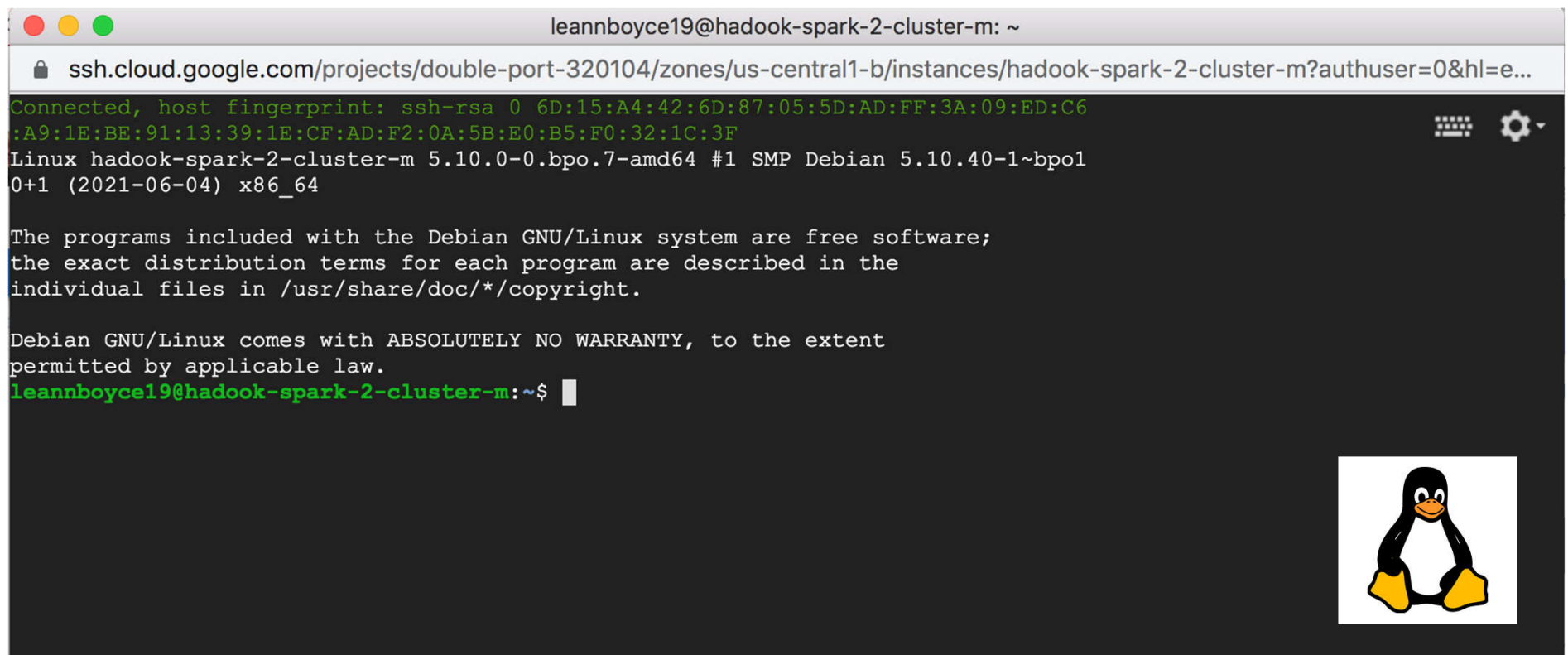
	Status	Name ↑	Zone	Recommendations	In use by	Internal IP	External IP	Connect	
<input type="checkbox"/>	✓	hadoop-spark-2-cluster-m	us-central1-b			10.128.0.2	34.72.230.181	SSH	⋮
<input type="checkbox"/>	✓	hadoop-spark-2-cluster-w-0	us-central1-b						⋮
<input type="checkbox"/>	✓	hadoop-spark-2-cluster-w-1	us-central1-b						⋮

Related Actions

- View Billing Report**
View and manage your Compute Engine billing
- Monitor VMs**
View outlier VMs across metrics like CPU and Network
- Explore VM Logs**
View, search, analyze, and download VM instance logs
- Setup Firewall Rules**
Control traffic to and from a VM instance
- Patch Management**
Schedule patch updates and view patch compliance on VM instances

