

# AUTOMATIZACIÓN DE BACKUPS EN TIEMPO REAL CON SUBIDA A UNA MÁQUINA VIRTUAL EN AZURE

Tema: Automatización de tareas en servidores locales y en la nube.



Miguel Gutiérrez García - U0295650 Adrian Dumitru - U0295652

## Índice

- 1. Introducción
- 2. Objetivo del proyecto
- 3. Infraestructura utilizada
- 4. Herramientas instaladas y configuraciones previas
- 5. Script de vigilancia y backup en tiempo real
- 6. Conexión de las máquinas mediante ssh
- 7. Automatización del Backup en Tiempo Real
- 8. Ejemplo de uso
- 9. Posibles mejoras
- 10. Conclusiones

#### 1. Introducción

Este proyecto desarrolla un sistema automatizado que permite realizar backups en tiempo real en una máquina virtual con AlmaLinux. El sistema detecta cambios en un directorio específico y, en caso de modificaciones, genera un archivo comprimido que se transfiere automáticamente a una máquina virtual ubicada en Azure. A diferencia de soluciones basadas en cron, este sistema se basa en la detección de eventos en tiempo real.

## 2. Objetivo del proyecto

El objetivo principal del proyecto es implementar un sistema de respaldo automatizado que funcione en tiempo real. Este objetivo implica varios sub-objetivos:

**Monitorización en tiempo real:** El sistema debe ser capaz de detectar cualquier modificación en el directorio protegido de forma inmediata.

**Automatización del proceso de respaldo:** La creación y transferencia de los archivos de respaldo debe realizarse de forma automática, sin intervención manual.

**Almacenamiento remoto:** Los respaldos se deben almacenar en una ubicación remota, en este caso, una máquina virtual en Azure, para garantizar la protección de los datos ante desastres locales.

#### 3. Infraestructura utilizada

La infraestructura del sistema de respaldo se compone de dos máquinas virtuales:

**Máquina virtual AlmaLinux (origen):** Esta máquina actúa como el servidor que contiene los datos a proteger.

**Máquina virtual Linux en Azure (destino):** Esta máquina virtual en la nube de Azure actúa como el repositorio remoto para los archivos de respaldo. La elección de Azure proporciona escalabilidad, alta disponibilidad y redundancia geográfica.

**Comunicación mediante SSH:** La comunicación entre las dos máquinas virtuales se realiza mediante SSH, utilizando un esquema de autenticación sin contraseña. Esto permite la transferencia segura y automatizada de los archivos de respaldo.

**Red pública:** Ambas máquinas virtuales están conectadas a una red pública, con el acceso controlado mediante reglas de firewall que restringen el tráfico al puerto 22 (SSH). Esto garantiza que solo las conexiones SSH autorizadas puedan acceder a las máquinas virtuales.

## 4. Herramientas instaladas y configuraciones previas

## En la máquina de origen (AlmaLinux):

**inotify-tools**: Este conjunto de herramientas proporciona una interfaz para monitorizar los eventos del sistema de archivos. El programa inotifywait se utiliza para detectar los cambios en el directorio protegido.

tar: Esta herramienta se utiliza para crear archivos comprimidos (archivos tar) de los datos que se van a respaldar. La compresión reduce el tamaño de los archivos de respaldo, lo que ahorra espacio de almacenamiento y ancho de banda durante la transferencia.

**scp**: Esta herramienta se utiliza para copiar archivos de forma segura a través de una conexión SSH. Se utiliza para transferir los archivos de respaldo desde el servidor AlmaLinux a la máquina virtual en Azure.

**openssh-clients**: Este paquete proporciona los clientes SSH necesarios para establecer la conexión segura entre las máquinas virtuales.

#### En la máquina destino (Azure):

**SSH habilitado**: El servidor SSH debe estar habilitado en la máquina virtual de Azure para permitir las conexiones entrantes desde el servidor AlmaLinux.

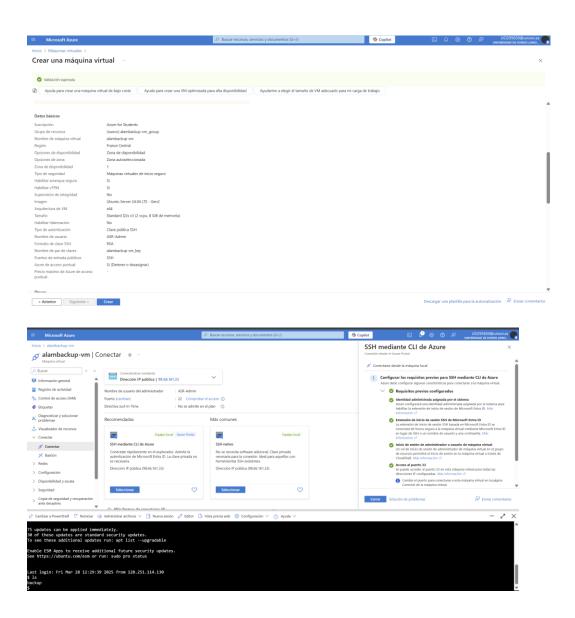
**Carpeta de backups creada** : Se debe crear una carpeta específica en la máquina virtual de Azure para almacenar los archivos de respaldo. Esto ayuda a organizar los respaldos y facilita su gestión.

