DATABASE SYSTEMS

CS – 355/CE – 373

Instructor: Maria N. Samad

November 20th, 2024

TYPES OF BACKUP

- An accidental loss of data assets can have a devastating impact on a business.
- Running proper backups at the right intervals can prevent loss of revenue and customers.
- Types of Backups for databases are:
  - ***Full Backup***
  - ***Incremental Backup***
  - ***Differential Backup***
- A ***Log Backup*** is also conducted

FULL BACKUP

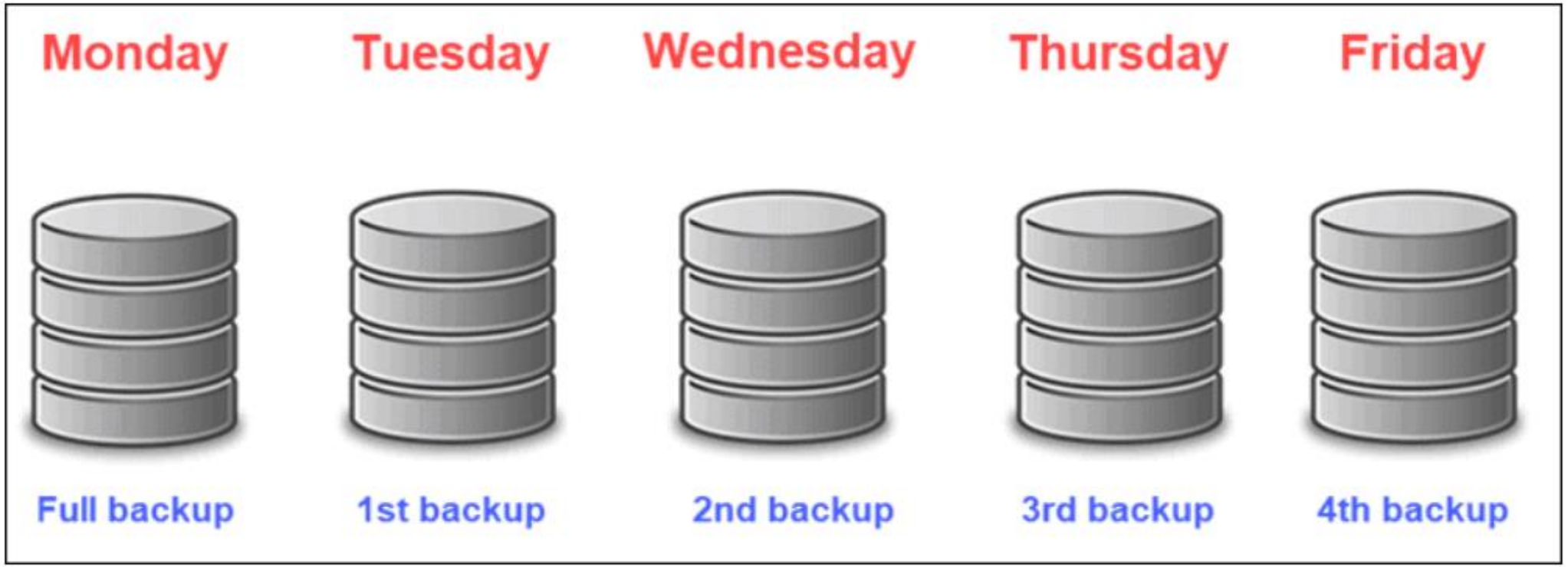
- A ***full backup*** involves copying the entire data set of the system into a separate partition or onto an external disk (Secondary Storage)
- It requires a lot of free disk space
- It takes up a lot of time
- It may not be practical on a daily basis, especially if the database is large
- It is more practical to schedule full backups on a weekly or biweekly basis, depending on the size of the database (daily backup is possible, provided the size of the database is small)

ADVANTAGES AND DISADVANTAGES

- **Advantages:**
  - Provides the best protection in terms of data recovery
  - Fast recovery of data in a single backup set

