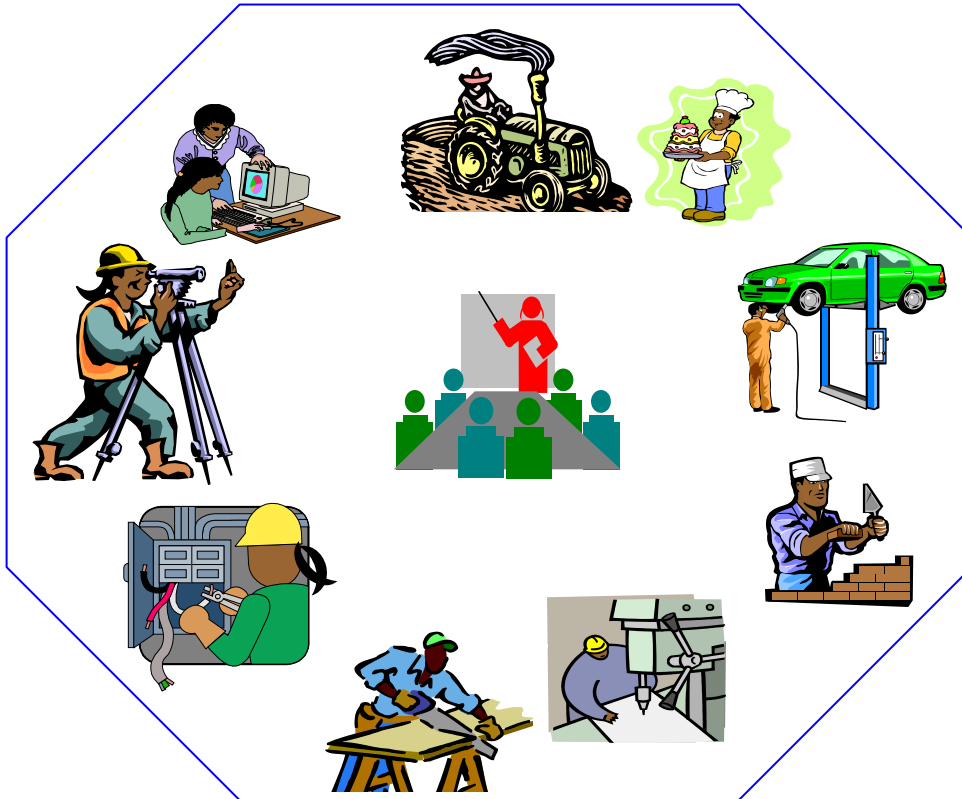




Database Administration Level III

Based on August, 2011, Version 3 Occupational
Standards (OS) and Curriculum



**Module Title: Completing Database Backup and
Recovery**

LG code: EIS DBA3 M011 1220 LO (1- 3) (LG40-42)

TTLM Code: EIS DBA3 TTLM011 1220 V1

December, 2020

Bishoftu, Ethiopia

Page 1 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020



Contents

LO #1- Review database architecture	1
Instruction sheet	1
Self-Check -1.....	6
Information sheet 2: Determining appropriate methods for back-up and recovery	7
Self-Check -2.....	13
Information sheet 3: identifying and examining Risks and failure scenarios	14
Self-Check -3.....	18
LO #2- Determine backup methods appropriate to database requirements	19
Instruction sheet	19
Information sheet 1: Evaluating range of back-up and restoration.....	20
Self-Check -1.....	23
Information sheet 2: Completing full off-line back-ups	24
Self-Check -2.....	29
Information sheet 3: Completing On-line file back-ups	30
Self-Check -3.....	33
Information sheet 4: Employing Disk mirroring and redundant array of inexpensive disks (RAID) hard disk configurations to keep copies of files.....	34
Self-Check -4.....	38
Information sheet 5: Arranging Off-site copies of back-up files.....	39
Self-Check -5.....	43
Operation sheet 1	44
LAP Test 1.....	47
LO #3- Establish recovery points and disaster recovery procedures	48
Instruction sheet	48
Information sheet 1: Determining database recovery points based on the back-up arrangements	49
Self-Check -1.....	53
Information sheet 2: Testing restore process.....	54
Self-Check -2.....	55
Information sheet 3: Completing restoration of the database to the point of failure without loss of committed transactions	56
Self-Check -3.....	60
Operation sheet1:	61
LAP Test 1:.....	63
References	64
Answer sheet.....	66



L #40

LO #1- Review database architecture

Instruction sheet

This learning guide is developed to provide you the necessary information regarding the following content coverage and topics:

- identifying architecture of a database file system
- determining appropriate methods for back-up and recovery
- identifying and examining Risks and failure scenarios

This guide will also assist you to attain the learning outcomes stated in the cover page. Specifically, upon completion of this learning guide, you will be able to:

- identify architecture of a database file system
- determine appropriate methods for back-up and recovery
- identify and examine Risks and failure scenarios

Learning Instructions:

Read the specific objectives of this Learning Guide.

1. Follow the instructions described below.
2. Read the information written in the “Information Sheets”. Try to understand what are being discussed. Ask your trainer for assistance if you have hard time understanding them.
3. Accomplish the “Self-checks” which are placed following all information sheets.
4. Ask from your trainer the key to correction (key answers) or you can request your trainer to correct your work. (You are to get the key answer only after you finished answering the Self-checks).
5. If you earned a satisfactory evaluation proceed to the next learning guide.



Information sheet 1: Identifying architecture of a database file system

1.1. Identifying architecture of a database file system

Database architecture focuses on database design and construction for large enterprise database systems that manage massive amounts of information for organizations. Database architecture includes setting the standards for the security and programming aspects of these databases, as well as figuring out how these databases will operate and function within existing structures.

The design of a DBMS depends on its architecture. Selecting the correct Database Architecture helps in quick and secure access to data. It can be centralized or decentralized or hierarchical. The architecture of a DBMS can be seen as either single tier or multi-tier. The tiers are classified as follows:

1.1.1. Single tier architecture

The simplest of Database Architecture are **1 tier** where the Client, Server, and Database all reside on the same machine. In other word, it keeps all of the elements of an application, including the interface, Middleware and back-end data, in one place. Developers see these types of systems as the simplest and most direct way.

- The database is directly available to the user. It means the user can directly sit on the DBMS and uses it.
- Any changes done here will directly be done on the database itself. It doesn't provide a handy tool for end users.
- The 1-Tier architecture is used for development of the local application, where programmers can directly communicate with the database for the quick response.

For example; when you install a DB in your system and access it to practise SQL queries it is tier-one architecture. But such architecture is rarely used in production



Figure 1.1: Single tier architecture

1.1.2. Two-tier Architecture

- The two-tier is based on Client Server architecture.
- It is like client server application.
- The direct communication takes place between client and server. There is no intermediate between client and server.
- Applications on the client end can directly communicate with the database at the server side. For this interaction, API's like: ODBC, JDBC are used.
- The user interfaces and application programs are run on the client-side.
- The server side is responsible to provide the functionalities like: query processing and transaction management.
- To communicate with the DBMS, client-side application establishes a connection with the server side
- 2 tier architecture provides added security to the DBMS as it is not exposed to the end user directly.

A two-tier architecture is a database architecture where

1. Presentation layer runs on a client (PC, Mobile, Tablet, etc)
2. Data is stored on a Server.

Example of Two-tier Architecture is a Contact Management System created using MS-Access.

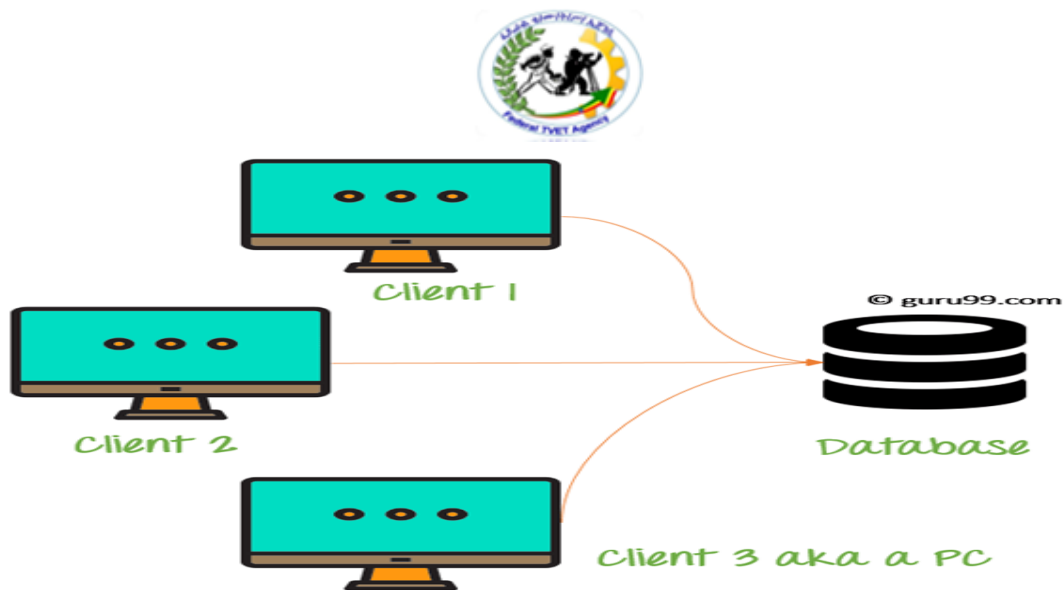


Figure 1.2. Two Tier architecture

1.1.3. Three-tier Architecture

It is an extension of the 2-tier architecture. A 3-tier architecture separates its tiers from each other based on the complexity of the users and how they use the data present in the database. The three tier architecture is the most popular DBMS architecture.

This architecture has different usages with different applications. It can be used in web applications and distributed applications. 3-tier architecture has following layers;

- **Database server (Data) Tier** – At this tier, the database resides along with its query processing languages. We also have the relations that define the data and their constraints at this level.
- **Application (Middle) Tier** – also called business logic layer and it processes functional logic, constraint, and rules before passing data to the user or down to the DBMS. This DBMS architecture contains an Application layer between the user and the DBMS, which is responsible for communicating the user's request to the DBMS system and send the response from the DBMS to the user.

For a user, this application tier presents an abstracted view of the database. End-users are unaware of any existence of the database beyond the application. At the other end, the database tier is not aware of any other user beyond the application



tier. Hence, the application layer sits in the middle and acts as a mediator between the end-user and the database.

- **User (Presentation) Tier** – End-users operate on this tier and they know nothing about any existence of the database beyond this layer. At this layer, multiple views of the database can be provided by the application. All views are generated by applications that reside in the application tier. Example your PC, Tablet, Mobile, etc.)

The goal of Three-tier architecture is:

- To separate the user applications and physical database
- Proposed to support DBMS characteristics
- Program-data independence
- Support of multiple views of the data

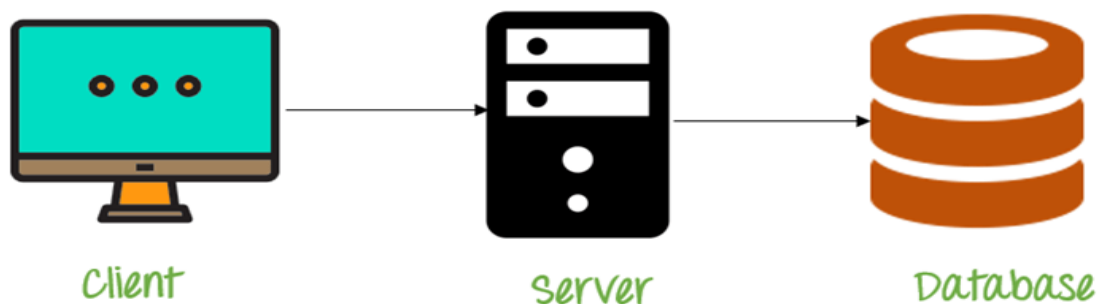


Figure 1.3. Three-tier Architecture Diagram

1.1.4. N-tier architecture

N-tier architecture would involve dividing an application into three different tiers. These would be the

- Logic tier,
- The presentation tier, and
- The data tier.



Self-Check -1 : Written Test

Directions: Answer all the questions listed below. Use the Answer sheet provided by your instructor to check whether your answer is correct or not

I. Choose the best answer (2 pts each)

1. At which architecture the database is directly available to the user.
 - A. Tier-one
 - B. Tier-two
 - C. Tier-three
 - D. N-tier
2. Programmers can directly communicate with the database for the quick response.
 - A. N-tier
 - B. Tier-two
 - C. Tier-one
 - D. Tier-three
3. Applications on the client end can directly communicate with the database at the server side through API's like ODBC and JDBC.
 - A. N-tier
 - B. Tier-three
 - C. Tier-one
 - D. Tier-two
4. Selecting the correct Database architecture helps promote;
 - A. quick data access
 - B. Ensure security
 - C. Facilitate data management
 - D. All

Note: Satisfactory rating – 4 and 10 points points

Unsatisfactory - below 4

Answer Sheet

Name: _____

Date: _____

Score = _____

Page 6 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020



Information sheet 2: Determining appropriate methods for back-up and recovery

Introduction

Backup is the process of creating a copy of data to protect against accidental or malicious deletion, corruption, hardware failure, ransom ware attacks, and other types of data loss. Data backups can be created locally, offsite, or both.

Restore is the process of retrieving data from a backup. This might mean copying data from backup media to an existing device or to a new device. It also could mean copying data from the cloud to a local device or from one cloud to another.

