TASK 6: File Compression & Backup

Tools: Primary: tar, gzip

Alternatives: zip, rsync

✅ TASK 6: File Compression & Backup (Linux)

🎯 Objective

Understand how to **compress files**, **take backups**, **automate them**, and **schedule backups** using Linux tools.

1️⃣ Create Sample Directories & Files

mkdir -p ~/backup_demo/docs mkdir -p ~/backup_demo/images echo "This is a document file" > ~/backup_demo/docs/file1.txt echo "Linux Backup Practice" > ~/backup_demo/docs/file2.txt touch ~/backup_demo/images/img1.jpg touch ~/backup_demo/images/img2.png

📁 Structure created:

backup_demo/

├── docs/

│ ├── file1.txt

│ └── file2.txt

└── images/

├── img1.jpg

└── img2.png

2️⃣ Compress Files Using tar and gzip

🔹 Create a tar.gz backup

tar -czvf backup_demo.tar.gz ~/backup_demo

Explanation:

- c → create
- z → gzip compression
- v → verbose
- f → file name

✅ Output file: backup_demo.tar.gz

3️⃣ Extract Compressed Files

tar -xzvf backup_demo.tar.gz

Explanation:

- x → extract
- z → gzip
- v → verbose
- f → file

4️⃣ Automate Backup Using a Script

🔹 **Create backup script**

nano backup.sh

🔹 **Script content:**

#!/bin/bash SOURCE=~/backup_demo DEST=~/backups DATE=$(date +%F) mkdir -p $DEST tar -czf $DEST/backup_$DATE.tar.gz $SOURCE echo "Backup completed on $DATE"

🔹 **Make script executable**

chmod +x backup.sh

🔹 **Run manually**

./backup.sh

5️⃣ Schedule Backup Using cron

🔹 **Open cron editor**

crontab -e

🔹 **Add job (Daily at 10 PM)**

0 22 * * * /home/username/backup.sh

🕒 Cron format:

minute hour day month weekday command

6️⃣ Verify Backups

ls ~/backups

tar -tzf backup_2026-01-25.tar.gz

✓ Ensures files are correctly backed up.

7️⃣ Alternative Tools (Optional Knowledge)

🔹 **zip**

zip -r backup.zip backup_demo

🔹 **rsync**

rsync -av ~/backup_demo ~/backup_sync

❓ What is cron?

cron is a **Linux scheduler** that runs commands automatically at fixed times/dates.

❓ Why do backups matter?

- Protect against **data loss**
- Recover from **system failure**
- Help during **accidental deletion**
- Essential for **disaster recovery**

❓ What is rsync?

rsync is a **fast file synchronization tool** that copies **only changed files**, saving time and space.

📦 Deliverables

✓ Backup archive: backup_YYYY-MM-DD.tar.gz
✓ Cron job configuration
✓ Automated backup script

🎓 Final Outcome

✅ Intern understands:

- File compression
- Backup strategies

- Automation using scripts
- Scheduling with cron
- Real-world Linux admin skills