



Project Report

Ministry of Sustainable Resource Management

Digital Data Backup, Archive and Retention Strategy

Submitted to:

Ms. Elaine Dawson
Manager, Data Services
Information Management Branch
Business & Information Services Division
Ministry of Sustainable Resource Management
P.O. Box 9352 STN PROV GOVT
Victoria, B.C.
V8W 9M2

KPMG LLP
St. Andrews Square II
800 – 730 View Street
Victoria, B.C. V8W 3Y7

Submitted by:

Gordon S. Gunn
Senior Principal
Government Services Practice

June 17, 2002

Table of Contents

Executive Summary	1
Scope and Objectives	4
Background	4
Purpose and Objectives	5
Current Environment	6
Digital Data Backup Processes	6
Digital Data Archive Processes	10
Project Datasets	15
Project Lifecycles	18
Archive Retention	19
Business Requirements	20
Digital Data Backup Requirements	20
Digital Data Archive Requirements	21
Digital Data Backup, Retention and Archive Strategy	24
Backup Strategy	24
Archive Strategy	26
Retention Strategy	31
Implementation Plan	35
Implementation Issues	57
Appendix A – Interview List	A - 1
Appendix B – Glossary of Terms	B - 1
Appendix C – Available Storage Medias	C - 1
Appendix D – Project Gantt Charts	D - 1
Appendix E – Interim Regional Strategy (Draft)	E - 1

Executive Summary

The British Columbia Ministry of Sustainable Resource Management (MSRM) provides provincial leadership for sustainable economic development of the province's land, water and resources. In British Columbia, a high proportion of land, water and associated resources are owned by the Crown and developed by private interests. MSRM has the lead responsibility for access to these natural resources and for the collection of information about them.

In addressing its mandate, MSRM has set five strategic goals, one of which is *“Effective delivery of integrated, science-based land, resource and geographic information.”* Related to this goal, MSRM has set objectives to establish common data and information standards and to reduce the cost of maintaining this information.

The backup, archiving and retention of land-based information are important elements of the Ministry's overall data management system. Two recent studies on hardware capacity in the Ministry concluded that capacity management is impacted by the lack of a Ministry-wide archive policy. As a result of these findings, MSRM engaged KPMG LLP to assist with the development of a Digital Data Backup, Archive and Retention Strategy. This report presents our understanding of the current backup and archive processes, identifies the business requirements for backup, archiving and retention of digital data, presents the strategic direction for establishing consistent approaches and provides a high-level Implementation Plan.

While the development of this strategy focused primarily on project datasets maintained by regional offices, it is important to highlight that the strategy is applicable to all digital data maintained by MSRM.

Backup

For the purposes of developing the backup strategy, we defined backup to be *“copying files to a second medium (a disk or tape) as a precaution in case required files are no longer available on the primary medium.”* In general, we found that MSRM has established reasonably consistent backup processes for its regional and Victoria servers that are effective in reducing the risk of data loss. These processes do vary in accordance with the needs of specific internal clients. We found that the backup processes are not formally documented to ensure consistent application across the Ministry.

Any backup process involves trade-offs between the cost and effectiveness of the process. Increased frequency of backup routines can enhance the effectiveness of the process (e.g., the ability to recover lost data) but at a higher cost. In reviewing the Ministry's backup processes, we noted that trade-offs have been made without full assurance that the processes meet the business requirements. For example, backups are preserved for a maximum of one year. Requests to recover data can only be met if the data was present on a server within the past year. The user community may not fully appreciate this limitation on the ability to recovery data from backups.

Monitoring the Ministry's ability to restore data, as requested by the user community, can validate the appropriateness of the backup processes.

As a result of our findings and conclusions, we have focused our Implementation Plan on the documentation and communication of backup processes and on establishing appropriate procedures to monitor the effectiveness of those processes.

Archiving

In developing the archive and retention strategy, we recognized that there are a number of alternative definitions of "archiving" and that clarity is required. For the purposes of developing an archive and retention strategy for MSRM, we used the following definition "*Digital data archiving is the process of moving semi-active digital data of historical value from on-line to off-line storage media, in order to optimize storage costs, and in a way that preserves the data so that it can be accessed and understood in the future.*"

Our review confirmed that the Ministry does not have a consistent approach to the archiving of digital data. We found that most regions and Victoria branches have some experience with archiving project and other datasets and that there are significant libraries of archived materials located throughout the Ministry. In most cases, we found that current archive practices do not provide reasonable assurance that archived materials can be located, restored and accessed when required at some future date.

It is our conclusion that the Ministry has a significant business requirement to preserve digital data for long periods of time and that the Ministry's ability to meet this requirement is seriously hampered by the lack of a consistent archiving strategy. We have presented a strategy that recognizes the opportunity to decentralize the implementation of archiving procedures while ensuring the process is consistent and effective across the Ministry.

Retention and Disposal

The retention and disposal of archived digital data is a records management issue and should be subject to the Ministry's records management policies and procedures. Currently, all archived materials are unclassified and retained indefinitely. This is equally true for all 'active' data stored online. It is our conclusion that the current archive processes do not ensure that the Ministry can access information from archived digital data and do not ensure that digital data is disposed of after the retention period is complete. In developing the archive and retention strategy, we focused on ensuring the Ministry is able to apply records management practices to archived digital data and to ensure that it will be able to access information from that archived material throughout the retention period.

Scope and Objectives

Background

The British Columbia Ministry of Sustainable Resource Management (MSRM) recently completed two studies, the Regional Data Warehouse Capacity Review and Interim Data Warehouse Headquarters Capacity Review, which concluded that:

- There is no province-wide backup, archiving and retention policy for MSRM GIS project data;
- Total regional server capacity is affected by the absence of a digital data archival system - data volumes can be reduced by up to 30% with formal data retention and archival policies and systems in place;
- Not all regions are technically capable of successful data backup and recovery; and,
- Digital data backup and archive practices are in place in Victoria but large volumes of project-oriented data require an inordinate amount of disk storage space.

Both reviews identified the need for the Ministry to develop a Digital Data Archive Strategy.