Recovery refers to the process of restoring data and operations (e.g., returning a server to normal working order following hardware failure).

Restore and recovery times can vary widely depending on the backup format and data recovery methods you choose. Additionally, restore needs also vary (e.g., restoring a single file vs. an entire server). Finally, critical data may live on workstations, local servers, and in the cloud. These are important considerations when selecting a backup and recovery solution. The most common type of database backups are:

- Logical backup - backup of data is stored in a human-readable format like SQL
- Physical backup - backup contains binary data

1.1. Determining appropriate methods for back-up and recovery

1.1.1. Types of Backup

There are different types of backup, and each backup process works differently.

Table 2.1. Comparison of backup type
A comparison of different types of backup

backup	Full	Mirror	Incremental	Differential
Backup 1	All data	All data Selected	-	-
Backup 2	All data	All data Selected	Changes from backup 1	Changes from backup 1
Backup 3	All data	All data Selected	Changes from backup 2	Changes from backup 1
Backup 4	All data	All data Selected	Changes from backup 3	Changes from backup 1



1. Full backups

The most basic and complete type of backup operation is a full backup. As the name implies, this type of backup makes a copy of all data to a storage device, such as a disk or tape. The primary advantage to performing a full backup during every operation is that a complete copy of all data is available with a single set of media. This results in a minimal time to restore data, a metric known as a recovery time objective. However, the disadvantages are that it takes longer to perform a full backup than other types (sometimes by a factor of 10 or more), and it requires more storage space.

Thus, full backups are typically run only periodically. Data centers that have a small amount of data (or critical applications) may choose to run a full backup daily, or even more often in some cases. Typically, backup operations employ a full backup in combination with either incremental or differential backups.

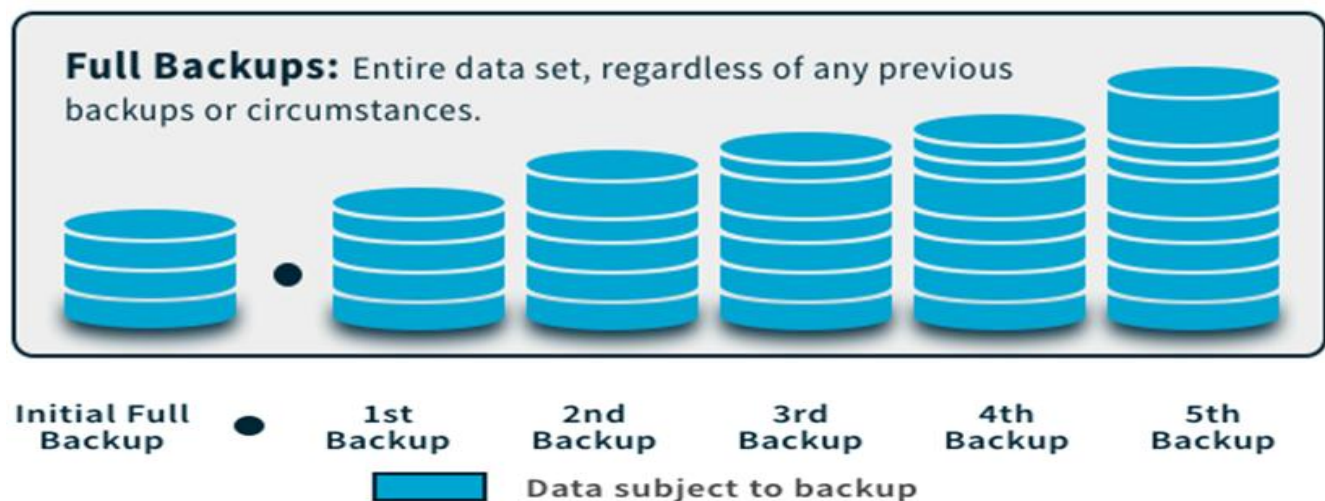


Figure 2.1. The way full backup perform

1. Incremental backups

An incremental backup operation will result in copying only the data that has changed since the last backup operation of any type. An organization typically uses the modified time stamp on files and compares it to the time stamp of the last backup. Backup applications track and record the date and time that backup operations occur in order to track files modified since these operations.



Because an incremental backup will only copy data since the last backup of any type, an organization may run it as often as desired, with only the most recent changes stored. The benefit of an incremental backup is that it copies a smaller amount of data than a full. Thus, these operations will have a faster backup speed, and require fewer medium to store the backup.

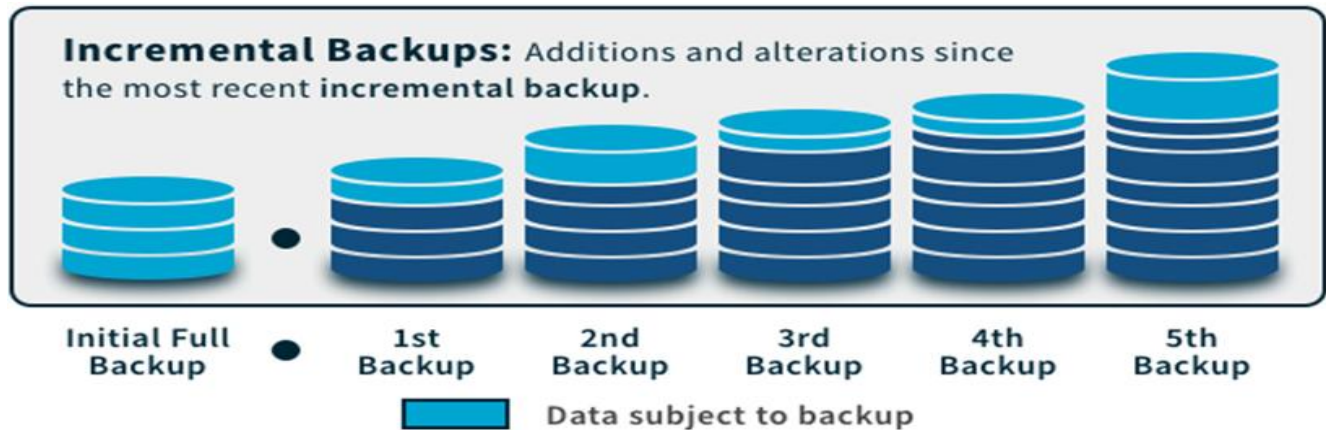


Figure 2.2. The way incremental backup perform

2. Differential backups

A differential backup operation is similar to an incremental the first time it is performed, in that it will copy all data changed from the previous backup. However, each time it is run afterwards, it will continue to copy all data changed since the previous full backup. Thus, it will store more backed up data than an incremental on subsequent operations, although typically far less than a full backup. Moreover, differential backups require more space and time to complete than incremental backups, although less than full backups.

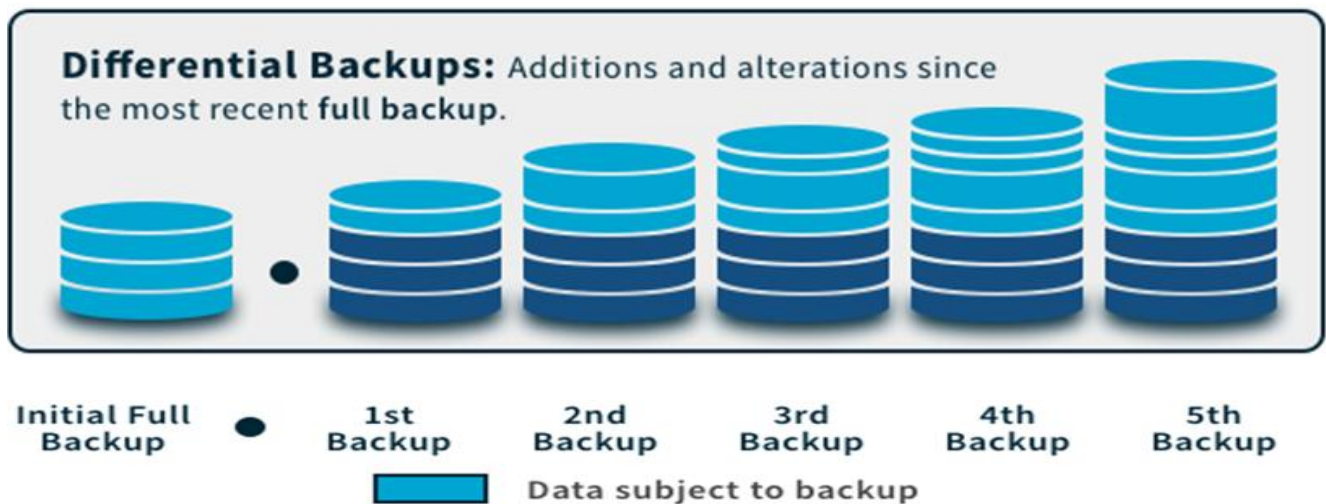


Figure 2.3. The way differential backup perform



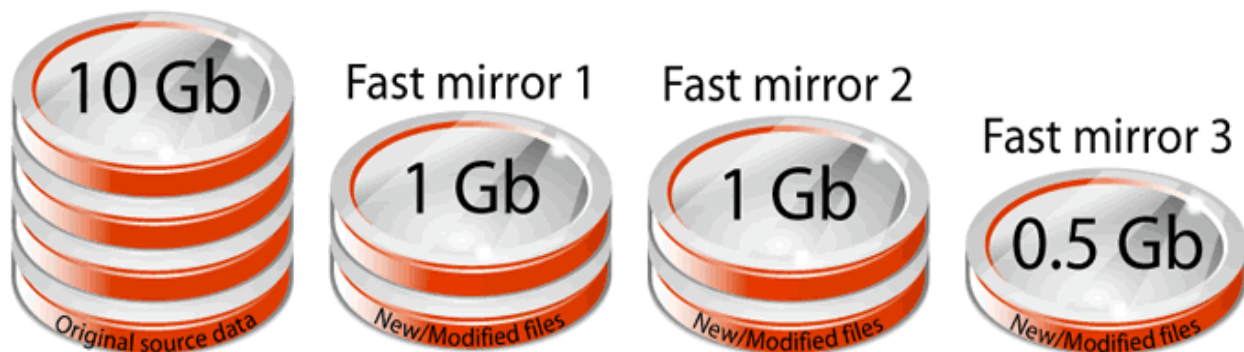
3. Mirror backups

A mirror backup is comparable to a full backup. This backup type creates an exact copy of the source data set, but only the latest data version is stored in the backup repository with no track of different versions of the files. The backup is a mirror of the source data. All the different backed up files are stored separately, like they are in the source.

One of the benefits of mirror backup is a fast data recovery time. It's also easy to access individual backed up files. Mirror backup is the fastest backup method because it copies files and folders to the destination without any compression.

One of the main drawbacks, though, is the amount of storage space required. It needs more storage space than any other backup type; password protection is not possible and cannot track different versions of files

With that extra storage, organizations should be wary of cost increases and maintenance needs. In addition, if there's a problem in the source data set, such as a corruption or deletion, the mirror backup experiences the same. As a result, it's a good idea not to rely on mirror backups for all your data protection needs, and to have other types of backup for the data. You'll want to follow the **3-2-1 rule** of backup, which includes three copies of data on two different media, with one copy off site.



Note: First time when it runs, mirror backup will back-up everything without archiving. After that only new/modified files.

Figure 2.5. The way mirror backup perform



1.1.2. Determining appropriate methods

To determine the type of backup strategy to be used there are different determinant factors such as overall business cost, performance, data protection levels, total amount of data retained and availability goals.

Do the right or appropriate backup for your organization. For organizations with small data sets, running a daily full backup provides a high level of protection without much additional storage space costs. Larger organizations or those with more data or server volume find that running a weekly full backup, coupled with either daily incremental backups or differential backups, provides a better option. Using differentials provides a higher level of data protection with less restore time for most scenarios and a small increase in storage capacity. For this reason, using a strategy of weekly full backups with daily differential backups is a good option for many organizations.

An organization must run a full backup at least once. For subsequent backups, it is possible to run either another full, an incremental or a differential backup. The first partial backup performed either a differential or incremental, will back up the same data. By the third backup operation, the data that is backed up with an incremental is limited to the changes since the last incremental. In comparison, the third backup with a differential will back up all changes since the first full backup, which was "Backup 1."

From these types of backup, it is possible to develop an approach for comprehensive data protection. An organization often uses one of the following backup settings:

- Full daily
- Full weekly + differential daily
- Full weekly + incremental daily

Performing a full backup daily requires the most amount of space, and will also take the most amount of time. However, more total copies of data are available, and fewer pieces of media are required to perform a restore operation. As a result, implementing this



backup policy has a higher tolerance to disasters, and provides the least time to restore, since any piece of data required will be located on at most one backup set.

As an alternative, performing a full backup weekly, coupled with running incremental backups daily, will deliver the shortest backup time during weekdays and use the least amount of storage space. However, there are fewer copies of data available and restore time is the longest, since an organization may need to use six sets of media to recover the necessary information. If data is needed from data backed up on Wednesday, the Sunday full backup, plus the Monday, Tuesday and Wednesday incremental media sets, is required. This can dramatically increase recovery times, and requires that each media set work properly; a failure in one backup set can impact the entire restoration.

Running a weekly full backup plus daily differential backup's delivers results in between the other alternatives. Namely, more backup media sets are required to restore than with a daily full policy, although less than with a daily incremental policy. Also, the restore time is less than using daily incremental backups, and more than daily full backups. In order to restore data from a particular day, at most two media sets are required, diminishing the time needed to recover and the potential for problems with an unreadable backup set.

