TAHO
Tasmanian Archive + Heritage Office

# Information Management Advice 25 Management of Backups

## Introduction

***Backup and archiving are useful tools for managing data, especially volumes of ageing data. Agency policies for retention and access necessitate that these tools should be used to implement complementary but different policies to support business processes.***

## Data Backup

A backup system is a system designed to create a copy of all relevant data for the purposes of recovery – disaster recovery or business continuity. The data are stored in a separate data storage medium; common bulk storage media include tape and hard disk.

Backing up electronic information is part of routine Information and Communications Technology (ICT) operations, to protect the data from accidental or malicious loss or corruption. Information that is stored on backup systems act as a reserve copy of the original, to be accessed if anything happens to the original data or system. Backup systems are designed to support recovering lost data to the last known good version and ensure that systems can be restored in the event of a disaster. Backup is a process used to support active data from loss or corruption. The backup process creates a copy of all relevant data for the purposes of recovery, usually in a simple container format. Finding a particular object based on its context in a backup that is one year old is very difficult, as only file names can be readily searched for. This is particularly true in extreme cases, such as searching for old data specifically related to a legal case e.g. as in eDiscovery.

Inactive data should be removed from the backup stream, rather than backed up over and over again. This type of ageing data, which is no longer being modified, should be managed by an archiving solution in order to reduce the primary storage capacity and the amount of data being backed up for operational recovery.

**Archiving** in the classical sense is the process of ensuring proper records are made and kept as evidence that can be relied on and used to support current activities. Such records are managed in ways to ensure that they can be reused and understood in the future. This can be for current business purposes, as evidence in legal proceedings, for accountability to internal or external stakeholders, or for future historical research.

**Digital Archiving** is defined as the identification, appraisal, description, storage, preservation, management and retrieval of digital records, including all of the policies, guidelines and systems associated with these processes, so that the logical and physical integrity of the records is securely maintained over time, despite the obsolescence of technology. That is, digital archiving is not simply the process of moving data from online, more expensive storage to cheaper, offline storage.

**Data Archiving** involves moving inactive data to cheaper storage medium so that it can be removed from the online system.

## Data Archiving

Backup systems are often confused with 'data archiving'. Data Archiving can be described as the need to retain ageing data for any business purpose other than recovery. Unlike backing up a system or network, which involves copying an entire system in case of system failure, data archiving involves moving selected, usually inactive, data to a cheaper storage medium so it can be removed from the online system.

Archiving products can effectively identify information that needs to be preserved based on any number of criteria from user-defined, to file metadata, to object context. When information is archived, it is generally stored with context so that when and if it becomes necessary to find it in the future it is relatively easy to do so using advanced search criteria.

Information suitable for archiving include records that are no longer actively being changed and frequently accessed, and information that is important (inactive long term temporary records) but is less frequently accessed. Organisations may use archiving for long term historical preservation of information as an approach to storage capacity reduction, and as a tool for shrinking the backup window ( by removing inactive data from backup) or to facilitate application decommissioning and e-discovery.

This is also known as hierarchical storage management (HSM). While HSM can be seen as the electronic equivalent to moving inactive physical records to secondary storage, HSM alone is not a recordkeeping system and is not sufficient to preserve records over time for their full retention periods. HSM can form part of an appropriate records management strategy.

## Backups are not, and should not be used as, a recordkeeping system

Backups are a necessary tool for business continuity purposes, but they should not be used for purposes other than for what they were designed. Some organisations use their backup software to create "archive copies" from backed-up data. Organisations that create daily, weekly, monthly, quarterly and even yearly copies of backup to tape, and store some of these backup tapes for many years (sometimes even decades) as an archive. Backup systems are not recordkeeping systems; they cannot be relied upon as the means to maintain access to organisational records.

## Risks of retaining backup copies

Some agencies may assume that keeping records on backup tapes or other media meets their responsibility to maintain digital records over time. However, data on backup tapes or in backup systems do not commonly include the metadata required to preserve their context and make them useable over time.

Backups also save data as a single mass of information, making searching and retrieval time-consuming and difficult. For example, a single tape could contain a wide variety of records and data, all with different records retention periods under an approved Retention and Disposal Schedule. Backups may contain not only data but also applications and operating systems.

Prolonged retention of backups may pose additional risks. Backup systems generally use propriety storage compression algorithms to increase the density. Without the correct software, there is a risk that information

could become irretrievable over longer periods. In addition, vendors frequently update their software with limited periods of backwards compatibility, further exacerbating this issue.

