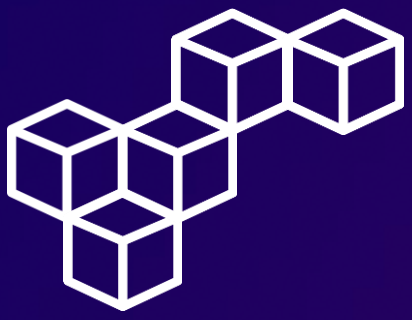
AWS re:Start

LAB

Working with Files

WEEK 2

aws re/start



Overview

Working with files in a Linux environment involves a range of essential tasks to manage data effectively. One common task is creating backups of entire folder structures, ensuring that data is securely stored and can be restored if needed. Logging the creation of backups with the date, time, and file name adds an extra layer of organization and accountability, making it easier to track and manage backups over time.

Additionally, transferring backup files to another folder or location is crucial for data redundancy and disaster recovery planning. This process ensures that backup files are stored in separate locations, reducing the risk of data loss due to hardware failures or other unforeseen events. By integrating these practices into file management workflows, users can safeguard their data and maintain efficient data management practices in Linux.

Note: This lab was made using Windows Subsystem for Linux.

Topics covered

- Create a backup file of an entire folder structure using tar
- Log the creation of the backup in a file with the date, time, and file name of the backup file
- Transfer the backup file to another folder

Task 1

Use SSH to connect to an Amazon Linux EC2 instance

Initial Preparations

In the AWS Management Console, select the EC2 instance and make note of the **Public IPv4 address**.

Download the private key file **labsuser.pem**. Change to the Downloads directory and modify the permissions on the key to be read-only (r--------).

Connect to the instance using SSH

Establish a connection to the EC2 instance using the ssh command, the key and the instance's public IPv4 address.

```
support@HP-Pavilion-Laptop:~/Downloads$ ssh -i labsuser.pem ec2-user@35.94.9.157
The authenticity of host '35.94.9.157 (35.94.9.157)' can't be established.
ED25519 key fingerprint is SHA256:Em7q7M/ig71zZY7FTDqlTtWnhyM6+2+yAnSBW1gd4X0.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added '35.94.9.157' (ED25519) to the list of known hosts.
   ,     #_
   ~\_  ####_        Amazon Linux 2
  ~~  \_#####\
  ~~     \###|       AL2 End of Life is 2025-06-30.
  ~~       \#/ ___
   ~~       V~' '->
    ~~~         /    A newer version of Amazon Linux is available!
      ~~._.   _/
         _/ _/       Amazon Linux 2023, GA and supported until 2028-03-15.
       _/m/'           https://aws.amazon.com/linux/amazon-linux-2023/

[ec2-user@ip-10-0-10-139 ~]$
```

aws re/start

Task 2

Create a backup

Check the current folder structure

To display the contents of the **CompanyA** top-level folder enter the command ls –R CompanyA.

```
[ec2-user@ip-10-0-10-139 ~]$ pwd
/home/ec2-user
[ec2-user@ip-10-0-10-139 ~]$ ls -R CompanyA
CompanyA:
Employees  Finance  HR  IA  Management  SharedFolders

CompanyA/Employees:
Schedules.csv

CompanyA/Finance:
Salary.csv

CompanyA/HR:
Assessments.csv  Managers.csv

CompanyA/IA:

CompanyA/Management:
Promotions.csv  Sections.csv

CompanyA/SharedFolders:
[ec2-user@ip-10-0-10-139 ~]$
```