Most of the advanced types of backup such as synthetic full, mirror and continuous data protection require disk storage as the backup target. A synthetic full simply reconstructs the full backup image using all required incremental backups or the differential backup on disk. This synthetic full may then be stored to tape for offsite storage, with the advantage being reduced restoration time. Finally, continuous data protection enables a greater number of restoration points than traditional backup options.

**Self-Check -2****Written Test**

Directions: Answer all the questions listed below. Use the Answer sheet provided by your instructor to check whether your answer is correct or not.

I. Choose the best answer (2 pts each)

1. Which of the following backups are faster to backup and more efficient from a storage perspective?
 - A. Differential
 - B. Incremental
 - C. Full
 - D. All
2. Point out the correct statement.
 - A. Using differential backups can save available space
 - B. Using full backups can save available space
 - C. Using partial backups can save available space
 - D. None
3. Backing up your SQL Server database is essential for _____ your data.
 - A. Protecting
 - B. Preventing
 - C. Replication
 - D. None

Note: Satisfactory rating – 4 and 8 points

Unsatisfactory - below 4 points

You can ask you teacher for the copy of the correct answers.

Answer Sheet

Name: _____

Date: _____

Score = _____



Information sheet 3: identifying and examining Risks and failure scenarios

3.1. Identifying and examining Risks and failure scenarios

A database is the means of organizing information so it can be easily managed, updated and retrieved. Losing a database would also mean losing the associated data. This means that if a business loses its databases for any number of reasons, with no backups stored, it is fair to assume that they will most likely lose the data too. There are many types of failures that can affect database processing. Some failures affect the main memory only, while others involve secondary storage. There are different scenarios or causes that could lead to a database loss. Some of them include;



Figure 3.1. Cause of database failure



3.1.1. Power Failure

Power failures can lead to hardware failure. The affected hardware components could be cables, power supplies or storage devices. It can render the data either inaccessible or simply result in data loss. One would need to isolate the affected area before investigating if the database was affected by the power failure.

3.1.2. Disk Failure

While power failures can lead to disk failure, they can also fail due to physical damage or a logical failure. Such failures are due to head crashes or unreadable media, resulting in the loss of parts of secondary storage. They are the most dangerous failures and one of the most common causes of data loss.

3.1.3. Human Error (Carelessness)

This is the failure due to unintentional destruction of data or facilities by operators or users. An employee may unintentionally delete some data or may modify the data unknowingly in a way that would cease the DBMS from interacting with the database effectively. When DBMS software is unable to interact with the database, it causes a ripple effect since the remaining third party applications relying on the DBMS to interact with the database also lose access to it. Human error is the number one cause of data loss.

3.1.4. Software Corruption

Companies using traditional in-house IT infrastructures are more at risk of software corruption than those relying on cloud-based services. While cloud vendors provide flexibility and scalability of resources, traditional IT environments have fixed sets of hardware resources which they manually upgrade.

When the number of end users in a company increase, the applications using the same resources are divided even further among the new users, causing problems such as freezing and crashing of the operating systems and applications in the middle of using the software. Crashing causes the end user to lose the unsaved data. Repeated crashing can



especially cause serious damage if the user is working on a database. These are logical errors in the program that is accessing the database, which cause one or more transactions to fail. Software failure may include a failure a failures related to software such as, operating systems, DBMS software, application Programs and so on.

3.1.5. Virus Infection

An enterprise cannot operate safely without the use of a good security solution. Cyber attacks are the biggest threat a company faces today and it is imperative that the security solution performs real-time scanning. Depending on the type of virus, it could have the ability to steal, corrupt, modify and even delete the complete database.

3.1.6. Natural Disasters

Natural disasters such as fire, floods, earthquake, tsunami, etc have the ability to destroy the entire infrastructure. In such an event, there is absolutely no way to even find, let alone recover, the data.

3.1.7. Disgruntled Employees

A disgruntled employee could provide essential and confidential information to outsiders, causing untold damage to an organization. And if the employee has access or gains unauthorized access to systems or applications, he/she can inject a virus or delete data to halt the company's day to day operations.

3.1.8. Hardware Failure

Hardware failure may include memory errors, disk crashes, bad disk sectors, disk full error and so on. Hardware failure can also be attributed to design errors, poor quality control during fabrication, overloading and wear out of mechanical parts.

3.1.8. System Crash

System crashes are due to hardware or software errors, resulting in the loss of main memory. This could be the situation that the system has entered an undesirable state, such as Dead Lock, which prevents the program form continuing with normal processing.



3.1.9. Network Failure

Network failure can occur while using a Client-server configuration or distributed database system where multiple database servers are connected y common network. • Network failure such as communication software failure or aborted asynchronous connections will interrupt the normal operation of the database system.

3.1.10. Sabotages

These are failures due to international corruption or destruction of data, hardware or users.

Page 17 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020

**Self-Check -3****Written Test**

Directions: Answer all the questions listed below. Use the Answer sheet provided in the next page:

Choose the best Answer (2 pts each)

1. Unintentional deletion or modification to data which causes data loss.
 - 1.2. Human Error
 - 1.3. Natural Disaster
 - 1.4. System crash
 - 1.5. Disk failure
2. A hardware or software errors that resulting in the loss of main memory.
 - A. Power failure
 - B. System crash
 - C. Sabotage
 - D. Virus
3. The failure that related to operating systems, DBMS and other application Programs.
 - A. Hardware failure
 - B. System crash
 - C. Software failure
 - D. Natural Disasters
4. A failure that related to fire, floods, earthquake and tsunami
 - A. Human Error
 - B. Natural Disaster
 - C. System crash
 - D. Disk failure

Note: Satisfactory rating - 4 points

Unsatisfactory - below 4points

You can ask you teacher for the copy of the correct answers.

Answer Sheet

Name _____ Date _____

Score = _____

Rating: _____

Page 18 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020



L #41	LO #2- Determine backup methods appropriate to database requirements
Instruction sheet	
This learning guide is developed to provide you the necessary information regarding the following content coverage and topics: • Evaluating range of back-up and restoration • Completing full off-line back-ups • Completing on-line file back-ups • Employing disk mirroring and redundant array of inexpensive disks (raid) hard disk configurations to keep copies of files • Arranging Off-site copies of back-up files This guide will also assist you to attain the learning outcomes stated in the cover page. Specifically, upon completion of this learning guide, you will be able to: • Evaluate range of back-up and restoration • Complete full off-line back-ups • Complete on-line file back-ups • Employ disk mirroring and redundant array of inexpensive disks (raid) hard disk configurations to keep copies of files • Arrange Off-site copies of back-up files	
Learning Instructions:	
Read the specific objectives of this Learning Guide. 1. Follow the instructions described below. 2. Read the information written in the “Information Sheets”. Try to understand what are being discussed. Ask your trainer for assistance if you have hard time understanding them. 1. Accomplish the “Self-checks” which are placed following all information sheets. 2. Ask from your trainer the key to correction (key answers) or you can request your trainer to correct your work. (You are to get the key answer only after you finished answering the Self-checks). 3. If you earned a satisfactory evaluation proceed to the next learning guide, 4. If your performance is unsatisfactory, see your trainer for further instructions.	



Information sheet 1: Evaluating range of back-up and restoration

1.1. Evaluating range of back-up and restoration

While each approach carries its own benefits and risks, organizations need to consider their need for performance, data protection, their total volume of data assets, and the cost of recovery. The following five factors can be used in making a decision about which backup schedule is right for you.

1. How active are your read/write activities?

If your databases and applications are actively being updated with new data at a high rate, known to database specialists as "write activities", full backups could be more efficient. If you are primarily using your data assets for reference without updating them, known as "read activities," you may not need full backups on a very consistent basis.

2. What is your tolerance for recovery time?

With a full backup on a daily basis, all of your assets are in a single set. While a full recovery isn't quite immediate, it can occur very quickly and doesn't require the combination of multiple types of backup files. If your tolerance to any downtime is zero, full backups represent the least risk.

3. How many of your data assets are actively being update?

Unless all of your data assets, applications, and databases are "living," running full backups on a very consistent basis may take more storage space than necessary.

4. How much storage space can you dedicate?

Running a full backup on a daily basis requires more than twice the storage space of differential or incremental in many cases. Assuming your business is actively using 25% of your data assets on a daily basis, running a full daily backup each weekday could require five times more storage space than a weekly full backup and a daily incremental or differential backup. At most organizations, the difference is significant.

Page 20 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020



5. How much data do you have?

For some organizations, running a full backup daily is actually the most cost-effective approach. These are typically organizations with minimal data assets, which can be a product of their industry, products, services, or a lack of multimedia data assets. If cost and storage space factors are not prohibitive, a full backup represents the easiest and fastest recovery.

Table 1.1. Comparison of backup type

	Full	Differential	Incremental
Storage Space	High	Medium to High	Low
Backup Speed	Slowest	Fast	Fastest
Restoration Speed	Fastest	Fast	Slowest
Media Required for Recovery	Most recent backup only	Most recent full backup & most recent differential backup	Most recent full backup & all incremental backups since full backup
Duplication	Stores a lot of duplicate files	Stores duplicate files	No duplicate files

1.1.1. The advantages and disadvantages of each backup type

- **Pros of Full Backups**

- ✓ Potential for fast, total recovery of data assets.
- ✓ Simple access to the most recent backup version.
- ✓ All back-ups are contained in a single version.
- ✓ Minimal time needed to restore business operations.

- **Cons of Full Backups**

- ✓ Requires the most storage space.
- ✓ Demands the most bandwidth.
- ✓ Relatively time-consuming to complete the backup process

- **Pros of Incremental Backups**



- ✓ Minimal time to complete backup.
- ✓ Requires the least storage space.
- ✓ Demands the least bandwidth.
- **Cons of Incremental Backups**
 - ✓ Recovery time may be slower.
 - ✓ Requires a full backup in addition to incremental backups for complete recovery.
 - ✓ Recovery requires the piecing together of data from multiple backup sets.
 - ✓ Small potential for incomplete data recovery if one or more backup sets has failed.
- **Pros of Differential Backups**
 - ✓ Requires less storage space than full backups.
 - ✓ Only two backups (last full and most recent incremental) are required for recovery.
- **Cons of Differential Backups**
 - ✓ Slower than incremental.
 - ✓ Requires an initial full backup for complete recovery.
 - ✓ IT will need to piece together two backup sets.
 - ✓ Potential for failed recovery if one or more backups is incomplete

**Self-Check -1****Written Test**

Directions: Answer all the questions listed below. Use the Answer sheet provided by your instructor to check whether your answer is correct or not

Choose the best answer (2pts each)

1. A backup type that require the highest storage space.
 - A. Full
 - B. Differential
 - C. Incremental
 - D. All
2. Recovery time may be slower.
 - A. Differential
 - B. Full
 - C. Incremental
 - D. All
3. A backup type with the fastest backup speed
 - A. Incremental
 - B. Differential
 - C. Full
 - D. All
4. Which backup type cannot store duplicate file?
 - A. Full
 - B. Incremental
 - C. Differential
 - D. All

Note: Satisfactory rating - 4points

Unsatisfactory - below 4points

You can ask you teacher for the copy of the correct answers.

Answer Sheet

Score = _____

Rating: _____

Name _____ Date _____



Information sheet 2: Completing full off-line back-ups

2.1. Completing full off-line back-ups

Off-line backup is also called cold backup or static backup. It is a database backup during which the database is offline and not accessible to update or the database operations are entirely stopped, and then the backup is performed. While the backup is in progress, no business operations can be performed. This requires the data access to be completely shut off from the front-end and neither online users nor background processes can access the database for the duration of the backup.

It is mostly accomplished before the beginning of the day or at the end of the day. Here a single backup version is made. Since no new data is added to this in real-time, the backup is performed swiftly and only once.

Cold backups consume fewer resources but have a limitation. The database cannot be accessed when the backup operation is in progress.

The advantage of this method is that users are still able to access the system during the backup. However, if the server crashes, the backup will also be gone. The risk that comes with a hot backup is that the data may be modified during the process, resulting in inconsistent data.