- **Disadvantages:**
  - Backup is time-consuming
  - Requires more storage space
  - May use a lot of bandwidth

A TYPICAL FULL BACKUP SCHEDULE



INCREMENTAL BACKUP

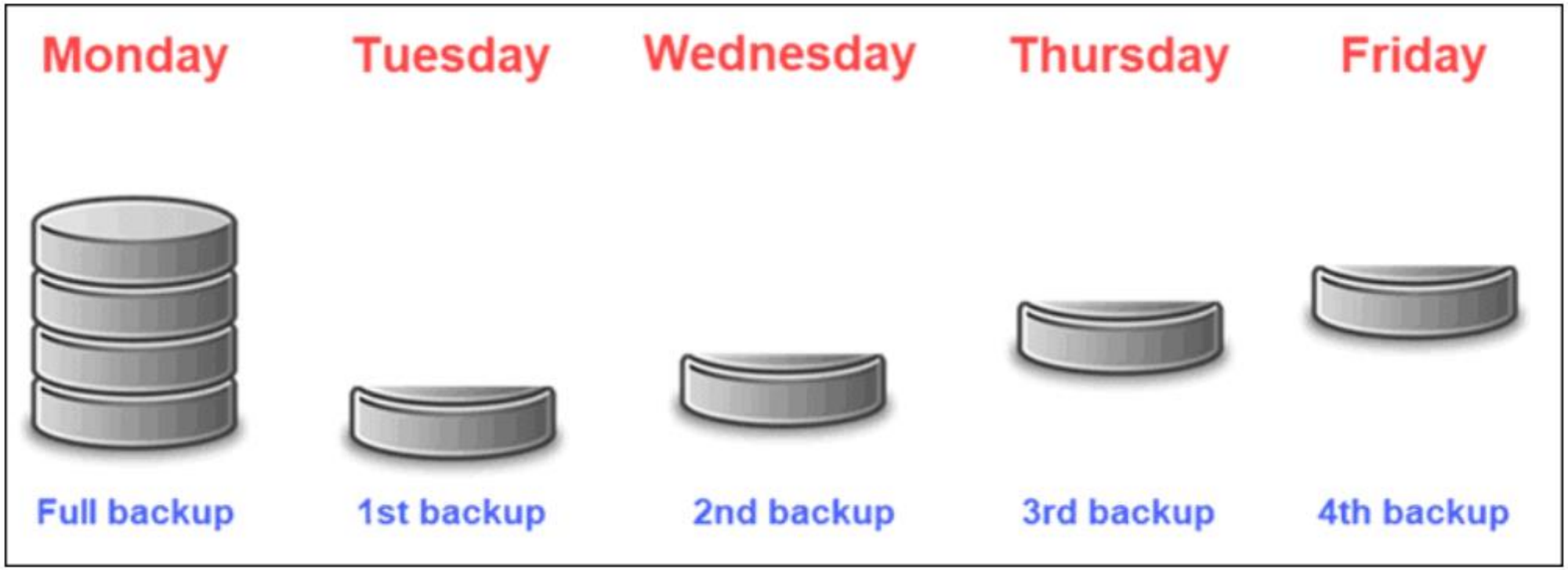
- An ***Incremental Backup*** stores all files changed since the last backup
- This last (most recent) backup may be a ***full***, ***differential*** or an ***incremental backup*** itself
- An incremental backup can be considered as a resource-friendly alternative to full backup
- It exclusively saves data that has been modified or added to the existing data volume

ADVANTAGES AND DISADVANTAGES

- **Advantages**:
  - Smaller backups that take up less storage space
  - Faster to backup
  - Uses less bandwidth

- **Disadvantages:**
  - Time-consuming to recover
  - Risk of failed recovery if there is damage to a segment in the backup chain