Backup the folder structure

To back up the entire **CompanyA** folder structure recursively, enter tar -csvpzf backup.CompanyA.tar.gz CompanyA.

```
[ec2-user@ip-10-0-10-139 ~]$ tar -csvpzf backup.CompanyA.tar.gz CompanyA
CompanyA/
CompanyA/Management/
CompanyA/Management/Sections.csv
CompanyA/Management/Promotions.csv
CompanyA/Employees/
CompanyA/Employees/Schedules.csv
CompanyA/Finance/
CompanyA/Finance/Salary.csv
CompanyA/HR/
CompanyA/HR/Managers.csv
CompanyA/HR/Assessments.csv
CompanyA/IA/
CompanyA/SharedFolders/
[ec2-user@ip-10-0-10-139 ~]$ ls
CompanyA  backup.CompanyA.tar.gz
[ec2-user@ip-10-0-10-139 ~]$
```

Task 3

Log the backup

Create a log file

Inside the **SharedFolders** directory, create a log file named **backups.csv** for logging the date, time, and file name of the backup tar file that you created **backup.CompanyA.tar.gz**. This log file indicates when you created backups and could be useful to avoid creating unnecessary backups in the future.

```
[ec2-user@ip-10-0-10-139 ~]$ cd /home/ec2-user/CompanyA
[ec2-user@ip-10-0-10-139 CompanyA]$ touch SharedFolders/backups.csv
[ec2-user@ip-10-0-10-139 CompanyA]$ echo "08 Apr 08 2024, 08:12, backup.CompanyA.tar.gz" | sudo tee SharedFolders/backups.csv
08 Apr 08 2024, 08:12, backup.CompanyA.tar.gz
[ec2-user@ip-10-0-10-139 CompanyA]$
```

Review the log file

To view the content of the newly created log file **backups.csv**, enter the command cat SharedFolders/backups.csv.

```
[ec2-user@ip-10-0-10-139 CompanyA]$ cat SharedFolders/backups.csv
08 Apr 08 2024, 08:12, backup.CompanyA.tar.gz
[ec2-user@ip-10-0-10-139 CompanyA]$
```

Task 4

Move the backup file

Transfer the backup file

Transfer the backup file from the **/home/ec2-user** directory to the **/home/ec2-user/CompanyA/IA** folder using the mv command.

```
[ec2-user@ip-10-0-10-139 CompanyA]$ pwd
/home/ec2-user/CompanyA
[ec2-user@ip-10-0-10-139 CompanyA]$ mv ../backup.CompanyA.tar.gz IA/
[ec2-user@ip-10-0-10-139 CompanyA]$
```

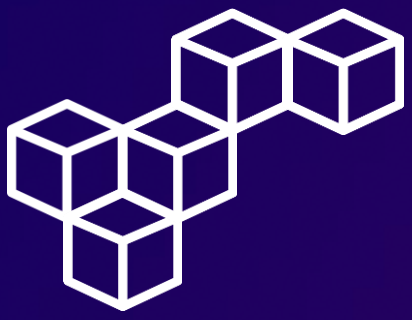
Verify backup location

Verify that the file **backup.CompanyA.tar.gz** was successfully moved to the **/home/ec2-user/CompanyA/IA** folder.

```
[ec2-user@ip-10-0-10-139 CompanyA]$ ls . IA
.:
Employees  Finance  HR  IA  Management  SharedFolders

IA:
backup.CompanyA.tar.gz
[ec2-user@ip-10-0-10-139 CompanyA]$
```

aws re/start



Conclusions

The tar & gzip commands

Creating backups using tools like tar and gzip streamlines the process of preserving entire folder structures. This method ensures data integrity and reduces storage space through compression, making it efficient for handling large volumes of data.

Logging the creation of backups

Logging backup creation activities is crucial for maintaining an organized record of backup events. This practice aids in tracking backup schedules, identifying data included in each backup, and facilitating troubleshooting during data recovery processes.

Transferring backup files

Transferring backup files in .tar.gz format offers convenience and space efficiency. It maintains file structures and permissions while reducing the risk of data loss by storing backups in multiple locations. This redundancy enhances data availability and resilience against unforeseen events.

aws re/start

Cristhian Becerra

cristhian-becerra-espinoza

+51 951 634 354

cristhianbecerra99@gmail.com

Lima, Peru