2.1.1. Taking Database Offline

i. Taking Database Offline using SSMS(SQL Server Management Studio)

1. Login to SQL Server Management Studio.
2. In the Object Explorer, select the **database** you want to **take offline** and right-click.
3. In the right-click menu go to Tasks >> **Take Offline**.
4. In the pop-up window, choose the check box under the Drop All Active Connections and click OK

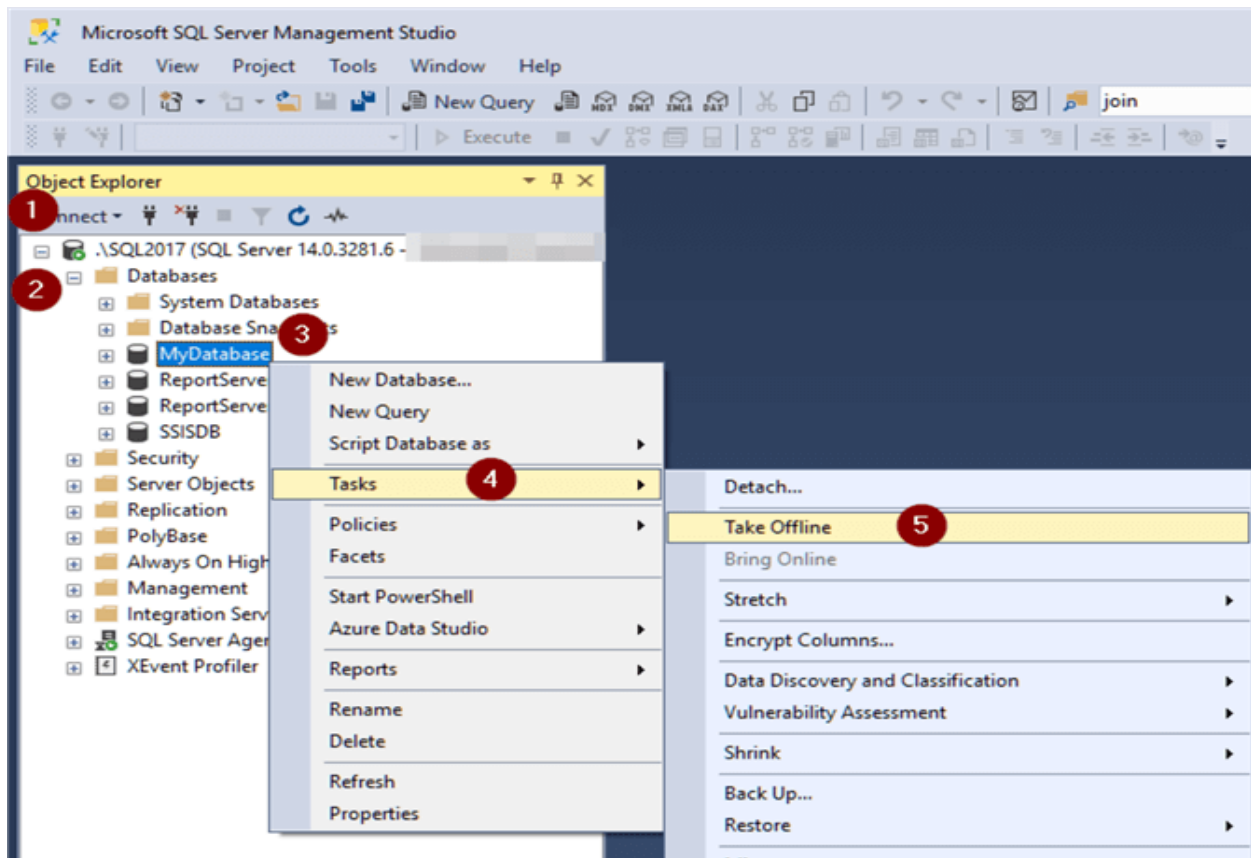


Figure 2.1. Taking Database Offline using SSMS

ii. Taking Database Offline using T-SQL

Open SQL Server Management Studio and open a Query Editor pane. Enter and execute the following code:

```
USE master  
GO
```

```
ALTER DATABASE <<Database-name>> SET OFFLINE
```

2.1.2. Take Full backup

i. Backing Up the Database Using T-SQL

```
BACKUP DATABASE << databaseName >> TO DISK = 'C:\Backupname.BAK'
```

The command is `BACKUP DATABASE databaseName`. The "TO DISK" option specifies that the backup should be written to disk and the location and filename to create the backup is specified.



- **Create a full SQL Server backup to multiple disk files**

This command uses the "DISK" option multiple times to write the backup to three equally sized smaller files instead of one large file.

```
BACKUP DATABASE << databaseName >>
TO DISK = 'C:\ Backupname _1.BAK',
DISK = 'D:\ Backupname _2.BAK',
DISK = 'E:\ Backupname _3.BAK'
GO
```

- **Create a full SQL Server backup with a password**

This command creates a backup with a password that will need to be supplied when restoring the database.

```
BACKUP DATABASE << databaseName >>
TO DISK = 'C:\ Backupname.BAK'
WITH PASSWORD = 'Q!W@E#R$'
GO
```

- **Create a full SQL Server backup with progress stats**

This command creates a full backup and also displays the progress of the backup. The default is to show progress after every 10%.

```
BACKUP DATABASE << databaseName >>
TO DISK = 'C:\ Backupname.BAK'
WITH STATS
GO
```

Here is another option showing stats after every 1%.

```
BACKUP DATABASE << databaseName >>
TO DISK = 'C:\ Backupname.BAK'
WITH STATS = 1
GO
```

- **Create a SQL Server backup and give it a description**

This command uses the description option to give the backup a name. This can later be used with some of the restore commands to see what is contained with the backup. The maximum size is 255 characters.



```
BACKUP DATABASE << databaseName >>
TO DISK = 'C:\ Backupname.BAK'
WITH DESCRIPTION = 'Full backup for << databaseName >>'
GO
```

- **Create a mirrored SQL Server backup**

One of the features found in the Enterprise Edition of SQL Server is the ability to take mirrored backups. Basically, taking a mirrored backup means creating additional copies of the backup media (up to three) using a single BACKUP command, eliminating the need to perform the copies with copy or robocopy.

This option allows you to create multiple copies of the backups, preferably to different locations.

```
BACKUP DATABASE <<databaseName>>
TO DISK = 'C:\backup_name.BAK'
MIRROR TO DISK = 'D:\ backup_name _mirror.BAK'
WITH FORMAT
GO
```

- **Specifying multiple options for SQL Server Backups**

```
BACKUP DATABASE <<databaseName>>
TO DISK = 'C:\ backup_name.BAK'
MIRROR TO DISK = 'D:\ backup_name _mirror.BAK'
WITH FORMAT, STATS, PASSWORD = 'Q!W@E#R$'
GO
```

2.2. Detaching and Attaching a Database

There may be a time in the life of our SQL Server database when we have to move it from one server to another, or in fact just from one hard drive to another if it is getting full. Possibly we can move our database from an old slower server to a new faster server or a server on a better network. By detaching and reattaching the database, we can do this simply and easily.



Detaching a database physically removes the details from the SQL Server master and msdb databases, but does not remove the files from the disk that it resides on. However, detaching the database from SQL Server will then allow you to safely move, copy, or delete the files that make up the database, if you so desired. This is the only way that a database should be physically removed from a server for moving it.

Syntax to detach;

```
sp_detach_db 'databaseName'
```

Attaching a database involves informing SQL Server of the name and the location of the data files and the transaction log files. This data can be placed anywhere on a computer, but it is recommended to place the data in a sensible location.

```
CREATE DATABASE 'databaseName'
```

```
ON (FILENAME='C:\Program Files (x86)\Microsoft SQL  
Server\MSSQL10.MSSQLSERVER\MSSQL\DATA\dd.MDF')  
FOR ATTACH
```

**Self-Check -2****Written Test**

Directions: Answer all the questions listed below. Use the Answer sheet provided by your instructor to check whether your answer is correct or not

Write Choose the best answer(2pts each)

1. The database cannot be accessed when the backup operation is in progress.
 - A. Offline
 - B. Online
 - C. Offsite
 - D. onsite
2. Which one of the following statements allows someone to take backup of database named ICT on local disk D:\ using DBA as a backup name?
 - A. BACKUP DATABASE DBA to DISK = D:\ICT.BAK
 - B. BACKUP DATABASE ICT to DISK = D:\DBA.BAK
 - C. BACKUP DATABASE DBA to DISK = 'D:\ICT.BAK'
 - D. BACKUP DATABASE ICT to DISK = 'D:\DBA.BAK'
3. It physically removes the database details from the SQL Server master and msdb databases, even if the files exist on the disk that it resides on.
 - A. Mirroring
 - B. Backing up
 - C. Detaching
 - D. All
4. It is also called cold backup or static backup
 - A. Offline
 - B. Online
 - C. Offsite
 - D. onsite

Note: Satisfactory rating - 4 points

Unsatisfactory - below 4 points

You can ask you teacher for the copy of the correct answers.

Answer Sheet

Name _____ Date _____

Score = _____

Rating: _____



Information sheet 3: Completing On-line file back-ups

3.1. Completing On-line file back-ups

It is also called a hot backup or dynamic backup. A hot backup is a backup performed while the database is open and available for use (read and write activity). It is performed in near real-time when the systems are up and running, and new data is continuously generated or captured.

In a hot backup, there is a time parameter involved as to when to perform a backup. This can range from seconds to minutes. The entire data is copied on the secondary location, and hence the relevant changes reflect in the new backup. Hot backups are a bit resource intensive as there are multiple iterations stored at a time. This allows the user to restore the backup to a required point.

The most important advantage here is the capability to continue business operations while the backup is in progress. The database is available at all times, and hence the business can continue as usual.

Table 3.1. Comparison of Hot Backup and Cold Backup

HOT BACKUP	COLD BACKUP
Also, known as dynamic backup.	Also, known as static backup.
Hot backups are resource intensive.	Cold backups consume fewer resources in comparison.
It can be performed when the systems are up and running.	Database operations need to be stopped when the backup is performed.
Database is available at all times.	Database can't be even accessed when the backup is in progress.



If you are an organization with business operations working around the clock and cannot afford any disturbance or downtime, then hot backup is the one for you.

This way, you can safeguard your data and keep your business applications and operations running. If you are an organization that has fixed working hours, then cold backup is better for you. The data which has been updated through the working day will be easily copied over without any hindrances.

3.1.1. Taking Database Online

Taking Database Online using SSMS(SQL Server Management Studio)

1. Login to SQL Server Management Studio.
2. In the Object Explorer, select the **database** you want to **take online** and right-click.
3. In the right-click menu go to Tasks >> **Take Online**.
4. In the pop-up window, choose the check box under the Drop All Active Connections and click OK

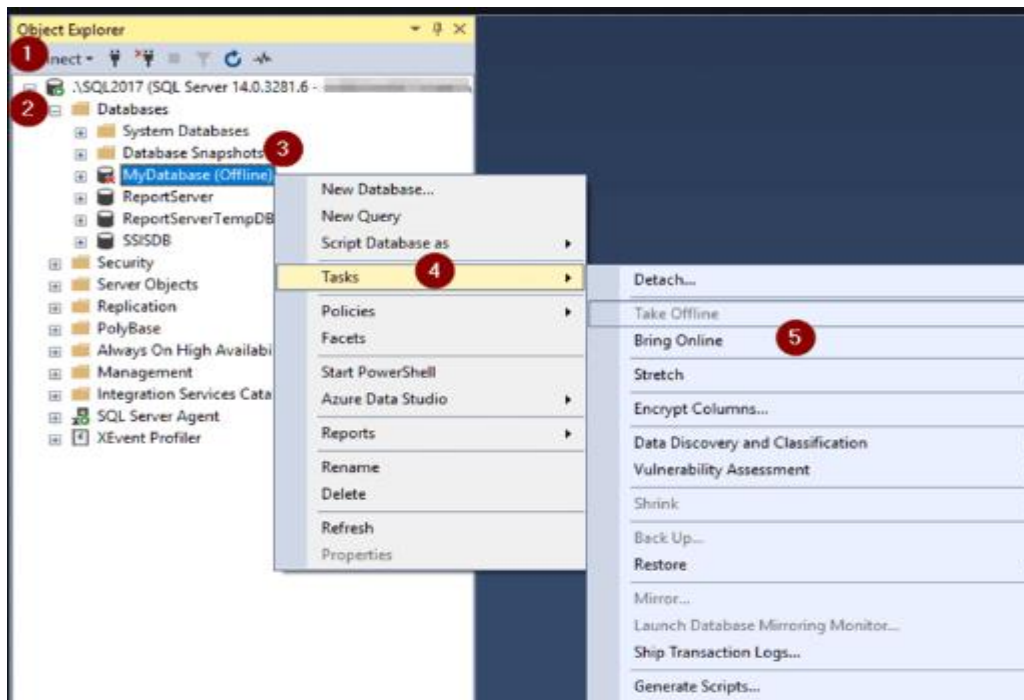


Figure 3.1 How to bring SQL server database online



Taking Database On line using T-SQL

Open SQL Server Management Studio and open a Query Editor pane. Enter and execute the following code:

```
USE master  
go  
ALTER DATABASE <<Database-name>> SET ONLINE
```

3.1.2. Take Full backup

Backing Up the Database Using T-SQL

```
BACKUP DATABASE << databaseName >> TO DISK = 'C:\Backupname.BAK'
```

**Self-Check -3****Written Test**

Directions: Answer all the questions listed below. Use the Answer sheet provided by your instructor to check whether your answer is correct or not

Choose the best answer (2pts each)

1. Backup can occur while the database is in use.
 - A. Online
 - B. Offline
 - C. Restoring
 - D. All
2. A resource intensive backup.
 - A. Dynamic
 - B. Online
 - C. Cold backup
 - D. A & B
 - E. B & C
3. It allows the users to continue their operations while the backup is in progress.
 - A. Dynamic
 - B. Offline
 - C. Cold backup
 - D. Offsite backup
4. Which one of the following syntax is correct?
 - A. ALTER <<Database-name>> DATABASE SET ONLINE
 - B. ALTER DATABASE <<Database-name> SET ONLINE
 - C. SET ONLINE DATABASE <<Database-name>>
 - D. None

Note: Satisfactory rating - 4 points

Unsatisfactory - below 4 points

You can ask you teacher for the copy of the correct answers.

Answer Sheet

Name _____ Date _____

Score = _____

Rating: _____



Information sheet 4: Employing Disk mirroring and redundant array of inexpensive disks (RAID) hard disk configurations to keep copies of files

4.1. Disk mirroring

Disk mirroring, also known as RAID 1, is the replication of data to two or more disks. Disk mirroring is a strong option for data that needs high availability because of its quick recovery time. It's also helpful for disaster recovery because of its immediate failover capability. Disk mirroring requires at least two physical drives. If one hard drive fails, an organization can use the mirror copy. While disk mirroring offers comprehensive data protection, it requires a lot of storage capacity.