A TYPICAL INCREMENTAL BACKUP SCHEDULE



DIFFERENTIAL BACKUP

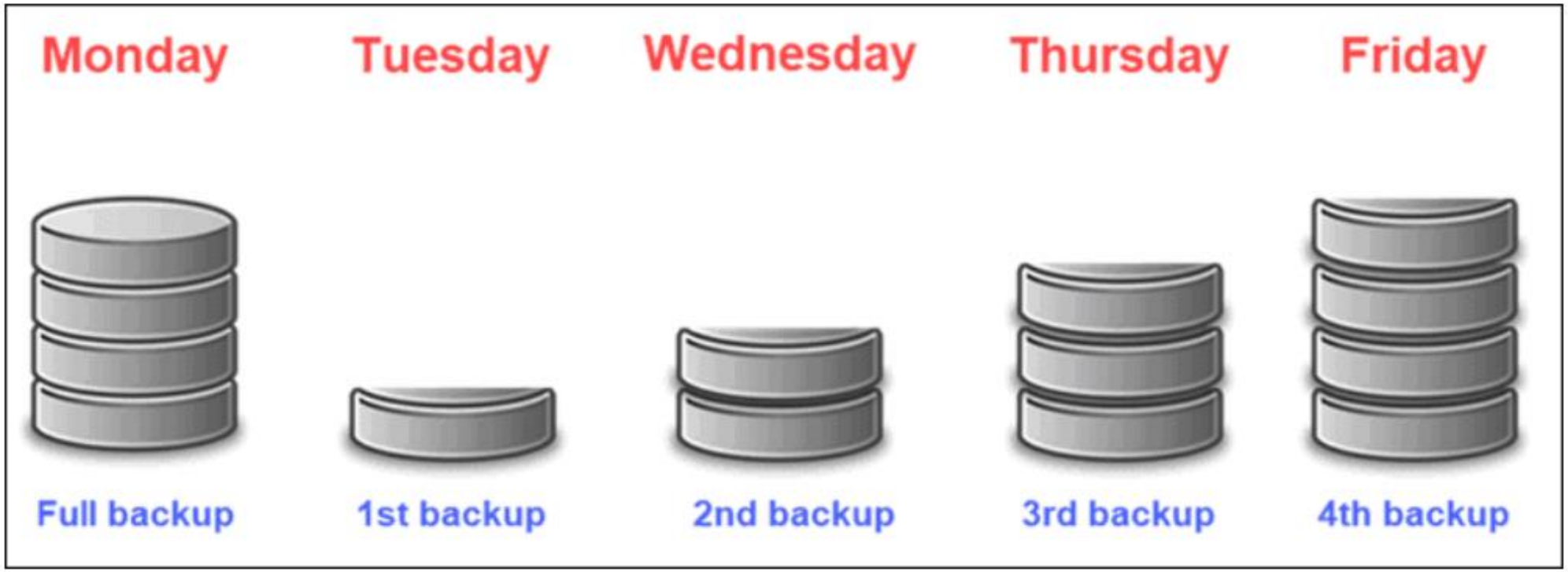
- A ***differential backup*** updates all data that has changed since the last **full backup**
- It relies on the most recent last full backup
- Recovery time is much faster, although somewhat slower than a full backup.
- This provides good data protection and a valid disaster recovery solution.

ADVANTAGES AND DISADVANTAGES

- **Advantages:**
  - Faster to back up compared to full backups
  - Takes up less space than full backups

- **Disadvantages:**
  - Takes up more space than incremental backups
  - Slower to back up compared to incremental backups

A TYPICAL DIFFERENTIAL BACKUP SCHEDULE



SUMMARY

The following table → provides a summary of backup types and their advantages/ disadvantages with respect to various aspects.