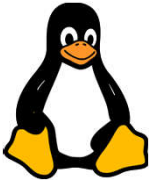


Exploring Remote Virtual Machine in the Cloud



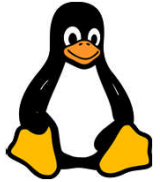
A terminal window titled "leannboyce19@hadoopk-spark-2-cluster-m: ~" showing an SSH connection to a Google Cloud VM. The address bar displays the URL: `ssh.cloud.google.com/projects/double-port-320104/zones/us-central1-b/instances/hadoopk-spark-2-cluster-m?authuser=0&hl=e...`. The terminal output shows the host fingerprint, the Linux distribution (Debian 5.10.40-1~bpo10+1), and the system architecture (x86_64). It also displays the Debian GNU/Linux warranty disclaimer. The prompt is `leannboyce19@hadoopk-spark-2-cluster-m:~$`. A small penguin icon is visible in the bottom right corner of the terminal window.

```
leannboyce19@hadoopk-spark-2-cluster-m: ~  
ssh.cloud.google.com/projects/double-port-320104/zones/us-central1-b/instances/hadoopk-spark-2-cluster-m?authuser=0&hl=e...  
Connected, host fingerprint: ssh-rsa 0 6D:15:A4:42:6D:87:05:5D:AD:FF:3A:09:ED:C6  
:A9:1E:BE:91:13:39:1E:CF:AD:F2:0A:5B:E0:B5:F0:32:1C:3F  
Linux hadook-spark-2-cluster-m 5.10.0-0.bpo.7-amd64 #1 SMP Debian 5.10.40-1~bpo1  
0+1 (2021-06-04) x86_64  
  
The programs included with the Debian GNU/Linux system are free software;  
the exact distribution terms for each program are described in the  
individual files in /usr/share/doc/*/copyright.  
  
Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent  
permitted by applicable law.  
leannboyce19@hadoopk-spark-2-cluster-m:~$
```



Exploring Remote Virtual Machine in the Cloud: Two users in Linux System

- “**root**” is the **super-user** in a Linux system
- “**root**” is the first user created during the process of installing a Linux OS.
- Most administration tasks, such as adding new users or managing file systems, are required to be done with the privilege of the “root” user.
- “**root**” has unlimited powers and can do anything on the system.
- The “**root**” account: Is also known as the “super-user” account



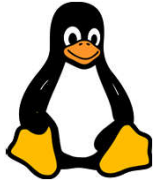
Exploring Remote Virtual Machine in the Cloud

- “**User**”: The normal user in the Linux system
- Based on the permission modes set for a folder, a **normal user**:
- May or may not access (read) the contents
- May or may not change (write) the contents
- May or may not execute the executable contents
- Based on the permission set for a file, a normal user:
- May or may not access (read) the contents of the file
- May or may not change (write) the contents of the file
- May or may not execute the file if it is an executable one
- To find out the permission modes of a directory/folder or a file::
 - Use the command line: “**ls -l**<directory/file name>

Exploring Remote Virtual Machine in the Cloud

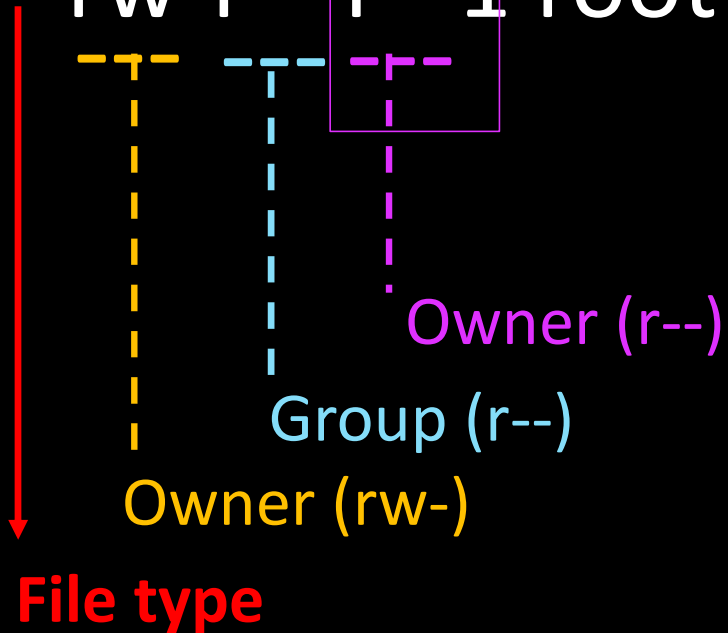
Permission	Description
Owner	Permission used by the assigned owner of the file or directory
Group	Permissions used by members of the group that owns the file or directory
Other	Permissions used by all users other than the file owner, and members of the group that owns the file or the directory