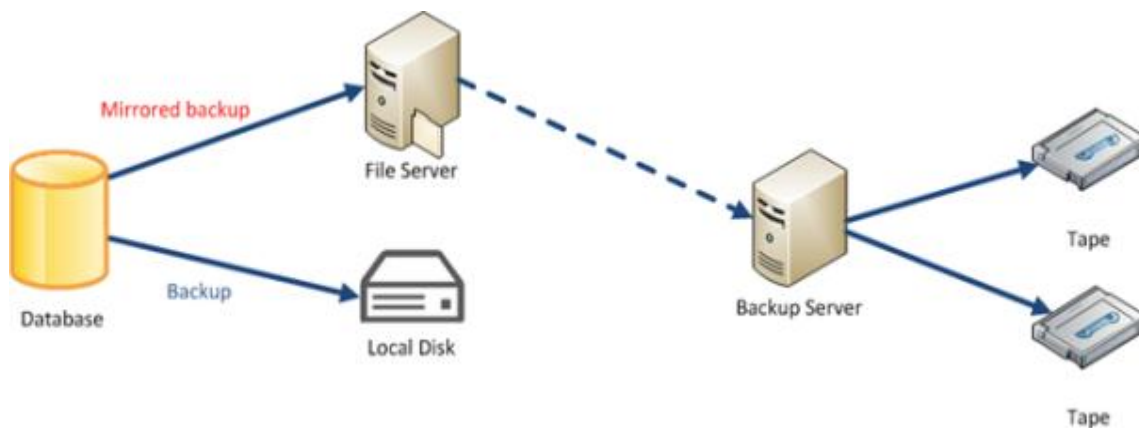


Figure 4.1 Database mirroring

4.2. RAID

RAID refers to redundancy array of the independent disk. It is a technology which is used to connect multiple secondary storage devices for increased performance, data redundancy or both. It gives the ability to survive one or more drive failure depending upon the RAID level used.

It consists of an array of disks in which multiple disks are connected to achieve different goals.

There are 7 levels of RAID schemes. These schemas are as RAID 0, RAID 1, ..., RAID 6.

These levels contain the following characteristics:

- It contains a set of physical disk drives.
- The operating system views these separate disks as a single logical disk.



- In this technology, data is distributed across the physical drives of the array.
- Redundancy disk capacity is used to store parity information.
- In case of disk failure, the parity information can be helped to recover the data.

4.2.1. RAID 0

RAID level 0 provides data stripping, i.e., a data can place across multiple disks. It is based on stripping that means if one disk fails then all data in the array is lost. This level doesn't provide fault tolerance but increases the system performance.

Pros of RAID 0:

- In this level, throughput is increased because multiple data requests probably not on the same disk.
- This level full utilizes the disk space and provides high performance.
- It requires minimum 2 drives.

Cons of RAID 0:

- It doesn't contain any error detection mechanism.
- The RAID 0 is not a true RAID because it is not fault-tolerance.
- In this level, failure of either disk results in complete data loss in respective array.

4.2.2. RAID 1

This level is called mirroring of data as it copies the data from drive 1 to drive 2. It provides 100% redundancy in case of a failure. Only half space of the drive is used to store the data. The other half of drive is just a mirror to the already stored data.

Pros of RAID 1:

- The main advantage of RAID 1 is fault tolerance. In this level, if one disk fails, then the other automatically takes over.
- In this level, the array will function even if any one of the drives fails.

Cons of RAID 1:

- In this level, one extra drive is required per drive for mirroring, so the expense is higher.

4.2.3. RAID 2

RAID 2 consists of bit-level striping using hamming code parity. In this level, each data bit in a word is recorded on a separate disk and ECC code of data words is stored on

Page 35 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020



different set disks. Due to its high cost and complex structure, this level is not commercially used. The same performance can be achieved by RAID 3 at a lower cost.

Pros of RAID 2:

- This level uses one designated drive to store parity.
- It uses the hamming code for error detection.

Cons of RAID 2:

- It requires an additional drive for error detection.

4.2.4. RAID 3

RAID 3 consists of byte-level striping with dedicated parity. In this level, the parity information is stored for each disk section and written to a dedicated parity drive. In case of drive failure, the parity drive is accessed, and data is reconstructed from the remaining devices. Once the failed drive is replaced, the missing data can be restored on the new drive. In this level, data can be transferred in bulk. Thus high-speed data transmission is possible.

Pros of RAID 3:

- In this level, data is regenerated using parity drive.
- It contains high data transfer rates.
- In this level, data is accessed in parallel.

Cons of RAID 3:

- It required an additional drive for parity.
- It gives a slow performance for operating on small sized files.

4.2.5. RAID 4

RAID 4 consists of block-level stripping with a parity disk. Instead of duplicating data, the RAID 4 adopts a parity-based approach. This level allows recovery of at most 1 disk failure due to the way parity works. In this level, if more than one disk fails, then there is no way to recover the data. Level 3 and level 4 both are required at least three disks to implement RAID.

In this level, parity can be calculated using an XOR function. If the data bits are 0,0,0,1 then the parity bits is XOR (0,1,0,0) = 1. If the parity bits are 0,0,1,1 then the parity bit is XOR (0,0,1,1)= 0. That means, even number of one results in parity 0 and an odd number of one results in parity 1.

This level allows us to recover lost data.

Page 36 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020



4.2.6. RAID 5

RAID 5 is a slight modification of the RAID 4 system. The only difference is that in RAID 5, the parity rotates among the drives. It consists of block-level striping with DISTRIBUTED parity. Same as RAID 4, this level allows recovery of at most 1 disk failure. If more than one disk fails, then there is no way for data recovery. This level was introduced to make the random write performance better.

Pros of RAID 5:

- This level is cost effective and provides high performance.
- In this level, parity is distributed across the disks in an array.
- It is used to make the random write performance better.

Cons of RAID 5:

- In this level, disk failure recovery takes longer time as parity has to be calculated from all available drives.
- This level cannot survive in concurrent drive failure.

4.2.7. RAID 6

This level is an extension of RAID 5. It contains block-level striping with 2 parity bits. In RAID 6, you can survive 2 concurrent disk failures. Suppose you are using RAID 5, and RAID 1. When your disks fail, you need to replace the failed disk because if simultaneously another disk fails then you won't be able to recover any of the data, so in this case RAID 6 plays its part where you can survive two concurrent disk failures before you run out of options.

Pros of RAID 6:

- This level performs RAID 0 to strip data and RAID 1 to mirror. In this level, striping is performed before mirroring.
- In this level, drives required should be multiple of 2.

Cons of RAID 6:

- It is not utilized 100% disk capability as half is used for mirroring.
- It contains very limited scalability.

**Self-Check -4****Written Test**

Directions: Answer all the questions listed below. Use the Answer sheet provided by your instructor to check whether your answer is correct or not

Choose the best answer (2 pts each)

1. A technology that used to connect multiple secondary storage devices in order to increase performance and data redundancy.
 - A. Mirroring
 - B. RAID
 - C. Detaching
 - D. None
2. Which RAID level is called a mirroring of data?
 - A. RAID 0
 - B. RAID 1
 - C. RAID 2
 - D. RAID 7
 - E. All level
3. How many disk drives are required to implement disk mirroring?
 - A. At least two physical drives.
 - B. At least one disk drive and two logical drives
 - C. At least two logical drives
 - D. None
4. Which one of the following statement is incorrect about RAID?
 - A. It contains a set of physical disk drives.
 - B. The operating system views these separate disks as a single logical disk.
 - C. Data is distributed across the physical drives of the array.
 - D. None

Note: Satisfactory rating - 4 points

Unsatisfactory - below 4 points

You can ask you teacher for the copy of the correct answers.

Answer Sheet

Name _____ Date _____

Score = _____

Rating: _____



Information sheet 5: Arranging Off-site copies of back-up files

5.1. Off-site back-up files

Offsite backup is the replication of the data to a server which is separated geographically from a production systems site. Offsite data backup may also be done via direct access, over Wide Area Network (WAN). An offsite backup is a backup process or facility that stores backup data or applications external to the organization or core IT environment.

It is similar to a standard backup process, but uses a facility or storage media that is not physically located within the organization's core infrastructure.

Offsite backups are primarily used in data backup and disaster-recovery measures. The core objective behind storing and maintaining data at a backup facility is to:

- Secure data from malicious attacks
- Keep a backup copy of data in case the primary site is damaged or destroyed

Cloud backup, online backup or managed backup are examples of offsite backup solutions that enable an individual or organization to store data at facilities that are geographically and logically external.

5.1.1. Advantages and disadvantages Offsite Storage Advantages

Offsite storage has several major advantages. Some of them are;

- **Scalability:** The cloud provider or managed service provider is responsible for offering storage that can be expanded on demand.
- **Cost & Value:** Offsite and cloud storage is extremely affordable; in most circumstances, you only pay for what you use. There is no initial outlay for expensive storage platforms and no additional maintenance or support contracts.
- **Fast Deployment:** With offsite cloud storage, the infrastructure is already in place, and the platform is already available to use. There is no waiting for purchasing and installation of hardware and the client can simply plug into the cloud storage and get started immediately.
- **Managed Storage Service:** The cloud provider will usually have a team of storage experts and subject matter experts who own the solution and manage the service for

Page 39 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020



its clients. Many providers offer automated object storage deployment, where blobs of data storage can be assigned and removed when needed.

- **Connectivity:** Offsite storage can be made available over an internet connection or a dedicated virtual private network for added security. This flexible approach to connectivity is great for clients, as it makes accessing data extremely easy. Dedicated network links can provide high-speed data pipes between the clients and the storage for ultra-fast connectivity.
- **Performance:** Offsite storage performance has gotten significantly faster in recent years. Network improvements have boosted the performance of storage and, in most circumstances, provide near real-time response rates.

Disadvantages

Some disadvantages of offsite storage include;

- **It can be difficult to access** the data when it is needed. For example, some of the offsite data servers will be routinely taken offline in order to perform preventive maintenance. During that time period, access to the offsite data will either be severely limited or completely cut off. In most cases, clients who use offsite storage are given plenty of notice when a site is going down for maintenance or repair.
- **Security and Privacy:** One of the major concerns with offsite cloud storage is the security and integrity of data. Data protection and privacy are extremely important for business organizations. If you choose to move data to an offsite storage provider, consideration must be given to compliance as well as the security measures in place to protect the data. Offsite storage must be protected from unauthorized access and should always be encrypted.
- **Compliance and Data Governance:** there are complex compliance and governance rules which must be adhered to when securing or destroying data.
- **Lifetime Costs:** Offsite storage is often affordable; however, there may be a risk of vendor lock-in when an organization is reliant on the provider's organization. This may result in a lifetime of monthly charges for data usage, so it's important to weigh up if the overall costs and benefits outweigh implementing local storage.



Speed: Although cloud/offsite performance is generally very good, some data-intense applications may perform better using local storage. Even if dedicated express routes are used, latency and network bottlenecks may impact performance.

Noisy Neighbours: When choosing offsite storage, it's important to understand whether you will be leveraging dedicated offsite storage or if you will be using shared storage.

5.2. Onsite Backup

In onsite storage, data and storage hardware are geographically located internally to your business or organization. You may have a computer room or data center onsite where the storage arrays are securely located. All your internal systems will have direct access to the storage within the same building or organization, usually over an internal Local Area Network (LAN).

Organizations supporting databases with a high data change rate will often employ an onsite backup strategy for quick recovery in the event of a failure.

Onsite storage usually entails storing important data on a periodic basis on local storage devices, such as hard drives, DVDs, magnetic tapes, or CDs. Offsite storage requires storing important data on a remote server, usually via the Internet, although it can also be done via direct access.

Advantages of onsite storage:

- Immediate access to data
- Less expensive
- Internet access not needed
- Control of your own data security
- Performance improvement

Disadvantage of onsite storage;

- In the event of a catastrophic event, onsite data storage can be destroyed. For example, if there is a fire in the building, or a water main bursts, the onsite servers can lose all the data that has been collected on them. In addition, onsite storage units can also be stolen, resulting in a loss of time, money, and data.
- Storage can be extremely expensive depending on the size of the storage array.



- Storage device will need to be managed, maintained, and upgraded in-house.

5.3. Hybrid storage

This article will focus on onsite and offsite storage advantages and disadvantages; however, for the benefit of completeness, it's important to mention hybrid storage. Hybrid storage is a mix of offsite and onsite storage approaches. A typical hybrid setup could be a local storage array at a head office which replicates data to an external cloud-based location; such a setup could be for backups or data integrity reasons.

Page 42 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020

**Self-Check -5****Written Test**

Directions: Answer all the questions listed below. Use the Answer sheet provided by your instructor to check whether your answer is correct or not

Choose the best answer (2 pts each)

1. Backing up the data to a server which is geographically separated from a production systems.
 - A. Online
 - B. Offline
 - C. Offsite
 - D. Onsite
2. A mix of offsite and onsite storage approaches for the benefit of completeness.
 - A. Dynamic
 - B. Hybrid
 - C. Cloud
 - D. Cold
3. One is not the advantages of onsite backup
 - A. May be destroyed during a catastrophic event
 - B. Control of your own data security
 - C. Immediate access to data
 - D. Less expensive

Note: Satisfactory rating - 4points

Unsatisfactory - below 4 points

You can ask you teacher for the copy of the correct answers.