Currently, the backup process for regional UNIX servers uses Legato Networker backup software running under UNIX. Data is written to DLTtape drives using hardware data compression to store approximately 55 gigabytes of data on each DLT type IV cartridge tape.

Most regional sites utilize a manual archive process to copy working data for completed projects from disk to tape or CD.

The Information Management Branch (IMB), in conjunction with the Corporate Services Division (CSD), is planning to implement a new technical infrastructure using Veritas NetBackup software to perform the backups and Veritas Global Data Manager to remotely control backup operations at regional sites. Existing DLTtape drives are planned to be moved from the UNIX servers to Windows 2000 servers. The UNIX backups will then be executed using a network connection between the Windows 2000 servers and the UNIX servers and will be controlled through a backup client program running on UNIX.

The Digital Data Archive Strategy must conform to and complement these province-wide MSRM technical upgrades.

Purpose and Objectives

The development of the Ministry-wide digital data backup, archive and retention strategy is intended to:

- Differentiate between digital data backup and archive processes;
- Deliver a Ministry-wide digital data backup, retention and archive strategy;
- Provide the operational framework for this strategy;
- Optimise the efficient use of regional and Victoria server capacity;
- Promote the consistent use within MSRM of available archive and recovery capability; and,
- Rationalize the data archive and recovery process to meet specific MSRM data retention requirements.

The process should:

- Evaluate existing digital data archive processes;
- Provide for the optimal deployment of Ministry backup/archive technological capabilities;
- Define local and centralized backup and archive processes for project datasets in co-operation with Victoria and regional data managers and system officers; and,
- Document a Ministry-wide Digital Data Backup, Archive and Retention Strategy framework and Implementation Plan.

The process for archived Geographical Information System (GIS) project data is to:

- Define the life-cycle of project datasets such as Land Use Plans;
- Establish criteria for project dataset retention and disposal;
- Identify relevant project datasets;
- Identify previously archived project data sets; and,
- Determine regional and Victoria requirements for acceptable project dataset recovery times and processes in co-operation with Victoria and regional data managers and systems officers.

The project scope did not include the selection or implementation of backup/archive software or hardware and was constrained by the archive/backup products already acquired. As well, the project was not asked to address the Ministry's needs for business continuity or disaster recovery plans.

Current Environment

Digital Data Backup Processes

In this section, we identify our understanding of the current digital data backup and recovery processes used by the Ministry. We recognize that there are some variances in the processes to meet the specific needs of Ministry business units that may not be fully reflected in this section.

Definition

The backing up of digital data can be defined as “**copying files to a second medium (a disk or tape) as a precaution in case required files are no longer available on the primary medium.**” The key characteristic of the digital data backup process is that data is stored in at least two locations in such a way that the impact of losing the primary source is adequately mitigated.

Backup involves saving all digital data (“official” records and “transitory” records) and metadata, and retaining that data for as long as there is an operational requirement.

Backup Processes

- Some Ministry data is located on hardware operated by Common IT Services (CITS). Backup procedures for this data are subject to standards and procedures adopted by CITS. These procedures are assumed to take into account the Ministry’s requirements for backups.
- The backup of Regional Office and Victoria servers is the responsibility of systems staff in the Corporate Services Division (CSD), with some procedures delegated to on-site systems staff, and are performed relatively consistently as follows:
 - **Daily.** Daily tapes are retained for one month, from the beginning to the end of the month, when monthly backups are performed. Daily tapes are recycled on a monthly basis. Daily tapes are retained on-site where the backups occur
 - **Weekly.** Weekly tapes are retained for one month, from the beginning to the end of the month, when monthly backups are performed. Weekly tapes are recycled on a monthly basis. Weekly tapes are retained on-site where the backups occur; and,

-
- **Monthly.** Monthly tapes are retained for 12 months and are sent offsite to FACS for secure storage. Monthly tapes are recycled on an annual basis.
 - In some limited cases, Ministry servers are not under the control of CSD staff. In these cases, backup processes may vary from the standard processes adopted by CSD staff. It is the intention of CSD staff to bring all Ministry servers under their central management.
 - Two different backup software technologies are currently in use:
 - *Legato Networker* is used for UNIX backups at Victoria and regional locations; and,
 - Veritas Backup Exec is used for NT/2000 backups at Victoria and regional locations.

There is a plan to replace the Legato backup infrastructure with Veritas NetBackup, but there have been problems in the implementation of this enterprise solution, which have delayed the migration of all backup infrastructure to Veritas.

- Oracle database backups (on Ministry servers) consist of daily exports and weekly cold backups. These exports and cold-backups allow for recovery to the end of day for the complete database (or schema) that is being exported or cold-backed up. This does not allow for point-in-time recoveries of data that may be modified or deleted during the day. The daily and weekly backups for Oracle databases are managed in a fashion consistent with the daily, weekly and monthly tape cycles. Hot backups (allowing point-in-time recovery) of Oracle databases may be occurring, but they are not part of the standard backup and recovery procedures for MSRM.
- Regional System Officers (RSOs) are responsible for:
 - Ensuring that the backups are successfully completed (this is only valid in locations where the RSO's currently manage the backups);
 - Changing tapes;
 - Sending tapes for offsite storage;
 - Requesting tape retrieval from offsite storage for recoveries; and,
 - Performing data recoveries. This is only valid in locations where the RSO's currently manage the backups.

Recovery Processes

- There is an average of one to two recovery procedures per month per regional office.
- Victoria branches report that recovery from backup tapes is an infrequent or rare event, but does occur from time-to-time.
- Based on the current backup schedule and tape rotation strategy, the following recovery scenarios are valid:
 - Recoveries from the end of each day are available for between 2 and 4 weeks from the date of the daily backup;
 - Recoveries for the end of week are available for 1 month from the date of the weekly backup; and,
 - Recoveries older than 1 month are only available at an end-month basis. This means that a file that was modified or deleted 6 weeks ago is only available for a recovery from the monthly backup created prior to this date. The file may or may not be available on the previous monthly backup tape.
- Where backup procedures vary from the Ministry's standard, recovery processes will also vary.

