DATA GOVERNANCE AND
MANAGEMENT TOOLKIT

January, 2021
Version 1

Data Backups

Table of Contents

INTRODUCTION

Backups are a key component of data security. In IT data management, a data backup is a copy of the current system data that is stored in a *separate location* from the main data. A file backup is a copy of the file stored in a *separate location* from the main file. If your main file is destroyed or corrupted for any reason, you can reinstate it from the backup file.

KEY POINTS IN THIS DOCUMENT

- A data backup is a copy of the current system data that is stored in a *separate location* from the main data.
- Data backups should be encrypted to protect the data.
- Physical access to data backups should be restricted and monitored.
- Securely destroy old backups and backup failures. For physical media, physical destruction is preferred. For cloud backups, the vendor should provide explicit terms for destruction within the terms of the contract.
- Take data residency into consideration when considering cloud-based data storage options and email servers.
- If your government is managing a backup of its own hardware systems, a rotating series of data backup devices must be used.
- Develop a backup policy that considers the factors outlined in the checklist below.
- Develop a data recovery plan that details how data will be restored or recovered from the backup device.
- Test backups periodically to ensure not only that the data is good and can be recovered, but to check that the media is functioning correctly and that the data recovery plan operates smoothly.

DATA BACKUP VS. DATA ARCHIVE

It is important to realize that a data backup is different from a data archive. While a **data archive** (discussed in the downloadable document "Data Archiving") is intended to capture the current state of the data and hold it for a long period of time for archival purposes (i.e. auditing, review, or future use), **data backups** are short lived and are intended to capture the state of the current system to serve a data **recovery plan**.

Data backup media types

Backups are available in different media depending on the size of the data set. Small data sets can be stored to a removal hard disk or USB type storage device, whereas some data sets require a large amount of data storage and will use digital backup tapes. Another option that can be used is cloud based data backups which eliminates the need for managing backup infrastructure, but can introduce complications related to data residency (see breakout box).

Data residency and the cloud

If you are storing data in the cloud (servers accessed over the internet), you must be aware of data residency issues. Data residency refers to the physical location of your data. If the server is located outside Canada, your data will be subject to the laws of the country it is located in. For example, data stored on a server in the United States may be accessed by the United States government under the USA PATRIOT Act. Ensure that you take data residency into consideration when considering cloud-based data storage options and email servers.

DATA BACKUP PRIVACY AND SECURITY

Regardless of your choice in media, data backups must be handled to ensure that the privacy and security of data is maintained. Data backups should be encrypted to protect the data. Physical access to data backups should be restricted and monitored. Securely destroy old backups and backup failures. For physical media, physical destruction is preferred. For cloud backups, the vendor should provide explicit terms for destruction within the terms of the contract.

DATA BACKUP SCHEDULE AND TIMING

Depending on the size of the data and how much the data changes, backups will be run on a regular schedule (e.g., daily or weekly). If your government is managing a backup of its own hardware systems, a rotating series of data backup devices need to be used. This will prevent a current backup from being overwritten with new data. While it may seem reasonable to overwrite backups with new data, backups do fail and using a rotating series of backup devices will help to reduce the risk of losing data. Additionally, consideration needs to be given to the amount of time that a backup takes to run. Some backups based on the size may take only minutes, whereas large data sets may take hours to backup to digital tape.

DATA BACKUP POLICY

When deciding on a **backup strategy or policy** your government needs to consider two key objectives:

1. **Recovery Time Objective (RTO)**

   The RTO is the maximum amount of time that your government has to bring the system back online before the data loss has a significant impact on your government's operations.

2. **Recovery Point Objective (RPO)**

   The RPO is the amount of data that your government can afford to lose before it has a significant impact on your government.

Both the RTO and RPO for the organization and data system will help to determine the type of backups used as well as the frequency of data backups.

TABLE 1: BASIC BACKUP TYPES

Type	Features
Normal (full backup)	Specific files are marked for backup. Files are marked as **backed up** by the system. If the file is changed, the file will be marked that a change has occurred and will be included in the next backup. A full backup is normally used as a baseline at the start of a backup schedule.
Copy	Similar to a **normal** backup, except that the backup state of the file is not tracked. This means that even if there has not been a change in the file, it will be backed up.
Daily	A backup of files that have changed during a daily period. A Daily backup is based on the timestamp of the files and generally copies all files included in the backup.

When planning a backup strategy, there are different backup based strategies that can be considered. Each strategy has pros and cons.