Answer Sheet

Name _____ Date _____

Score = _____

Rating: _____

Page 43 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020



Operation sheet 1: How to Take database Backup Using SSMS and T-SQL

i. Backup Using T-SQL

1. Backup database database_name to disk='path\backupname.bak'
2. Backup log database_name to disk='path\backupname.bak'

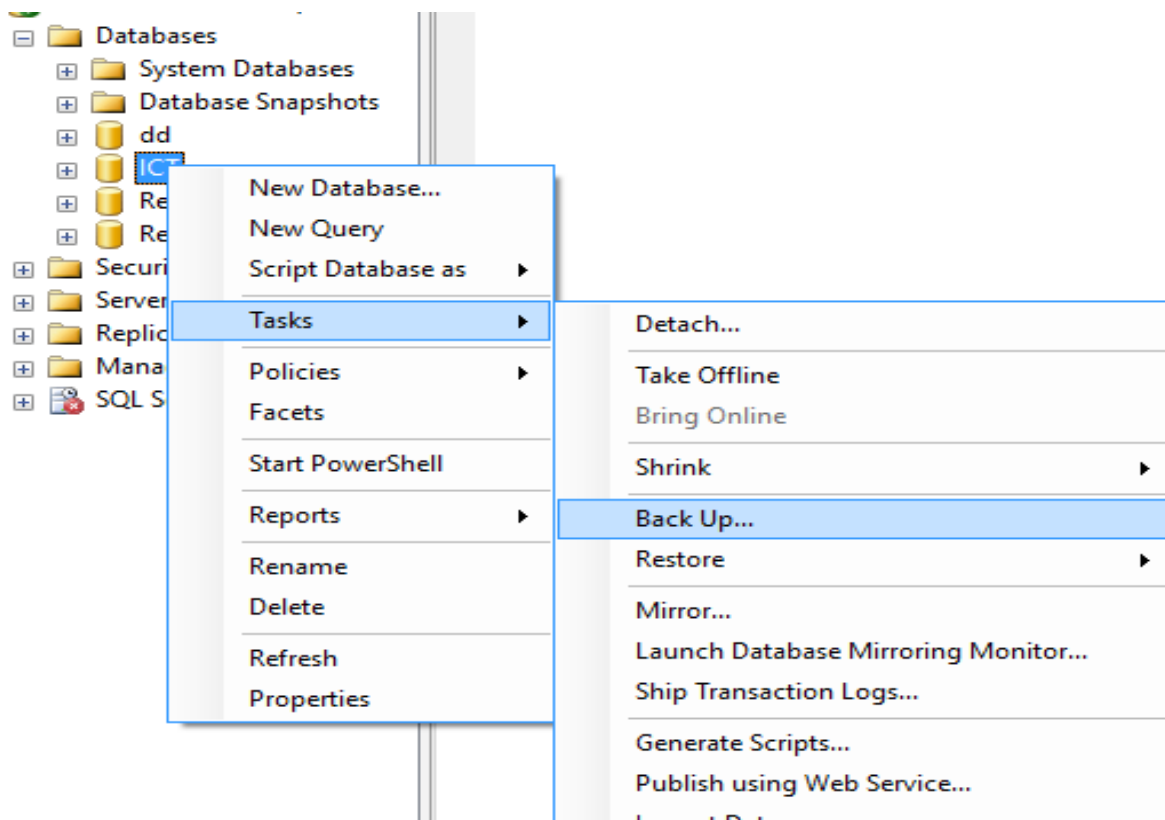
For example, to take the backup of a database named ICT into local disk D:\ using "ICTdata" as backup name; you should have to write the following SQL Command;

Backup database ICT to disk='D:\ICTdata.bak'

Backup log ICT to disk='D:\ICTdata.bak'

ii. Backup Using SSMS

1. Open SQL Server 2008
2. Connect the Server and point to your database
3. Make right click on your database and Select Task→ click on Backup





Back Up Database - ICT

Select a page

- General
- Options

Script Help

Source

Database: ICT

Recovery model: FULL

Backup type: Full

☐ Copy Only Backup

Backup component:

☒ Database

☐ Files and filegroups:

Backup set

Name: ICT-Full Database Backup

Description:

Backup set will expire:

☒ After: 0 days

☐ On: 03/01/2021

Destination

Back up to: ☒ Disk ☐ Tape

C:\Program Files (x86)\Microsoft SQL Server\MSSQL10.MSSQLSERVER\MSS

Add...

Remove

Contents

OK Cancel

Connection

Server: DESKTOP-QBA&JK2

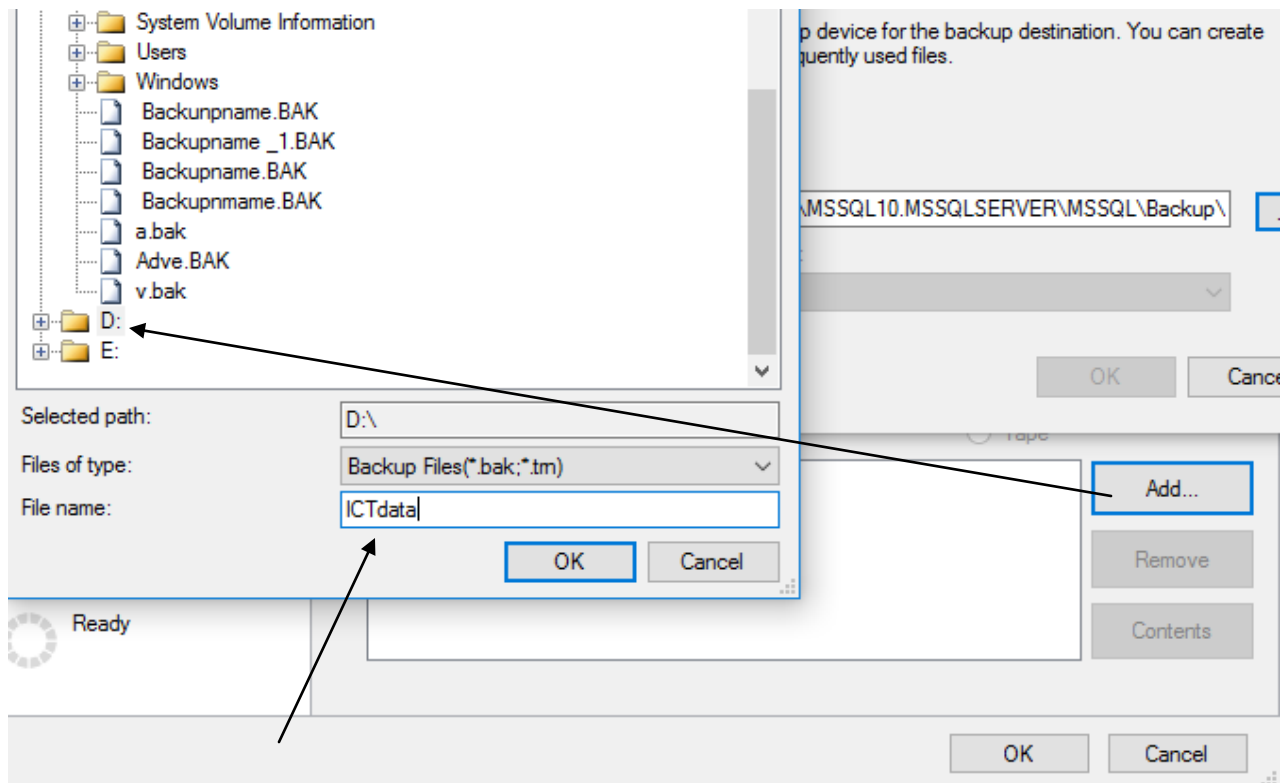
Connection: DESKTOP-QBA&JK2\End

[View connection properties](#)

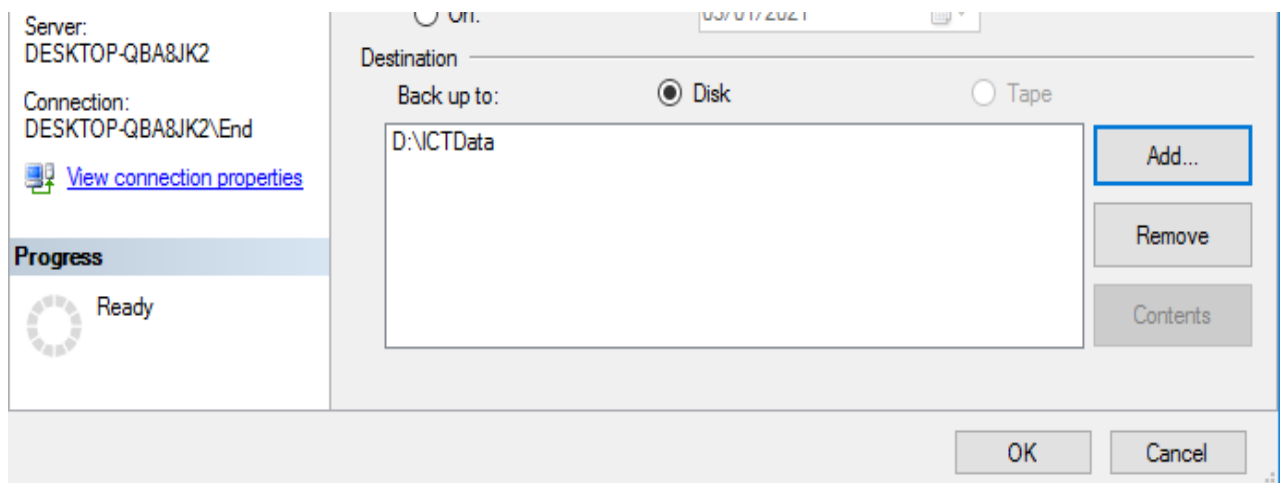
Progress

Ready

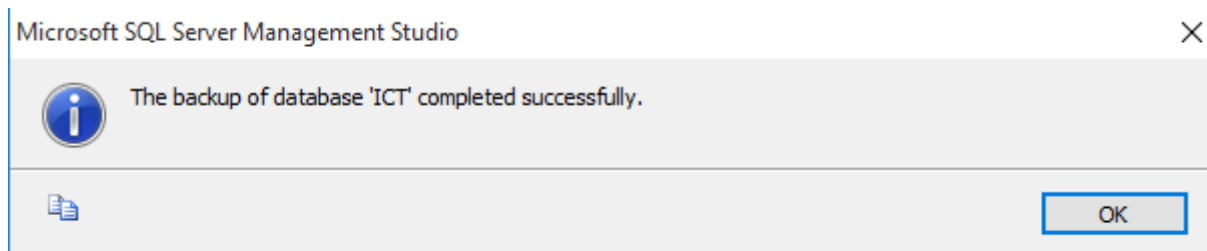
4. Select the backup destination and write the backup name



5. Click on OK button



6. Click on OK button





LAP Test 1: Take database Backup Using SSMS and T-SQL

Name: _____ Date: _____

Time allotted: 30

Starting time _____ Ending Time _____

Task1: Create a database named TVET.

Task 2: Create a folder named “**Data**” on the desktop and take the backup of your database in **Data** folder using “SSMSbackup” and “TSQLbackup” as a backup name using both SSMS and T-SQL respectively.



L #41	LO #3- Establish recovery points and disaster recovery procedures
Instruction sheet	
This learning guide is developed to provide you the necessary information regarding the following content coverage and topics: • Determining database recovery points based on the back-up arrangements • Testing restore process • Completing restoration of the database to the point of failure without loss of committed transactions This guide will also assist you to attain the learning outcomes stated in the cover page. Specifically, upon completion of this learning guide, you will be able to: • Determine database recovery points based on the back-up arrangements • Test restore process • Complete restoration of the database to the point of failure without loss of committed transactions	
Learning Instructions:	
Read the specific objectives of this Learning Guide. 1. Follow the instructions described below. 2. Read the information written in the “Information Sheets”. Try to understand what are being discussed. Ask your trainer for assistance if you have hard time understanding them. 3. Accomplish the “Self-checks” which are placed following all information sheets. 4. Ask from your trainer the key to correction (key answers) or you can request your trainer to correct your work. (You are to get the key answer only after you finished answering the Self-checks). 5. If you earned a satisfactory evaluation proceed to the next learning guide.	



Information sheet 1: Determining database recovery points based on the back-up arrangements

1.1. Database recovery point

Database recovery is the process of restoring the database to a correct (consistent) state in the event of a failure. In other words, it is the process of restoring the database to the most recent consistent state that existed shortly before the time of system failure.

There are many situations in which a transaction may not reach a commit or abort point. Some of them include;

- An operating system crash can terminate the DBMS processes
- The DBMS can crash
 - ✓ System failure(e.g. power outage)
 - ❖ affects all transactions currently in progress but does not physically damage the data (softcrash)
 - ✓ Media failures(e.g. Head crash on the disk)
 - ❖ damage to the database (hard crash)
 - ❖ need backup data
- The system might lose power
- Human error can result in deletion of critical data.

In any of these situations, data in the database may become inconsistent or lost.

When a DBMS recovers from a crash, it should maintain the following –

- It should check the states of all the transactions, which were being executed.
- A transaction may be in the middle of some operation; the DBMS must ensure the atomicity of the transaction in this case.
- It should check whether the transaction can be completed now or it needs to be rolled back.
- No transactions would be allowed to leave the DBMS in an inconsistent state.