Observations/Issues

With respect to the backup and recovery processes, we identified the following issues:

- There is a perception that even though backup and recovery procedures are in place, they are not well documented and/or communicated to the Regional Offices.
- There are no Service Level Agreements between regional offices and Victoria pertaining to backup and recovery functions. This has led to a misunderstanding of the expectations between regional and Victoria offices.
- A number of regional offices are regularly reaching the limit for tape capacity during backup procedures. Some sites have had to change the tape for their weekly backup on Monday morning so that it would complete; meaning that the backup would not complete until lunchtime on Monday.
- The rollout of centralized Veritas backups to regional offices has not been completely successful. A number of sites have had to revert back to their original backup software (i.e. Legato) and procedures.

-
- The Oracle Agent for Veritas is not currently being utilized to assist with Hot Backups of Oracle databases.
 - Database growth at Victoria branches is affecting backup windows. Both export and cold backup processes are sometimes exceeding the backup window time and/or tape space.
 - There is no currently-maintained corporate Disaster Recovery Plan (DRP) or Business Continuity Plan (BCP) that would assist with the recovery of backed up data in the event that corporate and/or regional servers were not available following a disaster. Data maintained on CITS hardware is subject to that organization's DRP.
 - Some MSRM data resides on hardware maintained by the Ministry of Forests. This data is subject to the backup standards and processes of the Ministry of Forests.
 - Database recovery of deleted or purged data is difficult due to the fact that separate recovery schemas may be required to be built in order to properly retrieve the deleted data. Disk space issues arise that may make it difficult to create these backup schemas.
 - There is no definition as to when tapes are retired. Currently, they are recycled based on the schedule described above, but there is no mechanism in place to determine when the recycled tapes should be replaced. We did not identify any issues relating to the failure of older tape media used for backups, but this may be an issue that should be addressed to ensure that tapes are not overwritten until the point of failure.
 - An ongoing issue that will need to be addressed is to ensure that all backup tapes remain readable by the current backup hardware and software. One item that was noted during interviews was that data backed up from a Novell system was not easily recoverable when the Novell systems were replaced. This issue will arise again if the Legato infrastructure is replaced with Veritas. A procedure must be put in place to ensure that data from monthly backup tapes are still recoverable.
 - The size of the backups and the time allotted to perform the backups will continue to be an issue for those offices with hardware and software that are not currently meeting their space/time requirements.
 - The Ministry is aware that some 'static' data replicated to the data warehouse is also routinely backed up. There is an opportunity for the Ministry to fine-tune its backup processes to reduce redundancy.

In summary, the current backup and recovery processes are generally consistent and generally meet the needs for all regional and Victoria offices. There is, however, a need to fully document and communicate the backup and recovery procedures to all Ministry offices.

Digital Data Archive Processes

Definition

There are numerous definitions of the word “archive”.

- As a noun, an archive is “a place in which public records or historical documents are preserved, or the material preserved.”
- As a verb, archiving means “to file or collect (as records or documents) in or as if in an archive”.
- Digital data archiving involves selective preservation of electronic information that has long-term value (“official” records) over an extended period of time. In the British Columbia government context, “archiving” does not occur until the Ministry records of long-term or permanent value are transferred to BC Archives.
- The BC Ministry of Forests defines information archiving as “the storing of valuable historical information in an organized manner so it can be easily accessed and understood in the future.”¹
- Archiving is the process of systematically making copies of your most current data and storing those copies in a safe, off-site location where they can be readily accessed if needed for restoration. Backup tapes should be regularly and quickly sent to an off-site archive for protection. Archiving ensures a chance of recovery from a data disaster if on-site copies of data are lost. Also, archiving is used for long-term data retention to safeguard operational needs and meet legal or regulatory requirements.²

¹ <http://www.for.gov.bc.ca/isb/datadmin/infoarc.htm>.

² <http://www.dlftape.com/Practices/Archive/Default.htm#>.

-
- Archiving and backup are two distinct processes. Backing up is essential to systems management in that it involves creating copies of hard disk files on tape (or other storage media) for ready retrieval, if needed. Archiving involves a detailed system for securely and systematically storing files in a safe, climate-controlled alternate location. It also includes a detailed cataloguing and retrieval process for data recovery and restoration when needed and takes into account media migration issues so that legacy data remains accessible as media device standards change over time.

For the purposes of defining a digital data archive strategy for the Ministry, we adopted the following definition:

“Digital data archiving is the process of moving semi-active digital data of historical value from on-line to off-line storage media, in order to optimize storage costs, and in a way that preserves the data so that it can be accessed and understood in the future.”

The key characteristic of this definition of digital data archiving is that the original data is removed from on-line storage and is only available in off-line storage media. Archived data is considered to be of long-term or permanent value to the Ministry.

Archive Processes

Data archiving within the Ministry is heavily oriented around datasets, with particular emphasis on large project datasets (e.g. Sustainable Resource Management Plan (SRMP) datasets defined at the Landscape Unit level). Archiving a project’s datasets to off-line media (e.g., CD-ROM or tape) is the principal method of archiving. Archiving typically occurs at the end of a project lifecycle, but moving a snapshot of project files to off-line media at certain stages of a project (a point-in-time archive) is also common.