TABLE 2: BASIC BACKUP STRATEGIES

Type	Features	Considerations
Incremental	An **incremental** backup will start with a full backup and then will automatically only include changes to data since the last backup. An incremental backup is intended to be run on a regular schedule. Files are marked as backed up so that unless they change in the future, they will not be included in subsequent backups. This strategy is used when the volume of data to be protected is extremely large and prohibits a full backup in the backup time period.	**Pros**: Faster backup time and less storage space. After the first backup, backups can be faster and require less storage space as only the changes between backup intervals are stored. **Cons: Slower recovery time.** As each backup only stores incremental changes, the recovery process starts with recovering the full backup and then applying each incremental

		change. This can lead to a long recovery period as well as increasing risks of data loss as all backups in the backup period must be maintained.
Differential	A **differential** backup will start with a full backup, and similar to an **incremental** backup, will backup files that have changed. nlike an **incremental** backup, files will not be marked as being backed up. This means that once a file changes, it will always be included in all subsequent backups.	**Pros:** Faster recovery time. Upon recovery, only the **initial full backup** and the last **differential** backup need to be applied as the differential will maintain all changes since the full backup. This also means that not every differential backup must be maintained as each backup contains all changes. **Cons:** Slower backup time and more storage space. As the differential backup captures all changes during every backup, the size of the backup can be larger and take longer to complete.

The following checklist establishes key items to consider in your data backup policy.

Table 3: Data Backup Checklist

Who is involved?	✓ Create a list of key stakeholders
What data is going to be backed up?	✓ Select data for backup ✓ Establish RPO and RTO
What media will be used for backup and how will it be labelled?	✓ Select the type of media and storage solution for backup ✓ Select an appropriate labelling strategy
What type of backup will be run?	✓ Select a type of backup
When and how will backups run?	✓ Establish a backup schedule ✓ Automate the backup process

	✓ Establish logging and notification mechanism in the event of backup failure
How will the data be protected to ensure the privacy and security of the data will be maintained?	✓ Select appropriate encryption strategy for backup ✓ Establish access rules for backups ✓ Establish a logging and audit policy for accessing backups ✓ If data is being moved across the network (private or public), establish appropriate encryption for data in transit
What is the rotation cycle for backup media?	✓ Establish a rotation cycle for backup media
Where will backups be stored and transported?	✓ Establish a secure storage location for backups ✓ Establish a transportation policy for backups. Consider how data is moved and if the format changes
Who is responsible for managing and monitoring backups?	✓ Establish a backup management team and assign responsibilities ✓ Establish a backup operator schedule ✓ Review and audit logs to ensure backups are running nominally ✓ Test backups periodically for correctness ✓ Establish how long backups are kept, and how to securely destroy old backups and backup failures
How will data be recovered in the event of a failure?	✓ Develop and test a Data Recovery Plan ✓ Plan continuous scheduled review and audit of Data Recovery Plan

SULINGITUK GOVERNMENT: A data backup policy in action

Fictional Case Study

Sulingituk Government maintains a small database of information on an in-house server. The data changes slowly with a few changes to the data over any given week, so a differential strategy will be used.

Based on the database size, Sulingituk decided to use a digital linear tape backup device (DLT) to backup all data at the end of each week. A set of four backups tapes are used in the system. The backups run automatically on Thursday evening. The backup takes 4 hours to perform and the data is encrypted while being stored to the DLT device. Thursday evening was selected as the system is not used during the evenings.

On Friday morning, the current backup tape is removed from the system and replaced with the next tape. Tapes are labeled with the current backup date and the oldest backup is chosen to hold the next backup of data. This allows three weeks of data backups to be held on tapes.

The backup tape is secured and moved to an offsite storage location by the designated team member, who logs that the backup was completed and the backup tape was moved offsite. At the same time, the oldest backup tape is retrieved to be used for the next backup cycle. The team member returns the previous backup tape, makes a log entry of the return and installs it into the server for the next backup cycle.

The backup checklist summarizes the organization’s backup policy:

Who is involved?	✓ Designated team members
What data is going to be backed up?	✓ All data
What media will be used for backup and how will it be labelled?	✓ DLT ✓ Tapes will be labelled with date of backup
What type of backup will be run?	✓ Differential
When and how will backups run?	✓ Weekly Thursday evenings automatically
How will the data be protected to ensure the privacy and security of the data will be maintained?	✓ Data is encrypted when written to tape
What is the rotation cycle for backup media?	✓ Weekly rotation of 4 tapes

Where will backups be stored and transported?	✓ Moved by designated member to secure off site location ✓ Movements are logged and audited ✓ Previous tape is returned to backup pool of tapes
Who is responsible for managing and monitoring backups?	✓ Team member has been selected to manage process ✓ Supervisor reviews audit logs on a weekly basis.

DATA RECOVERY PLANS

When an organization loses data, the impact can range from a minor loss in productivity to a large scale shutdown of the organization. Having a solid backup strategy along with the systems, processes, and tools in place is key in the ability of your government to recover from a loss in data. Your government should have a strategy in place and be prepared to execute it in the event of a data loss.

A data **recovery plan** details how data will be restored or recovered from the backup device. It details the steps that are required as well as what services could be impacted during a backup recovery.

There is a chance that a portion of your government's data may be lost in the event of a disaster (i.e., a hard drive failure). If there has been a significant change in the live data since the last backup interval, there is a possibility that the recovered data may not reflect the current state.

A critical item to consider with backups is to check to ensure that they are valid. Data sets have been lost in industry due to the fact that an error occurred during the backup process. While the backup may appear to be successful, it is critical to test backups periodically to ensure not only that the data is good and can be recovered, but to check that the media is functioning correctly and that the data recovery plan operates smoothly.