## 5. Creación de la Máquina Virtual en Azure

La máquina virtual en Azure se creó a través del portal de Azure, dado que ofrece almacenamiento remoto seguro y acceso controlado a través de credenciales y reglas de firewall. Se eligió un sistema operativo Linux para garantizar compatibilidad con la máquina de origen y simplificar la gestión de archivos.

También lo elegimos debido a los créditos proporcionados de manera gratuita para crear los recursos.

Configuramos los ajustes iniciales, como el nombre de la máquina virtual, el grupo de recursos, la región y el sistema operativo.



## 6. Conexión de las máquinas mediante ssh

Para garantizar una conexión segura y automatizada entre la máquina de origen (AlmaLinux) y la máquina en Azure, se utiliza autenticación basada en claves SSH. Este método elimina la necesidad de ingresar una contraseña manualmente cada vez que se realiza la transferencia del backup.

#### Pasos para la Configuración de SSH:

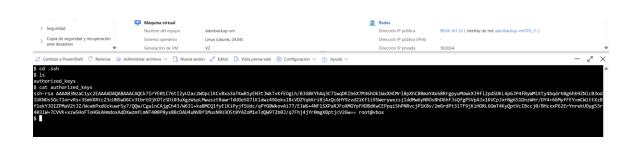
**1. Generación de Claves SSH en la Máquina de Origen.** Se genera un par de claves SSH con el comando:

ssh-keygen -t rsa -b 4096

- t rsa indica que se usará el algoritmo RSA.
- b 4096 define una clave de 4096 bits, aumentando la seguridad.

Esto crea un archivo id\_rsa (clave privada) y id\_rsa.pub (clave pública) en ~/.ssh/.

**2. Copia de la Clave Pública en la Máquina de Azure.** La clave pública debe agregarse al archivo authorized\_keys de la máquina de destino para permitir autenticación sin contraseña.



## 7. Automatización del Backup en Tiempo Real

Este script es el núcleo del sistema de respaldo y se encarga de monitorizar los cambios en el directorio protegido y de crear y transferir los backups a la máquina virtual en Azure.

```
UO295650 - AlmaLinux 9 GUI base [Corriendo] - Oracle VirtualBox
  GNU nano 5.6.1
                                                                                  usr/local/bin/watch and backup.sh
LOG="/var/log/watch_backup.log"
exec >> "$LOG" 2>&1
echo "=== Script iniciado $(date) ==="
DIR_VIGILADO="/root/DirectorioImportante"
USUARIO_REMOTO="asrtrabajo"
 IP_REMOTA="98.66.161.33"
RUTA_REMOTA="/home/asrtrabajo/backups"
# Crear carpeta remota si no existe
ssh "$USUARIO_REMOTO@$IP_REMOTA" "mkdir -p $RUTA_REMOTA"
  cho "Vigilando cambios en $DIR_VIGILADO..."
inotifywait -m -r -e modify,create,delete,move "$DIR_VIGILADO" --format '%w%f %e' |
 hile read archivo evento;
          "Cambio detectado: $archivo ($evento)"
    sleep 5
     FECHA=$(date +%d-%m-%Y_%H-%M)
     ARCHIVO="/tmp/backup_${FECHA}.tar.gz"
tar -czf "$ARCHIVO" -C "$(dirname "$D
     scp "$ARCHIVO" "$USUARIO_REMOTO@$IP_REMOTA:$RUTA_REMOTA"
     echo "Backup subido correctamente: $ARCHIVO"
```

#### Explicación del script:

El script utiliza la herramienta inotifywait para detectar los cambios que se producen en el directorio que se desea proteger. inotifywait es una herramienta que forma parte del paquete inotify-tools y permite monitorizar eventos del sistema de archivos, como la creación, modificación, eliminación o movimiento de archivos y directorios.

#### **Configuración inicial:**

- Se define la ruta del directorio que se va a monitorizar (DIR\_VIGILADO).
- Se definen las credenciales y la ruta de la máquina virtual en Azure (USUARIO\_REMOTO, IP\_REMOTA, RUTA\_REMOTA).
- Se crea el directorio de destino en la máquina virtual de Azure, si no existe.