Current digital data archiving processes vary between regions, and within regions. Where a process is in place, it has some or all of the following components:

- Software is used to move data to off-line storage media (Legato Networker, Legato Archiver, Veritas Backup-Exec, Unix TAR command and CD burning tools were identified);
- Off-line storage media is used for archived data (CD-ROM, DLTtape, 8 mm tape and other tape formats were identified);
- Hardware used for backups is typically used for archiving to DLTtape. CD-Burners are used to move files to CD-ROMs;

-
- One region places all completed project datasets into a standard Unix directory (ArcDone) as an indication that the data is available for archiving.
 - Data is typically archived after a full backup of the data has occurred through normal backup processes;
 - Archived data is typically stored on only one media (all on CD-ROM, or all on DLTtape);
 - Most regions create only one copy of the archives. Two regions create two copies of the archives;
 - The storage media are physically stored in a variety of locations:
 - On-site:
 - Fire-resistant safes are used by many regions;
 - Cabinets; and,
 - Desks.
 - Off-site – Local:
 - Ministry buildings;
 - File storage facility;
 - Employee homes; and,
 - Contractor homes.
 - Off-site – Remote (where 2 copies are archived, the second copy is typically sent to a remote storage facility):
 - File storage facility in Victoria
 - Information about what has been archived is identified and tracked by the majority of regions (e.g. what is on the archive, where the archive is stored, a regional project number);
 - A variety of tracking systems are used to store the information about archived datasets, including:
 - MS Access databases;
 - Spreadsheets; and,
 - Unix files.
 - In one region, application software used to read the archived data is sometimes archived along with the datasets;

-
- Recovery of information from archived datasets is generally a rare occurrence;
 - The Registries and Resource Information Division—Sub-Surface Tenure and Mineral Titles Branch has significant experience with archiving digital maps onto microfilm. The Branch has been archiving images onto microfilm since 1966. Prints of maps are produced and digitized on a tablet. The digital image is then copied to microfilm for preservation purposes. Microfilm has been adopted as an archiving media over digital media due to the durability and readability of microfilm as compared to other media. The archived material is retained to respond to queries concerning sub-surface tenures and to provide as evidence in legal proceedings should they arise.
 - The Sub-Surface Tenure and Mineral Titles Branch saves a digital snapshot of a subset of their attribute database on a monthly basis. As well, the Branch saves a digital snapshot of all the titles reference maps on an annual basis.
 - Recovery of information for registries datasets occurs relatively frequently as this historic information is sometimes needed to resolve issues with respect to current title.

Observations/Issues

- There are several business drivers within the Ministry for archiving data, including a desire to gain operational savings by moving data to lower cost media, legislative requirements for the preservation of data, and requirements to maintain evidence suitable for use in litigation.
- While legislative requirements are provided through provincial acts, such as the *Document Disposal Act*, no interpretation or further direction regarding archiving could be identified in Ministry policies, standards or procedures. To varying degrees the Ministry's regional offices have developed their own procedures for archiving but these are inconsistent between regions, and sometimes within a region.
- Estimates of the amount of online data that the regions could archive varied from 0% to 25% and two regions provided estimates of the amount of data they had archived-to-date (200GB and 250GB). The primary opportunity exists within the Unix server environment, as this is where most project datasets are stored. The opportunity to archive digital data in the Windows NT/2000 environment is considered to be of secondary importance.

-
- Tracking of the archiving process is limited and uncoordinated. There is insufficient metadata available to determine how much data has been archived within the Ministry, whether it could be accessed if required, or if the current procedures are cost effective.
 - Some problems have been encountered in the past with recovering data from archive media, including difficulties reading older tapes (e.g. 8 mm format), corrupted CD-ROMs and the inability to locate archived materials.
 - A primary concern for the preservation of digital data is the loss of information due to technological obsolescence of hardware and software, and the fragile medium, such as CD-ROMs or magnetic tape, that digital data is stored on. Recovery of the data requires the hardware, operating systems and backup software to restore the data. Viewing information in digital form requires a context, the hardware, operating systems and application software needed to view the data. The ability to access information from datasets is an issue for both active and semi-active data. While archiving requires preservation of the bits and bytes, it is equally important to preserve the ability to access and read the data. The fact that active data resides on-line media does not guarantee that the Ministry will retain the technical ability to read the data. As semi-active data tends to be older, the risk that the Ministry will not be able to access information from these datasets is inherently higher than for active data.

In preserving the ability to read digital information, the Ministry has two primary options:

- It can preserve the digital data and the context in which it was created. To preserve the context, the Ministry must preserve the hardware, operating systems and application software with which the digital data was created; and/or,
- It can migrate digital data from one hardware/software configuration to another, or from one generation of information technology to subsequent generations.

In most cases, the first option is not feasible for long-term archives as the cost and technical challenge of maintaining obsolete hardware and software is prohibitive. In the second option, we must face the challenge of maintaining data integrity as data is migrated from one context to another.

A third option for consideration would be to store the data in a non-proprietary format rather than its “native” proprietary format. This option creates a risk that data integrity may be affected by changing the data format. As discussed later, it may be more appropriate to save the archived files in both “native” and “non-proprietary” formats.

- Archiving digital data to non-digital media (e.g. microfilm) limits the ability to further digitally process the data.

Project Datasets

MSRM maintains a wide range of digital data stored on production servers distributed amongst the Ministry’s regional and Victoria offices. The majority of this digital data is routinely backed up to reduce the potential for data loss. The extent of digital data stored on-line directly impacts the storage requirements and, ultimately, the cost of the information technology environment.

Digital data exists to support a range of business needs. Much of the Ministry’s digital data exists to support production applications, data warehouses and e-service websites. In most cases, these types of digital data must be stored on-line to support continuous business processes.

Digital data also exists to support a variety of projects undertaken by the Ministry. Project datasets can include spatial data (including datasets from the Terrestrial Resource Information Mapping (TRIM) system), attribute data, digital images, presentations and reports. Project datasets are considered to be operational datasets and, therefore, are generally subject to the requirements of ORCS. A key characteristic of operational records is that the specific value of the information must be assessed in defining the retention and disposition schedule. Administrative records subject to ARCS are generally subject to a standard, pre-defined retention and disposition schedule.