In case of any type of failures, a transaction must either be aborted or committed to maintain data integrity.

Transaction log plays an important role for database recovery and bringing the database in a consistent state in the event of failure. Transactions represent the basic unit of recovery in a database system. The recovery manager guarantees the atomicity and durability properties of transactions in the event of failures. During recovery from failure, the recovery manager ensures that either all the effects of a given transaction are permanently recorded in the database or none of them are recorded. A transaction begins with successful execution of a <T, BEGIN>” (begin transaction) statement.

1.1.1. Database Recovery Techniques

For fast restoration or recovery of data, the database must hold tools which recover the data efficiently. It should have atomicity means either the transactions showing the consequence of successful accomplishment perpetually in the database or the transaction must have no sign of accomplishment consequence in the database.

So, recovery techniques which are based on deferred update and immediate update or backing up data can be used to stop loss in the database.

- **Immediate Update:** As soon as a data item is modified in cache, the disk copy is updated.
- **Deferred Update:** All modified data items in the cache are written either after a transaction ends its execution or after a fixed number of transactions have completed their execution.
- **Shadow update:** The modified version of a data item does not overwrite its disk copy but is written at a separate disk location.
- **In-place update:** The disk version of the data item is overwritten by the cache version.



1.1.2. Two approaches of Recovery

- **Manual Reprocessing**

In a **Manual Reprocessing** recovery approach, the database is periodically backed up (a database save) and all transactions applied since the last save are recorded

If the system crashes, the latest database backup set is restored and all of the transactions are re-applied (by users) to bring the database back up to the point just before the crash.

Several shortcomings to the Manual Reprocessing approach:

- ✓ Time required to re-apply transactions
 - ✓ Transactions might have other (physical) consequences
 - ✓ Re-applying concurrent transactions in the same original sequence may not be possible.
- **Automated Recovery with Rollback / Roll forward**
 - ✓ Introduce a **Log** file – this is a file separate from the data that records all of the changes made to the database by transactions. Also referred to as a **Journal**.
 - ✓ This transaction log Includes information helpful to the recovery process such as: A transaction identifier, the date and time, the user running the transaction, before images and after images.
 - ❖ **Before Image**: A copy of the table record (or data item) before it was changed by the transaction.
 - ❖ **After Image**: A copy of the table record (or data item) after it was changed by the transaction.
 - ✓ The Automated Recovery process uses both rollback and roll forward to restore the database.
 - ❖ **Rollback**: Undo any partially completed transactions (ones in progress when the crash occurred) by applying the before images to the database. **UNDO** the transactions in progress at the time of failure.
 - ❖ **Roll forward**: Redo the transactions by applying the after images to the database. This is done for transactions that were committed before the crash. **REDO** the transactions that successfully complete but did not write to the physical disk.
 - ✓ **Checkpoint** is a mechanism where all the previous logs are removed from the system and stored permanently in a storage disk. **Checkpoint** declares a **point** before which the **DBMS** was in consistent state, and all the transactions were



committed. **Checkpoints** can also be taken (less time consuming) in between database saves.

- ❖ The DBMS flushes all pending transactions and writes all data to disk and transaction log.
- ❖ Database can be recovered from the last checkpoint in much less time.

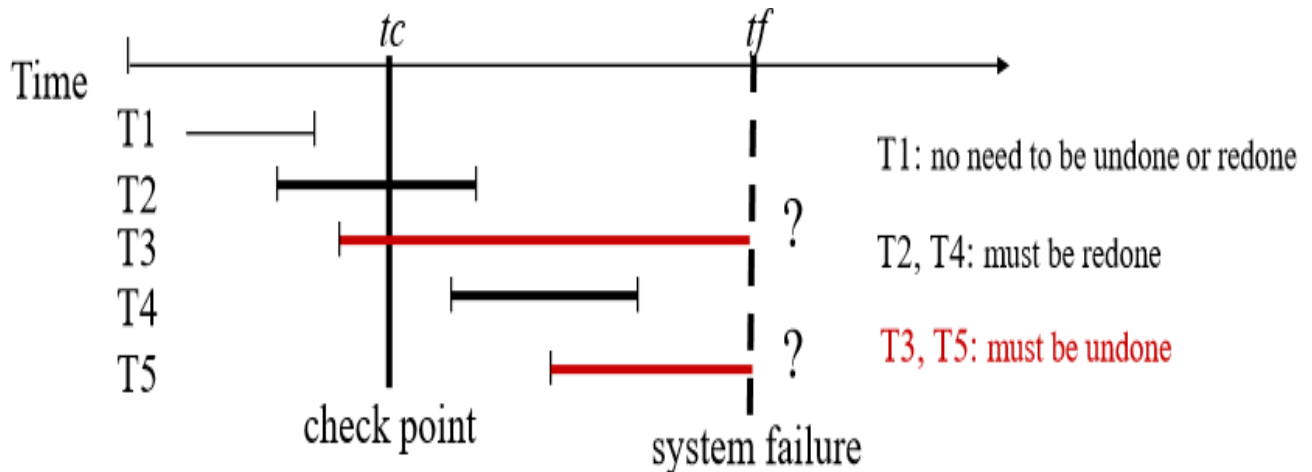


Figure 1.1. Recovery with Rollback / Roll forward

For example in the figure above;

STEP1: UNDO-list = list of transactions given in the checkpoint record = {T2, T3}

REDO-list = { }

STEP2: Search forward through the log, starting from the checkpoint, to the end of log:

- If a 'BEGIN TRANSACTION' is found => add to UNDO-list {T2, T3, T4, T5}
- If a 'COMMIT' is found => remove from UNDO-list to REDO-list

UNDO-list = {T3, T5} (System works backward through the log, undoing the UNDO-List.)

REDO-list = {T2, T4} (System then works forward through the log, redoing the REDO-List)

**Self-Check -1****Written Test**

Directions: Answer all the questions listed below. Use the Answer sheet provided by your instructor to check whether your answer is correct or not

Choose the best answer(2pts each)

1. _____ the process of restoring the database to the consistent state in the event of a failure.
 - A. Restore
 - B. Attach
 - C. Recovery
 - D. All
2. A process by which DBMS REDO the transactions that successfully complete before failure but did not write to the physical disk.
 - A. Recover
 - B. Roll Forward
 - C. Rollback
 - D. Restore
3. A recovery method by which the modified version of a data item is written at a separate disk location.
 - A. Differed Update
 - B. Shadow update
 - C. Immediate update
 - D. None
4. Factors that enforce the transaction not to reach a commit or abort point.
 - A. Human error
 - B. operating system crash
 - C. DBMS crash
 - D. All

Note: Satisfactory rating - 4 points

Unsatisfactory - below 4 points

You can ask you teacher for the copy of the correct answers.

Answer Sheet

Score = _____

Rating: _____



Information sheet 2: Testing restore process

2.1. Testing restore process

Test Database recovery testing is used to ensure that the database is recovered. Recovery testing allows you to find out whether the application is running properly and to check retrieving invaluable data that would have been lost if your recovery method is not properly setup.

You also check if several critical processes are running smooth to ensure that the data recovery will pass smoothly through the testing phase.

The key aim of backup testing is to ensure the business can retrieve its data and continue operations. Businesses should test that they can restore files, folders and volumes from backups on a storage volume, user and application basis. Backup testing should be regular and routine. In an ideal world, businesses would test every backup, but that is rarely practical.

Common Steps in Database Backup and Recovery Testing

In database recovery testing, you need to run the test in the actual environment to check if the system or the data can actually be recovered in case of any disasters and any other unforeseen events in the business environment.

The common actions performed in Database Recovery Testing –

- Testing of database system
- Testing of the SQL files
- Testing of partial files
- Testing of data backup
- Testing of Backup tool
- Testing log backups

These policies should set out the recovery point objective (RPO) and the recovery time objective (RTO).

The RPO sets out how old the most recent backup can be, or put another way, the amount of data loss the organisation can tolerate and still operate. The RTO specifies how quickly systems must be recovered. Unless the business tests recovery, CIOs will not know if they can meet the RTO and RPO, or if recovery works at all.

**Self-Check -2****Written Test**

Directions: Answer all the questions listed below. Use the Answer sheet provided by your instructor to check whether your answer is correct or not

Choose the best answer(2pts each)

1. It sets out the amount of data loss the organisation can tolerate.
 - A. Recovery point objective
 - B. Recovery time objective
 - C. Recovery
 - D. All
2. The main aim of backup testing is;
 - A. To check the functionality of the server
 - B. To ensure that the business can retrieve its data and continue operations.
 - C. To check if the system or the data can be properly managed
 - D. None
3. It specifies how quickly systems must be recovered.
 - A. RPO
 - B. RTO
 - C. CIO
 - D. All

Note: Satisfactory rating - 3 points

Unsatisfactory - below 3 points

You can ask you teacher for the copy of the correct answers.

Answer Sheet

Score = _____

Rating: _____

Name: _____ Date: _____



Information sheet 3: Completing restoration of the database to the point of failure without loss of committed transactions

3.1. Restore a database to a point in time

A point-in-time recovery can be used to return the database data and database object to its functional state prior to detrimental action has been performed.

The ability to perform this kind of recovery depends on a recovery model set for the database. The database must be in either the Full or Bulk-Logged recovery model. In case the Simple recovery mode was used, this recovery method is not possible.

In case of using the Bulk-Logged recovery model some errors may occur and recovery to a point in time might fail. An error will be thrown in case when any bulk-logged operations were performed. As such operations are minimally logged; there is not sufficient data in a particular transaction log.

When you issue a RESTORE DATABASE or RESTORE LOG command the WITH RECOVERY option is used by default.

If you restore a "Full" backup the default setting it to RESTORE WITH RECOVERY, so after the database has been restored it can then be used by your end users.

If you are restoring a database using multiple backup files, you would use the WITH NORECOVERY option for each restore except the last.

If your database is still in the restoring state and you want to recover it without restoring additional backups you can issue a RESTORE DATABASE. WITH RECOVERY to bring the database online for users to use.

The RESTORE ... WITH NORECOVERY option puts the database into a "restoring" state, so additional backups can be restored. When the database is in a "restoring" state no users can access the database or the database contents.

When you issue a RESTORE DATABASE or RESTORE LOG command; the WITH NORECOVERY option allows you to restore additional backup files before recovering the



database. This therefore allows you to get the database as current as possible before letting your end users access the data.

This option is not on by default, so if you need to recover a database by restoring multiple backup files and forget to use this option you have to start the backup process all over again.

The most common example of this would be to restore a "Full" backup and one or more "Transaction Log" backups.

3.1.1. Restore a database using T-SQL

- **Restore full backup and one transaction log backup**

The first command does the restore and leaves the database in a restoring state and second command restores the transaction log backup and then makes the database useable.

```
RESTORE DATABASE <<DatabaseName>> FROM DISK = 'C:\BackupName.BAK'  
WITH NORECOVERY  
GO  
RESTORE LOG <<DatabaseName>> FROM DISK = 'C:\BackupName.TRN'  
WITH RECOVERY  
GO
```

- **Restore full backup and two transaction log backups**

This restores the first two backups using NORECOVERY and then RECOVERY for the last restore.

```
RESTORE DATABASE <<DatabaseName>> FROM DISK = 'C:\ BackupName.BAK'  
WITH NORECOVERY  
GO  
RESTORE LOG <<DatabaseName>> FROM DISK = 'C:\ BackupName.TRN'  
WITH NORECOVERY  
GO  
RESTORE LOG <<DatabaseName>> FROM DISK = 'C:\ BackupName1.TRN'  
WITH RECOVERY
```



GO

- **Restore full backup, latest differential and two transaction log backups**

This restores the first three backups using NORECOVERY and then RECOVERY for the last restore.

```
RESTORE DATABASE <<DatabaseName>> FROM DISK = 'C:\ BackupName.BAK'  
WITH NORECOVERY
```

GO

```
RESTORE DATABASE <<DatabaseName>> FROM DISK = 'C:\ BackupName.DIF'  
WITH NORECOVERY
```

GO

```
RESTORE LOG <<DatabaseName>> FROM DISK = 'C:\ BackupName.TRN'  
WITH NORECOVERY
```

GO

```
RESTORE LOG <<DatabaseName>> FROM DISK = 'C:\ BackupName1.TRN'  
WITH RECOVERY
```

GO

3.1.2. Restore a database using SQL Server Management Studio

When restoring using SSMS the WITH RECOVERY option is used by default, so there is nothing that needs to be set but this can be set or changed on the options page when restoring.