Permission	Access for a file	Access for a directory
Read	You can display file contents and copy the file.	You can list the directory contents with the ls command.
Write	You can modify the file contents.	You can modify the contents of a directory, such as by deleting a file. You must also have the execute permission for this to happen.
Execute	You can execute the file if it is an executable. You can execute a shell script if you also have read and execute permissions.	You can use the cd command to access the directory. If you also have read access, you can run the ls l command on the directory to list contents. If you do not have read access, you can run the ls command as long as you know the file name.



Exploring Remote Virtual Machine in the Cloud

```
-rw-r--r-- 1 root root 0 July 19 23:49 file
```



Permission Type

r = Readable

w = Writeable

x = Executable

- = Denied

Exploring Remote Virtual Machine in the Cloud

Few Basic Linux Command Lines

whoami :

- Display the username of the current user

pwd:

- Print working directory: Display the full path of the current working directory

cd<sub directory name> name>:

- Change directory (forward) to a sub folder

cd<full path of a directory>:

- Change directory to the folder of which the full path is specified

cd...:

- Change directory (backward) to the folder that is immediately one level before this one

- NOTES : cd .." is the real command line to go a folder.

Exploring Remote Virtual Machine in the Cloud

Few Basic Linux Command Lines

mkdir<a directory name>:

- Create a new directory

rmdir<a directory name> :

- Delete a directory if it is empty

rm<a file name>:

- Delete a file

rm-rf<a directory name>:

- Delete all the files and sub-directories of a directory

rm*.txt:

- Delete all the text files with the suffix .txt

Exploring Remote Virtual Machine in the Cloud

Linux command line to explore remote server

```
Connected, host fingerprint: ssh-rsa 0 6D:15:A4:42:6D:87:05:5D:AD:FF:3A:09:ED:C6
:A9:1E:BE:91:13:39:1E:CF:AD:F2:0A:5B:E0:B5:F0:32:1C:3F
Linux hadook-spark-2-cluster-m 5.10.0-0.bpo.7-amd64 #1 SMP Debian 5.10.40-1~bpo1
0+1 (2021-06-04) x86_64
```

```
The programs included with the Debian GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.
```

```
Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
```

```
Last login: Sun Jul 18 01:20:21 2021 from 35.235.244.34
```

```
leannboyce19@hadook-spark-2-cluster-m:~$ whoami
```

```
leannboyce19
```

```
leannboyce19@hadook-spark-2-cluster-m:~$ pwd
```

```
/home/leannboyce19
```

```
leannboyce19@hadook-spark-2-cluster-m:~$ ls -l
```

```
total 0
```

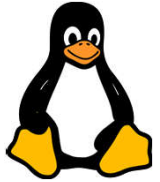
```
leannboyce19@hadook-spark-2-cluster-m:~$ mkdir DATA
```

```
leannboyce19@hadook-spark-2-cluster-m:~$
```

Exploring Remote Virtual Machine in the Cloud

Linux command line to explore remote server

```
leannboyce19@hadoop-spark-2-cluster-m:~$ mkdir DATA
leannboyce19@hadoop-spark-2-cluster-m:~$ ls -l
total 4
drwxr-xr-x 2 leannboyce19 leannboyce19 4096 Jul 18 01:22 DATA
leannboyce19@hadoop-spark-2-cluster-m:~$ cd DATA
leannboyce19@hadoop-spark-2-cluster-m:~/DATA$ pwd
/home/leannboyce19/DATA
leannboyce19@hadoop-spark-2-cluster-m:~/DATA$ ls -l
total 0
leannboyce19@hadoop-spark-2-cluster-m:~/DATA$
```



Exploring Remote Virtual Machine in the Cloud

Don't forget to turn off your clusters in GCP!!