These project datasets have a number of common characteristics:

- The datasets are typically large in size (from 10 megabytes to 50 gigabytes);
- Most project datasets are stored on UNIX servers. Some project datasets are stored on Windows/NT/2000 servers;
- Current spatial datasets are created using a small range of GIS applications (e.g., ESRI’s ArcInfo V8.02, MicroStation, MapInfo, etc.);
- Attribute data is generally stored in Oracle relational databases;

-
- Project datasets represent the Ministry's best information about land-related resources at a point-in-time and are used to support the development of important land use plans. With respect to such plans, there is an ongoing potential for public queries with respect to this digital data and the potential for litigation concerning the planning decisions made. Accordingly, project datasets typically have long-term value to government. The datasets contain the underlying information upon which important and legally binding land use decisions are made, such as the establishment of bio-diversity targets. Established plans consist of core or underlying base data, such as TRIM and forest cover, as well as resultant planning data from multiple resource analyses.
 - There may be points during the project lifecycle when the project team determines that a "snapshot" of the data is required to preserve an historical record of the data;
 - As projects by definition have a defined start and finish date, there is typically a point where the datasets are complete and remain static thereafter; and,
 - Once the related project is completed, there can be a long periods of time during which access to the digital data is not required. This creates the potential to archive this digital data and remove it from on-line storage.

Types of Projects

In developing the strategy, we attempted to identify the types of projects that generate significant datasets. Based upon our interviews with regional and Victoria staff, we identified the following types of projects that generate large datasets:

- Land Resource Management Plans (LRMPs);
 - Involve very large datasets;
 - Developed over multiple years;
 - Relate to a Timber Supply Area (TSA);
 - Few in number;
 - Cover a ten-year planning horizon;
- Sustainable Resource Management Plans (SRMPs) – these replace Land Use Plans (LUPs) but planning is still performed at the landscape unit (LU) level of information:
 - Prepared subsequently to a LRMP;
 - Multiple landscape units in each TSA;

- There are a total of 1,248 landscape units defined in the province, each varying in size.
- Landscape Level Plans are now being used which aggregate landscape units (maximum of 130 in the province);
- Timber Supply Reviews (TSRs):
 - Prepared for each TSA and TFL;
 - Set five-year limits for Allowable Annual Cut (AAC);
- Wildlife Strategic Plans/Habitat Analyses and Studies;
- Spatial Analyses;
- Site Selection Studies (occasionally performed from other government organizations);
- Coastal Planning;
- Treaty Negotiations Support;
- Sectoral Analysis;
- Pre-Tenure Planning—planning processes related to Oil, Gas and other resource extraction processes; and,
- Forest Development Plans (expected to evolve into ‘Sustainable Resource Management Plans’).

Other Datasets that could be archived

During our review, we identified other types of datasets that are not associated with a project that may require archiving, including:

- Production system datasets—datasets such as forest cover and Titles registries are maintained on a continuous basis and, unlike project datasets, are never considered to be complete. However, there are times when Ministry staff wish to or must retain a “point-in-time” version of these spatial datasets to preserve the history of the datasets. Other data that may be used as a point-in-time for planning purposes are Terrestrial Ecosystem Mapping (TEM) and Predictive Ecosystem Mapping (PEM) data. TEM provides detailed habitat information such as soil, flora and fauna collected locally by ecologists and wildlife biologists. PEM is resultant data from habitat modelling using TRIM digital elevation modelling, forest or vegetative cover, biogeoclimatic sub-zone and variant, habitat units and other terrestrial data; and,

-
- Land Tenure Registries—MSRM is responsible for maintaining information about a wide-range of land tenures (e.g., forest tenures, mineral tenures, water tenures, crown land tenures, etc.). Some of the tenure records maintained go back more than 100 years and it is critical for the Ministry to be able to determine when a change occurred that affected title. As the registries typically retain only the current tenure information, “point-in-time” snapshots of registry information is required to be archived in order to preserve the historical record. Given the long-term nature of the archive, the process must ensure longevity, security and ease of access.
 - Digital Image Management (DIM)—The Base Mapping and Geomatic Services Branch maintains a large database of digital images (primarily aerial photographs) that are stored online. The total database of images currently requires more than 9 terabytes of storage space and the size of the database is more than doubling each year. While the Branch is using online and near-line storage, there may be an opportunity to store some older images on off-line media.
 - Application Consolidation—A project is underway in MSRM to consolidate applications and their related data. Prior to consolidation, there will likely be a requirement to archive copies of both application code and related data.
 - Other Organizations—MSRM provides computing resources for other government organizations (e.g., Ministry of Water, Land and Air Protection, Ministry of Forests, Water and Land BC Inc., Oil and Gas Commission, etc.). In general, MSRM provides file storage services and are not directly responsible for the data. Service Level Agreements and Memorandums of Understanding have been drafted to document the service relationship. These agreements have addressed the need for MSRM to provide backup service for these datasets. No provision has been made for archiving of these datasets.

Project Lifecycles

Our review has found that while all regions work on the same types of projects, the timing of this work varies greatly. The timing of planning processes appears to be region-specific and the duration of planning projects can also vary widely. While the timing of projects varies, we did find that each project has a unique lifecycle that includes a number of common phases. Land use plans (e.g. SRMPs) often include the following phases:

- **Boundary Definition**—the area for which a plan is being developed is defined;

- **Data Collation**—available data concerning the plan area is gathered;
- **Data Analysis**—the available data is analyzed to support the setting of planning objectives and targets;
- **Data Cleanup and Update**—data may be modified or added to in order to provide a complete basis for defining planning objectives and targets;
- **Set Draft Objectives**—draft objectives and targets are documented and provided to the public for review and comment;
- **Revision**—data and plans are revised based upon comments received;
- **Finalize the Plan**—the final plan is documented and publicly released (SRMP establishment through the Regional Directors SRM); and,
- **Monitoring**—periodic comparison of plan targets against actual results is likely to become a more important activity in the future.

For each project, there is the potential to archive project datasets after the associated plan has been finalized. In addition, there may be milestone points during a project where a “point-in-time” version of the datasets is required for preservation.

From an archivist’s perspective, records typically have three life-cycle phases:

- **Active**—frequent access is required;
- **Semi-active**—infrequent access is required; and,
- **Final disposition**—access is no longer required, records are destroyed or transferred to archives.

Project data that is ‘semi-active’ should be available to be archived to off-line storage. Project data that is in ‘final disposition’ should not be stored at the Ministry.

Archive Retention

Across the Ministry, we found a common retention period ascribed to all archived digital data, “permanent retention”. There is no formal process in place for classifying digital data or assigning retention and disposal schedules. By default, because of the cost of collection, the potential need for scenario re-creation, and long-term temporal value, all archived datasets are preserved permanently.