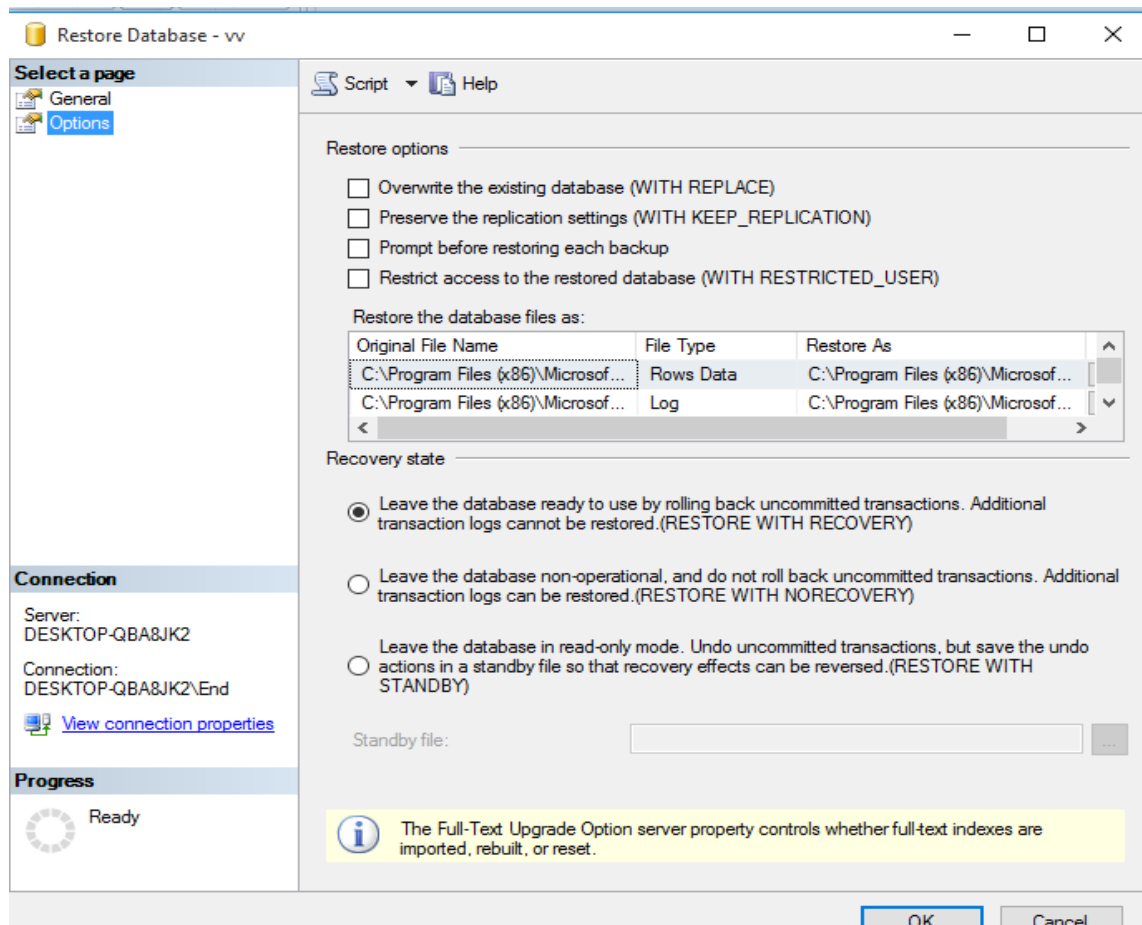


Figure 3.1: To restore a database backup using the WITH RECOVERY option

**Self-Check -3****Written Test**

Directions: Answer all the questions listed below. Use the Answer sheet provided by your instructor to check whether your answer is correct or not

Choose the best answer(2pts each)

3. To perform the point-in-time recovery the database recovery model must be in;
 - A. Full recovery model
 - B. Bulk-Logged recovery model
 - C. Simple recovery model
 - D. A and B
 - E. A and C
4. Which one is the default setting if you restore a Full backup?
 - A. RESTORE WITH RECOVERY
 - B. RESTORE WITH NORECOVERY
 - C. There is no default setting
5. If you are restoring a database using multiple backup files, which one is the suitable option for each restore except the last?
 - A. RESTORE WITH RECOVERY
 - B. RESTORE WITH NORECOVERY
 - C. Only RESTORE
 - D. None
6. Which one of the following command does the restore and leaves the database in a restoring state
 - A. RESTORE DATABASE <<DatabaseName>>FROM DISK = 'C:\BackupName.BAK'
WITH NORECOVERY
 - B. RESTORE LOG <<DatabaseName>> FROM DISK = 'C:\BackupName.TRN'
WITH RECOVERY
 - C. Neither A nor B

Note: Satisfactory rating - 4 points

Unsatisfactory - below 4 points

You can ask you teacher for the copy of the correct answers.

Answer Sheet

Score = _____

Rating: _____

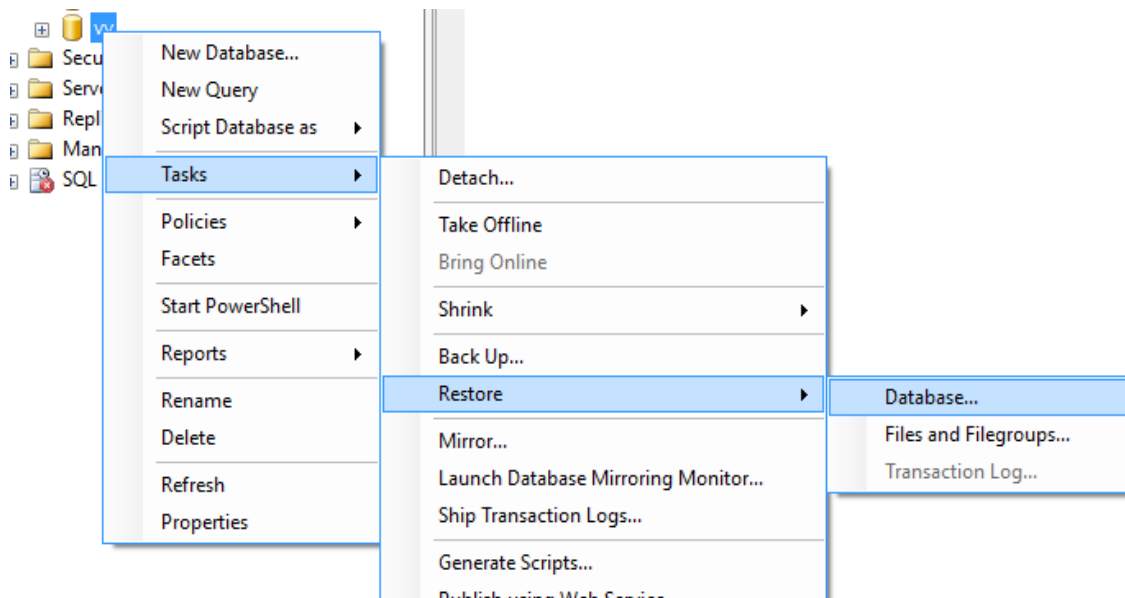
Page 60 of 69	Federal TVET Agency Author/Copyright	TVET program title:- Database Administration Level III	Version -1
			December 2020



Operation sheet1: Restoring Database from the backup

Steps to Restore database named ICT from Backup

1. Open SQL Server 2008
2. Select the Master database
3. Make right click your database (VV)
4. Point on Tasks→Restore→click on Database



5. Specify the location of your backup

Select or type the name of a new or existing database for your restore operation.

To database:

To a point in time:

Source for restore

Specify the source and location of backup sets to restore.

☐ From database:

☒ From device:

Select the backup sets to restore:

Restore	Name	Component	Type	Server	Database	Position	First L
☒		Database	Full	DESKTOP-QBA8JK2	vv	1	1900

6. Select the recovery state



Restore Database - vv

Select a page: General, Options

Script Help

Restore options

- ☐ Overwrite the existing database (WITH REPLACE)
- ☐ Preserve the replication settings (WITH KEEP_REPLICATION)
- ☐ Prompt before restoring each backup
- ☐ Restrict access to the restored database (WITH RESTRICTED_USER)

Restore the database files as:

Original File Name	File Type	Restore As
vv	Rows Data	C:\Program Files (x86)\Microsof...
vv_log	Log	C:\Program Files (x86)\Microsof...

Recovery state

- ☒ Leave the database ready to use by rolling back uncommitted transactions. Additional transaction logs cannot be restored. (RESTORE WITH RECOVERY)
- ☐ Leave the database non-operational, and do not roll back uncommitted transactions. Additional transaction logs can be restored. (RESTORE WITH NORECOVERY)
- ☐ Leave the database in read-only mode. Undo uncommitted transactions, but save the undo actions in a standby file so that recovery effects can be reversed. (RESTORE WITH STANDBY)

Standby file:

The Full-Text Upgrade Option server property controls whether full-text indexes are imported, rebuilt, or reset.

OK Cancel

7. Click OK button

Restore Database - vv

Select a page: General, Options

Script Help

Destination for restore

Select or type the name of a new or existing database for your restore operation.

To database: vv

To a point in time: Most recent possible

Source for restore

Specify the source and location of backup sets to restore.

Microsoft SQL Server Management Studio

The restore of database 'vv' completed successfully.

OK

Connection: Server: DESKTOP-QBA8JK2

Transaction Log DESKTOP-QBA8JK2 vv

8. Click OK again.

Restoring database Using T-SQL

1. Use master a database
2. Restore database **VV** from disk='D:\ICTdata.bak' with recovery



LAP Test 1: Restoring database from the backed up location

Name _____

Starting time _____ Ending Time _____

Task1: Restore TVET database from the previous backup location (under LO2).



References

1. Database Backup & Recovery B.H.GARDI COLLEGE OF ENGINEERING & TECHNOLOGY (MCA Dept.)
2. <https://www.guru99.com/dbms-architecture.html> (Lo1 inf.sh1)
3. [https://learn.org/articles/What is Database Architecture.html](https://learn.org/articles/What_is_Database_Architecture.html) (Lo1 inf.sh1)
4. <https://www.javatpoint.com/dbms-architecture> (Lo1 inf.sh1)
5. <https://searchdatabackup.techtarget.com/feature/Full-incremental-or-differential-How-to-choose-the-correct-backup-type> (Lo1 inf.sh2)
6. <https://www.datto.com/blog/data-backup-and-recovery-methods-the-basics-you-need-to-know> (Lo1 inf.sh2)
7. <https://www.backup4all.com/mirror-backup-kb.html> (Lo1 inf.sh2)
8. <https://www.atlantech.net/blog/full-backup-vs.-incremental-backup-vs.-differential-backup-which-is-best> (Lo2 inf.sh.1)
9. <https://www.hitechnectar.com/blogs/hot-backup-vs-cold-backup/> (Lo2 inf.sh3)
10. <https://www.remsoftware.com/info/difference-between-offline-backup-and-online-backup> (Lo2 inf.sh3)
11. <https://intellegensinc.com/resources/blog/difference-between-onsite-and-offsite-data-backup/> (Lo2 inf.sh5)
12. <https://www.atlantic.net/hipaa-compliant-cloud-hosting-services/pros-and-cons-of-onsite-and-offsite-data-storage/> (Lo2 inf.sh5)
13. <https://spaghettidba.com/2011/12/29/mirrored-backups-a-usefu-feature/> (Lo2 inf.sh5)
14. <https://www.mssqltips.com/sqlservertutorial/112/recovering-a-sql-server-database-that-is-in-the-restoring-state/> (Lo3 Inf.sh. 1)
15. https://www.tutorialspoint.com/dbms/dbms_data_recovery.htm (Lo3 Inf.sh. 1)



Acknowledgement

We wish to extend thanks and appreciation to the FTA, Oromia TVET bureau, and different college instructors, academic and government agencies that took vital role and donated their expertise and resource for the developments of this TTLM.

We would like also to express our appreciation to the Experts from different organizations with in the line FTA that made the developments of this TTLM possible in terms of financing and supporting technically.

The development of this TTLM held on December 2020 at Bishoftu, Ethiopia.

The trainers and Experts who are participated technically on the development of this TTLM

No	Name	Organization	Educational background & Level	Profession /Job title	Email address	Mobile
1.	Ayansa Ergiba	Ambo TVETC	Msc. ICT mgt	Instructor	aergiba@gmail.com	0917851343
2.	Ejigu Birhanu	Nekemte TVETC	Msc. ICT mgt	Instructor	ejigubiranu2011@gmail.com	0906566892
3.	Endale Lema	Adama PTC	Msc. ICT mgt	Instructor	endhiywet@gmail.com	0913292212
4.	Keresag Gadisa	Nekemte TVETC	Msc. ICT mgt	Instructor	keresag2010@gmail.com	0920420664
5.	Meseret Tezera	Atilet Kenanisa PTC	Msc. ICT mgt	Instructor	YafetMes@gmail.com	0911751285

This TTLM was developed under TVET program for Database Administration level III on the **December 2020** at **Bishoftu**, Ethiopia



Answer sheet for L #40-#42

Answer Key Module Title: Complete database backup and recovery

LO#1: Review database architecture

Self-Check

Written Test

Directions: Answer all the questions listed below. Use the Answer sheet provided in the next page:

I Self-check 1

1. A
2. C
3. D
4. D

Self-check 2

1. B
2. A
3. A

Self-check 3

1. A
2. B
3. B
4. B

LO #2- Determine backup methods appropriate to database requirements

Self-check 1

1. A
2. C
3. A
4. B

Self-check 2

1. A
2. D
3. C
4. A

Self-check 3

1. A
2. D
3. A
4. B

**Self-check 4**

1. B
2. B
3. A
4. D

Self-check 5

1. C
2. B
3. A

LAP Test 1:

1. backup database TVET to disk='C:\Users\End\Desktop\data\TSQLbackup.bak'
2. backup log TVET to disk='C:\Users\End\Desktop\data\TSQLbackup.bak'

LO #3- Establish recovery points and disaster recovery procedures**Self-check 1**

1. C
2. B
3. B
4. D

Self-check 2

1. A
2. B
3. B

Self-check 3

1. D
2. A
3. B
4. A

LAP Test 1:

Restore database TVET from disk='C:\Users\End\Desktop\data\TSQLbackup.bak'