Information in backup systems is discoverable in the event of litigation. If backups are kept for a long period, records previously destroyed in accordance with an authorised Retention and Disposal Schedule may still be discoverable.

Agencies should be aware that under the new provisions of the *Right to Information Act 2009*, agencies may be required to search a backup system for a document that is otherwise considered to be unloadable or non-existent if it is:

- a record that is required to be kept under the *Archives Act 1983*,
- has not been lawfully disposed of under the *Archives Act 1983*, and
- it is likely that the required document has been kept in, and is retrievable from, the backup system.

## Management of backup systems

The appropriate management of backup processes, including the storage media used, is usually the responsibility of Information and Communications Technology (ICT) staff within an agency. Advice on backup requirements, including the appropriate disposal of media, is contained in Information Security Manual developed by the Office of egovernment. [1]

Information managers should liaise with ICT staff to ensure that backups are kept only for the appropriate period specified by the agency. The *General Retention and Disposal Schedule for Administrative Records (DA2157)* states that backups are only required to be retained for as long as required for business/ administrative purposes. Agencies are required to determine appropriate retention periods, considering such factors as:

- the backup cycle (such as daily, weekly or monthly)
- whether backups are incremental or full backups
- the criticality and rate of change of the application or data being backed up, and
- how frequently backups are tested to ensure that the system can be recovered from the backups produced.

Decisions on the appropriate retention of backups should be documented in the agency's backup procedures.

If an agency becomes aware that a backup system holds records that are required to be kept and are not already captured in an appropriate recordkeeping system, steps should be taken to copy the record into a recordkeeping system.

[1] http://www.egovernment.tas.gov.au/

## Data Management

Good data management is something government agencies should strive for, and it is more important now than ever before as we see the volumes of data, particularly ageing data increase exponentially. Agencies need a defined Data Management Strategy that identifies how data will be managed through its life cycle as a documented set of policies and practices. Using backup as a means of providing operational recovery, and archiving records in a recordkeeping system for longer-term records retention, is an important first step.

Key points in a data management strategy are:

- Backups are used for operational recovery of recent data only. Once the backup set has passed the time when it realistically be used for recovery, the backup set should be deleted according to the retention and disposal authority. Agencies should evaluate whether or not they will need to perform a file, system or other recovery that takes them to a state older than 90 days. Will the agency really need to restore a server to the state it was in three months ago or longer? If not, your backup retention periods should be set accordingly.
- A Data Archive should be implemented for all information retention other than data protection for operational recovery. A data archive can be used for, e-discovery, the historical preservation of data, or as information repositories for future data mining. Retention periods associated with archived data and information vary greatly, and should align with policies set by the legal and information management unit, the business owner of the data and the agency's responsibilities under the *Archives Act 1983*. When an application or user needs an object that is older than 90 days, chances are that application or user knows something about the object it is looking for (i.e., the context of the file, e-mail or document). A data archive support contextual searches and the retrieval of single or multiple objects to support a specific retrieval requirement.

The access policies surrounding archived information drive the storage location, data formats, the type of repository the information resides in, role-based access and the selection of tools for mining the records. Developing these access policies is critical to putting a useful recordkeeping system in place, beyond simply defining retention policies.

# More information

Contact your agency's Information/ Records Manager for further information and advice on state records within your agency.

## Further Advice

For more detailed advice, please contact:

Government Information Strategy Unit
Tasmanian Archive and Heritage Office
91 Murray Street
HOBART TASMANIA 7000
Telephone: 03 6165 5581
Email: gisu@education.tas.gov.au

## Acknowledgements

This guideline is based on the following State Records QLD documents:

- .Management of Backups. [2]
- Evolving Best Practices for Backup, Archiving and Tape: Strategies for Alignment, Gartner Research Paper, and 7 June 2011

### Information Security Classification

This document has been security classified using the Tasmanian Government Information Security classification standard as PUBLIC and will be managed according to the requirements of the Tasmanian Government Information Security Policy.

### Document Development History
### Build Status

| Version | Date | Author | Reason | Sections |
|---|---|---|---|---|
| 2.0 | March 2015 | Christine Woods | Template | All |
| 1.0 | 23-4-2012 | | Initial Release | All |

### Amendments in this Release

| Section Title | Section Number | Amendment Summary |
|---|---|---|
| All | All | Document imported into new template |

### Issued: April 2012

**Ross Latham**
State Archivist

[2] http://www.archives.qld.gov.au/Recordkeeping/GRKDownloads/Documents/ManagementOfBackups.pdf