Business Requirements

The management of digital data is an integral part of efficient and effective business processes. Without effective management of digital data, an organization can expend resources unnecessarily in attempting to locate or recreate required information. As well, storage costs for digital data can be reduced by the appropriate application of archiving and disposal processes.

Recorded information is a primary means by which government can be held accountable for its actions, and as such, information in all forms must be clearly identifiable, readily accessible and properly maintained for the duration of its life cycle.

Digital Data Backup Requirements

Detailed below are a list of the business requirements that drive the backup and recovery infrastructure and procedures for the Ministry.

Mandatory Requirements

- All production data (attribute and spatial databases, documents, profiles, file systems, etc.) must be backed up on a daily basis. These daily backups must be retained for a minimum of one month;
- All project datasets must be incrementally backed up on a daily basis after it has been moved into a production directory. These project datasets will be managed as part of the daily production file system backup described above;
- Daily and weekly backups must be stored locally. Depending upon the tape backup hardware, the tapes should be rotated out on a daily basis into secure, fire protected storage;
- Local (daily and weekly) backup data must be recoverable within one business day of the request;
- Monthly backups will be shipped for offsite storage at FACS in Victoria;
- Offsite backup data must be recoverable within a 72-hour period (some users may prefer a faster response time and this may be considered in a client-specific Service Level Agreement);
- Ministry-wide policies and procedures should exist for backup and recovery (this will not include data stored at CITS or the Ministry of Forests);

-
- Service Level Agreements should be negotiated with the MSRM client base. These Service Level Agreements need to identify roles and responsibilities, response times, recovery requirements; and,
 - There needs to be sufficient tape space in all locations to support full, incremental and/or differential backups for all production data.

Optional Requirements

- A tracking system should be available to properly track and manage the daily, weekly and monthly tapes for MSRM;
- Backup schedules should be synchronized for both NT and Unix backups;
- The backup of redundant data across multiple servers should be minimized;
- A single backup and recovery software infrastructure should be implemented for both Unix and NT architectures across MSRM; and,
- A consistent, manageable backup and recovery infrastructure should be implemented for MSRM. The goal for this infrastructure will be to move towards centralized management of backup and recoveries.

Digital Data Archive Requirements

Based upon our understanding of the Ministry's responsibilities concerning the preservation of archived digital data, we identified the following requirements:

Mandatory Requirements

- There is a requirement for a Ministry policy and procedure framework to provide direction on the archiving of data;
- There is a requirement to document and assign roles and responsibilities for implementation of archiving processes;
- There is a need to preserve historical information on land use and land tenures, generally on a permanent basis;
- There is a business requirement to ensure that archiving processes are cost effective. The archiving of data is primarily a mechanism for ensuring the efficient use of resources and should provide costs savings or cost avoidance. All capital and operational costs associated with the data must be considered including hardware, software, and ongoing operational costs such as backups and virus scanning. The cost and likelihood of having to recover the data must also be considered;

- When data is archived, sufficient metadata must be recorded in a central location to manage that archive, including:
 - A record of what data is on the archive and how it is stored;
 - A record of the data custodian/organizational unit responsible for the archived data;
 - Identification of where the archive is physically located;
 - Identification of how to recover data from the archive (including the required hardware, operating system, software for retrieving datasets from the archive and software for processing the datasets); and,
 - Identification of the classification of the archived data for purposes of assigning the retention and disposal schedule for that archive.
- Metadata requirements for spatial and attribute data need to include other data attributes unique to the data format. For example, a project is currently underway in the Ministry to consider the metadata requirements for spatial data. The project has initially identified over 100 different metadata requirements.
- Metadata information should be stored with the archived data.
- A minimum of two copies of the data should be archived. Each physical copy should be stored in separate locations.
- Archived digital data must either be stored in a manner that is directly usable during the retention period, or there are procedures, hardware and software to copy, restore, convert, transform, or otherwise process the archive into a form that is usable.
- There is a requirement to store archive media at constant, lower temperatures and humidity levels to maximize longevity. Proper handling and storage of magnetic tapes is critical.
- There is a requirement to periodically migrate archived data to new media as the original media deteriorates or as the hardware or software required to process the archive is decommissioned. For practical data permanence, some vendors recommend refreshing the archive media every five years. This can avoid common problems such as media deterioration. It also allows an organization to stay within the approximately six-year window for media generation changes.
- The Ministry should be able to restore archived data to on-line storage within 24 hours of a request to restore. The requirements for response times were generally identified as 1 to 3 days although several regions indicated that “urgent” recoveries were required at times.

- The generation of backups or the copying of data for distribution is driven by other factors and is not considered “archiving” for the purpose of this project.

Optional Requirements

In addition to the mandatory requirements identified above, we identified several requirements that would enhance the archival process:

- There is a strong interest among regional staff for the data recovery capability to reside locally.
- Digital data can be archived to multiple media types (e.g., CD-ROM, DLTtape, 8 mm tape, etc.). Multiple copies of data on a single piece of media leaves that data at risk. Data availability is enhanced through media redundancy on reliable media.
- Digital data should not be compressed when it is archived. Data compression can optimize storage capacity but can impact the integrity of data (if the compression is not ‘lossless’). In addition, compressed data must be capable of being uncompressed when restored. Hardware compression is particularly limiting in that the decompression can only occur using that specific hardware.
- There is a requirement for a ‘**preservation master**’ to be included in the digital data archive. Digital data is generally stored in proprietary formats that restrict the range of software that can read the data. Most data formats can be converted to non-proprietary formats that broaden the range of software that can read the data. By preserving the data in both its native format and a non-proprietary format (the preservation master), the Ministry can increase the ability to access information in the future.

An example of a non-proprietary format is the Spatial Archive Interchange Format (SAIF)³, a Canadian standard, developed to store GIS data in an interchangeable format that can be used to create versions in a wide range of proprietary formats.