#### Monitorización de cambios:

- Se utiliza inotifywait para monitorizar los eventos de modificación, creación, eliminación y movimiento en el directorio especificado.
- El script entra en un bucle infinito (while true) que se ejecuta continuamente, esperando a que se produzca un evento.

#### Creación del backup:

- Cuando inotifywait detecta un cambio, el script crea un archivo tar comprimido (.tar.gz) del directorio.
- El nombre del archivo de backup incluye la fecha y la hora del evento, lo que permite tener un registro de cuándo se realizó cada copia de seguridad.

#### Transferencia del backup:

- El script utiliza scp para transferir el archivo tar comprimido a la máquina virtual en Azure.
- scp es una herramienta que permite copiar archivos de forma segura a través de una conexión SSH.

#### Eliminación del archivo temporal:

• Después de transferir el archivo de backup, el script elimina el archivo temporal .tar.gz del sistema local.

#### Registro de eventos:

• El script registra todos los eventos y acciones en un archivo de registro (/var/log/watch\_backup.log), lo que facilita la resolución de problemas y el seguimiento de la actividad del sistema de respaldo.

#### Guardar solo la última copia de seguridad:

Para guardar solo la última copia de seguridad en lugar de guardar todas, se puede modificar el script para que, antes de crear un nuevo backup, elimine el anterior. Esto se puede hacer añadiendo una línea al script que ejecute el comando ssh para eliminar el archivo de backup anterior en la máquina virtual de Azure.

#### Conversión del script en un servicio

Para que el script se ejecute automáticamente en segundo plano al iniciar el sistema, es necesario convertirlo en un servicio

Para ello es necesario la creacion de un archivo de unidad de systemd. Este archivo contiene la configuración del servicio, incluyendo el comando que se va a ejecutar, el usuario bajo el que se va a ejecutar, y las dependencias del servicio.

```
U0295650-AlmaLinux 9 GUI base [Corriendo] - Oracle VirtualBox

GNU nano 5.6.1

Petc/systemd/system/watch_backup.service

[Unit]

Description=Backup automático con detección de cambios

After=network.target

[Service]

ExecStart=/usr/local/bin/watch_and_backup.sh

Restart=always

User=root

WorkingDirectory=/root

StandardOutput=append:/var/log/watch_backup.log

StandardError=append:/var/log/watch_backup.log

[Install]

WantedBy=multi-user.target
```

Después de crear el archivo de unidad, se debe recargar la configuración de systemd para que reconozca el nuevo servicio.

Tras habilitar el servicio, podemos ver que el comando "sudo systemctl status watch\_backup.service" devuelve:

Por lo que funciona correctamente.

## 8. Ejemplo de uso

Creamos dos directorios distintos, uno desde la terminal y otro desde la interfaz gráfica. Esto lo hicimos para asegurarnos de que el sistema de respaldo detecta los cambios hechos tanto por la terminal como por la interfaz gráfica. Dentro del directorio "dirprueba" creamos un archivo de texto ("pruebaConsola.txt").



```
UO295650 - AlmaLinux 9 GUI base [Corriendo] - Oracle VirtualBox

[root@vbox ~] # nano /var/log/watch_backup.log

[root@vbox ~] # mkdir DirectorioImportante/dirprueba

[root@vbox ~] # echo "Fichero prueba" > DirectorioImportante/dirprueba/pruebaConsola.txt

[root@vbox ~] # ■
```

Log tras la creación de los ficheros y del archivo de texto:

```
👺 UO295650 - AlmaLinux 9 GUI base [Corriendo] - Oracle VirtualBox
[root@vbox ~]# nano /var/log/watch_backup.log
[root@vbox ~]# mkdir DirectorioImportante/dirprueba
[root@vbox ~]# echo "Fichero prueba" > DirectorioImportante/dirprueba/pruebaConsola.t
[root@vbox ~]# cat /var/log/watch_backup.log
=== Script iniciado vie 28 mar 2025 20:14:21 CET ===
ssh: connect to host 98.66.161.33 port 22: Network is unreachable
Vigilando cambios en /root/DirectorioImportante...
Setting up watches. Beware: since -r was given, this may take a while!
Watches established.
Cambio detectado: /root/DirectorioImportante/dirprueba (CREATE,ISDIR)
Backup subido correctamente: /tmp/backup_28-03-2025_20-29.tar.gz
Cambio detectado: /root/DirectorioImportante/dirprueba/pruebaConsola.txt (CREATE)
Backup subido correctamente: /tmp/backup_28-03-2025_20-29.tar.gz
Cambio detectado: /root/DirectorioImportante/dirprueba/pruebaConsola.txt (MODIFY)
Backup subido correctamente: /tmp/backup_28-03-2025_20-30.tar.gz
Cambio detectado: /root/DirectorioImportante/PruebaInterfaz (CREATE,ISDIR)
Backup subido correctamente: /tmp/backup_28-03-2025_20-31.tar.gz
[root@vbox ~]#
```