	FULL	INCREMENTAL	DIFFERENTIAL
Description:	copies the entire data set	full backup + changes since the previous backup	full backup + changes since the full backup
Backup time:	time-consuming	fast to back up	faster than a full backup but slower than an incremental
Recovery time:	fast recovery	slow recovery	faster than incremental but slower than full
Storage space:	requires a lot of storage space	requires less storage space	requires less storage space than a full backup, but more than an incremental
Bandwidth:	uses a lot of bandwidth	uses less bandwidth	uses less bandwidth than a full backup, but more than an incremental backup

EXAMPLE

- A company is planning to set up a backup schedule for its databases and considering to choose between the following alternative schedules suggested by its team members. They have asked you to assist in the decision making.

Alternate Schedule	Details
Schedule A	Full backups on Sundays at 9:00 am Differential backups Mondays to Saturdays 9:00 am
Schedule B	Full backups on Sundays at 9:00 am Incremental backups Mondays to Saturdays 9:00 am
Schedule C	Full backups on Sundays at 9:00 am Differential backups Mondays to Saturdays 9:00 am Incremental backups Mondays to Saturdays at 12 noon, 6:00 pm and 12 midnight

EXAMPLE

- In case a disaster occurs at 2:55 pm on a Tuesday resulting in loss of hardware (including storage), answer the following:
  1. How much data (in terms of time) will be lost in each of the backup schedules?
  2. Which schedule would be faster to restore? Why?
  3. What will be that schedule's restore strategy in case of the failure on the given time?
  4. For the other two remaining schedules, what will be their restoring strategies?

EXAMPLE (SOLUTION)

- In case a disaster occurs at 2:55 pm on a Tuesday resulting in loss of hardware (including storage), answer the following:
  1. **How much data (in terms of time) will be lost in each of the backup schedules?**
     - ***Schedule A*****:** 5 hours and 55 minutes
     - ***Schedule B*****:** 5 hours and 55 minutes
     - ***Schedule C*****:** 2 hours and 55 minutes

EXAMPLE (SOLUTION)

- In case a disaster occurs at 2:55 pm on a Tuesday resulting in loss of hardware (including storage), answer the following:
  2. **Which schedule would be faster to restore? Why?**
     - As ***schedule C*** has the least amount of data loss, it will be quicker to recover the data from that schedule.

EXAMPLE (SOLUTION)

- In case a disaster occurs at 2:55 pm on a Tuesday resulting in loss of hardware (including storage), answer the following:
  3. **What will be that schedule's restore strategy in case of the failure on the given time?**
     - Restore Incremental backup taken on Tuesday at 12:00pm
     - Restore Differential backup taken on Tuesday morning at 9:00am
     - Finally restore from Full backup taken on Sunday
     - We do NOT need to restore Monday's differential backup in between because Tuesday's differential backup will cover all the changes made since its Full Backup on Sunday

EXAMPLE (SOLUTION)

- In case a disaster occurs at 2:55 pm on a Tuesday resulting in loss of hardware (including storage), answer the following:
  4. For the other two remaining schedules, what will be their restoring strategies?
     - ***Schedule A:***
       - Restore Differential backup taken on Tuesday morning at 9:00am
       - Finally restore from Full backup taken on Sunday
       - Similarly in this case, we do NOT need to restore Monday's differential backup in between because Tuesday's differential backup will cover all the changes made since its Full Backup on Sunday

EXAMPLE (SOLUTION)

- In case a disaster occurs at 2:55 pm on a Tuesday resulting in loss of hardware (including storage), answer the following:
  4. For the other two remaining schedules, what will be their restoring strategies?
     - ***Schedule B:***
       - Restore Incremental backup taken on Tuesday morning at 9:00am
       - Restore Incremental backup taken on Monday morning at 9:00am
       - Finally restore from Full backup taken on Sunday
       - Here, we will need to restore backups from both Monday and Tuesday, as they each contain part of the updates and we need them both in order to restore the complete information

DATABASE BACKUP AND RECOVERY

- Activity Sheet

DATABASE BACKUP AND RECOVERY

- Activity Sheet Solution:
  - Database Backup and Recovery Schedules Solution