- The Ministry should consider whether the restore process should include a “priority restore service” in the archive process to meet urgent requests for restoration of archived data. These rare requests can likely be met if the ability to restore locally is maintained.

³ <http://home.gdbc.gov.bc.ca/fmebc/>.

Digital Data Backup, Retention and Archive Strategy

Backup Strategy

Vision

The vision for the Ministry's Digital Data Backup Strategy is as follows:

By September 1, 2002, MSRM will have in place Ministry-wide policies and procedures that support a consistent and effective process for the backup of digital data that meets the Ministry's requirements to protect valuable information from damage or loss.

Strategic Objectives

The objectives of the backup strategy are as follows:

To provide consistent, documented and robust procedures and infrastructure that will ensure the timely recovery of all production data and project datasets.

To maintain the metadata concerning digital backup in an accessible form, sufficient to identify the existence of the data, the location of the data and the media format.

To monitor the effectiveness of the backup and recovery procedures and infrastructure.

Scope

The backup strategy is applicable to all digital data stored on Ministry-managed file and database servers. The strategy is not applicable to Ministry data stored on equipment managed by CITS or other government organizations.

Critical Success Factors

In order to realize the strategic objectives, the Ministry must achieve the following critical success factors:

- Backup data must be maintained in a way that protects it from loss, damage, degradation, loss of information, and other threats to its physical integrity:
 - Backup data must be maintained on media and in formats that ensure it is readable and accessible for the duration of the retention period; and,

- Backup data must be housed in environmental conditions that meet its preservation, retrieval and security requirements;
- The complete metadata of the backup data must be preserved for the entire retention period for the backup:
 - Metadata must be linked to backup data in a way that ensures that it is identifiable and recoverable, and the context of its creation and use is maintained; and,
 - Backup data that is moved to different media must maintain its context and authenticity.
- Backup data must be recoverable in the period specified in service level agreements:
 - Daily and weekly backups must be recoverable within one business day;
 - Monthly backups must be recoverable in at least 72 hours (unless other timeframes are defined in Service Level Agreements); and,
 - Recovery times, number of recovery requests and the success/failure of all recovery requests must be tracked. This data will be used to better define the business requirements for backup and recovery.
- A full backup must be completed in the specified backup window in an unattended fashion.
- Recovery of end-of-day backups will be available for a minimum of one month.
- Recovery requests for data older than one month will be available for month-end backups only.

Key Performance Indicators

In order to assess the ongoing effectiveness of the Ministry's digital data backup strategy, the following performance indicators should be adopted. Targets should be established for each performance indicator and actual performance should be monitored and reported against the targets:

- **% of backups performed in backup window**—the percentage of backups that are successfully completed during the specified backup window is an indication of the effectiveness of the backup process.
- **% of successful restores**—The ultimate success of the digital data backup process is dependent on data being successfully restored from backed up media when required. The percentage of restoration requests that are successfully met should be tracked and reported.

Archive Strategy

Vision

The vision for the Ministry's Digital Data Archive Strategy is as follows:

By December 1, 2002, MSRSM will have in place Ministry-wide policies and procedures that support a consistent and effective process for the archiving of digital data that meets the Ministry's requirements for the identification and preservation of information with long-term value. Archive processes will be separate and distinct from backup processes to ensure neither process is disrupted or impacted by the other.

By January 1, 2003, all archived digital data in the possession of the Ministry will adhere to the Ministry's archive policy.

Responsibility for the maintenance of the Ministry's archive policy and procedures will reside centrally with the Information Management Branch. Responsibility for initiation of archiving procedures will be decentralized to data custodians and on-site systems staff. Technical support can be provided by either a centralized or decentralized service model.

The following exhibit highlights the key steps in the proposed archiving process:

Exhibit I – Proposed Archiving Process

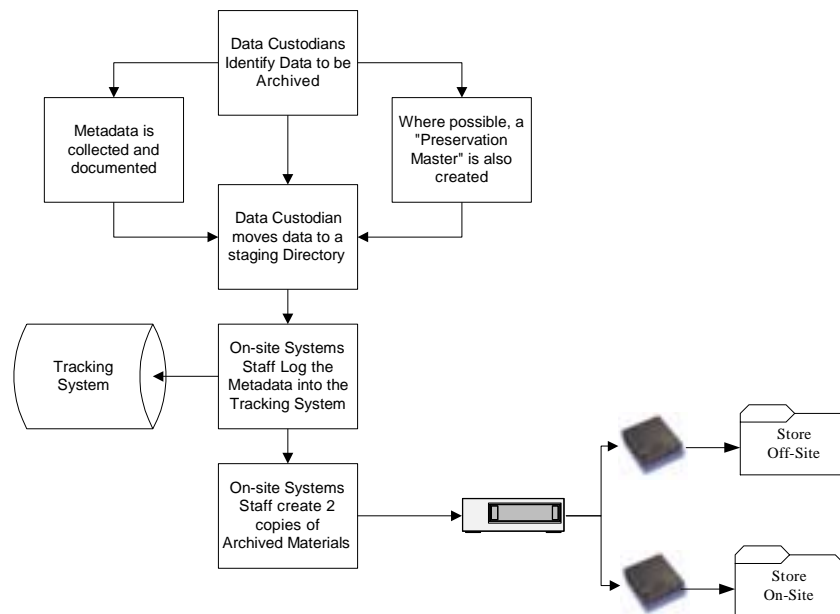
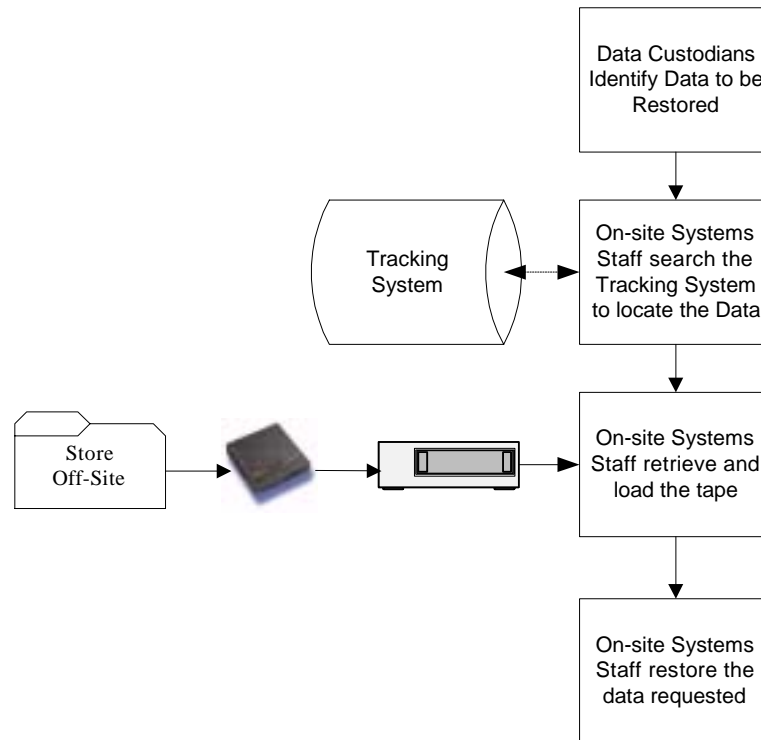