#### Conexión a la máquina virtual de Azure:

Realizamos una conexión SSH con la máquina virtual en Azure, usando su dirección IP pública. De esta manera comprobamos que la máquina azure está funcionando y que la conexión está bien configurada.

```
🌉 UO295650 - AlmaLinux 9 GUI base [Corriendo] - Oracle VirtualBox
[root@vbox ~]# ssh asrtrabajo@98.66.161.33
Welcome to Ubuntu 24.04.2 LTS (GNU/Linux 6.8.0-1021-azure x86_64)
 * Documentation: https://help.ubuntu.com
 * Management:
                  https://landscape.canonical.com
                  https://ubuntu.com/pro
 * Support:
 System information as of Fri Mar 28 23:23:12 UTC 2025
  System load: 0.0
                                  Processes:
                                                          134
 Usage of /: 6.6% of 28.02GB
                                  Users logged in:
  Memory usage: 4%
                                  IPv4 address for eth0: 10.0.0.4
  Swap usage: 0%
 \star Strictly confined Kubernetes makes edge and IoT secure. Learn how MicroK8s
  just raised the bar for easy, resilient and secure K8s cluster deployment.
  https://ubuntu.com/engage/secure-kubernetes-at-the-edge
Expanded Security Maintenance for Applications is not enabled.
75 updates can be applied immediately.
30 of these updates are standard security updates.
To see these additional updates run: apt list --upgradable
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
Last login: Fri Mar 28 18:57:51 2025 from 89.141.187.79
asrtrabajo@alambackup-vm:-$
```

#### Verificación de los backups en Azure:

En la siguiente captura, se ve el contenido del directorio donde se guardan los backups en la máquina virtual de Azure. Se ven varios archivos con la extensión ".tar.gz". Estos son los archivos comprimidos que contienen las copias de seguridad.

```
Sackups asrtrabajo@alambackup-vm:-$ cd backups/
asrtrabajo@alambackup-vm:-$ cd backups/
asrtrabajo@alambackup-vm:-$ backups$ ls
backup_28-03-2025_19-55.tar.gz backup_28-03-2025_20-13.tar.gz backup_28-03-2025_20-20.tar.gz
backup_28-03-2025_20-09.tar.gz backup_28-03-2025_20-15.tar.gz backup_28-03-2025_20-21.tar.gz
backup_28-03-2025_20-10.tar.gz backup_28-03-2025_20-16.tar.gz backup_28-03-2025_20-22.tar.gz
backup_28-03-2025_20-12.tar.gz backup_28-03-2025_20-18.tar.gz backup_28-03-2025_20-27.tar.gz
backup_28-03-2025_20-12.tar.gz backup_28-03-2025_20-18.tar.gz backup_28-03-2025_20-27.tar.gz
backup_28-03-2025_20-12.tar.gz backup_28-03-2025_20-18.tar.gz backup_28-03-2025_20-27.tar.gz
```

Por último, descomprimimos uno de los archivos de backup en la máquina virtual de Azure. Se ve que contiene el "DirectorioImportante", el subdirectorio "dirprueba" y el archivo "pruebaConsola.txt".

## 9. Posibles mejoras

#### Uso de almacenamiento cifrado en el servidor de destino

Para garantizar la seguridad de los backups, se podría y debería utilizar almacenamiento cifrado en el servidor de destino.

Para ello se podría utilizar encfs.

sudo dnf install encfs -y encfs /ruta/encriptada /ruta/montaje

Esto asegura que los archivos solo puedan leerse tras montar el directorio con la clave adecuada.

#### Verificación de integridad antes de eliminar el backup local

Actualmente, el script elimina el archivo local después de la transferencia al servidor remoto sin verificar que la copia remota se haya realizado correctamente. Para evitar pérdidas de datos, se podría calcular un hash antes y después de la transferencia. Con este procedimiento, el archivo sólo se eliminará localmente si la verificación de integridad es exitosa.

#### 10. Conclusiones

En este proyecto, se ha desarrollado un sistema de respaldo automatizado y en tiempo real para AlmaLinux, con el destino de los respaldos siendo una máquina virtual en Azure. El sistema utiliza la detección de eventos del sistema de archivos para disparar el proceso de respaldo, lo que proporciona una solución más eficiente y proactiva que los métodos tradicionales basados en cron. Los resultados de las pruebas demuestran que el sistema funciona correctamente, creando y transfiriendo backups de forma automática cuando se producen cambios en el directorio protegido. El sistema es robusto, eficiente y fácil de usar, y proporciona una solución valiosa para la protección de datos en entornos AlmaLinux.