Exhibit II highlights the key steps in the proposed process for restoring data from archive.

Exhibit II – Proposed Archive Restoration Process



Strategic Objectives

The objectives of the archive strategy are stated as follows:

To maximize the movement of semi-active digital data from on-line storage to off-line storage in order to rationalize the Ministry’s storage costs.

To maintain metadata concerning digital data archives in an accessible form, sufficient to identify the data stored, the custodian/organizational unit storing the data, the physical location, the context required to read the data and its classification for retention and disposition.

To preserve the digital data archive in a form that ensures it can be copied, restored, converted, transformed, or otherwise processed into a usable form.

Scope

The archive strategy is applicable to all digital data initially stored on Ministry-managed file and database servers whose status changes to “semi-active”. This includes digital data in all formats, including: spatial data; attribute data; data warehouses; word processing documents; spreadsheets; etc.

Critical Success Factors

In order to realize the objectives of the archive strategy, the Ministry must achieve the following critical success factors:

- An overall media and data management infrastructure. The very creation of a digital archive creates an accompanying custodial mandate for proper long-term management.
- A clearly defined, consistently applied digital data archiving policy.
- General awareness within the Ministry of the benefits of archiving and a champion for the use of digital archiving.
- Clear roles and responsibilities for identifying, classifying, archiving, storing, restoring and disposing of archived digital data. Including:
 - Identifying the opportunity to archive digital data;
 - Requesting an archive;
 - Generating the archive;
 - Maintaining the archive, including identifying factors that could affect the availability of an archive such as hardware or software obsolescence;
 - Establishing a security matrix to define who can request a restoration from an archive;
 - Identifying the need to recover data from an archive;
 - Requesting the recovery of data from an archive;
 - Recovering data from the archive;
 - Identifying changes in the metadata associated with an archive, such as a change to the date for disposal; and,
 - Dealing with old archives or archives coming from other organizations that do not follow the Ministry’s standards.
- Clear identification of the Data Custodians for each business area;

-
- The recording and maintenance of sufficient metadata about archives including:
 - A standard identifier that uniquely identifies the ministry and the archive within the ministry;
 - The quantity of data archived, the physical media containing the archive and where that media is stored;
 - What data will be on the archive and any dependencies on data not part of that archive;
 - The classification of the archived data identifying its retention and disposal schedule;
 - How the data should be disposed of after the retention period (destroyed or transferred to another specified party);
 - The procedure to recover the archived datasets and what hardware, operating system and archiving software is required to recover those datasets; and,
 - What tools are required to use the datasets being archived, (e.g. ArcInfo on AIX 4.1, Microsoft Access 97 on Windows NT).
 - A mechanism for identifying potential changes to the technical environment (e.g. hardware, archiving software, database software) and dealing with those potential changes.
 - Evaluating the impacts of the changes on the archive and recovery process including backward compatibility of the products, the cost of maintaining a copy of the current environment that is capable of recovering and accessing the data currently archived and the cost of converting existing archives into a format that can be recovered by the new technical environments.
 - Supporting mechanisms to effectively and efficiently create, maintain, restore from, and delete archives, including:
 - Defined workflow and technical procedures for the creation, maintenance, recovery and deletion of archives;
 - Sufficient hardware and other resources to perform archiving functions in the time frames required by the users of the archive process;
 - A database of all information required to make use of the archives; and,

- Ongoing procedures required for maintaining the archive media (e.g. copy data to a new tape ever 5 years).
- The development and monitoring of Service Level Agreements to enable management and optimisation of the archiving process.

Key Performance Indicators

In order to assess the ongoing effectiveness of the Ministry's digital data archive strategy, the following performance indicators should be adopted. Targets should be established for each performance indicator and actual performance should be monitored and reported against the targets:

- **Total size of the Digital Data Archive**—the direct benefit of archiving digital data is to reduce the storage capacity of the Ministry's server infrastructure. The total size of the archive can be used to estimate the ongoing benefit associated with reduced server capacity requirements.
- **Volume of data archived and number of new archives created over the year**—archiving is most effective as an ongoing process, and not as an emergency process when disk capacity is exceeded.
- **Percentage of Archived Digital Data that adheres to Archive Standard**—once an archive standard has been adopted, it will be important to migrate all existing archived data to this standard. It will also be important to minimize instances where data is archived but still present on on-line storage.
- **Number of restores**—As one of the justifications for archiving data is that access to semi-active data is limited, this justification should be validated through tracking of the number of restorations requested. A high demand for restoration will reduce the overall benefits of data archiving. This should be used in conjunction with the backup KPI of the number of restores since both backups and archives will be options for recovery at times.
- **Average restore time**—The user community will support data archiving if data can be restored easily and promptly when required. The Ministry should be able to track and report on the average time taken to service a request to restore.
- **Percentage utilization of server capacity**—the ultimate objective of archiving to preserve server capacity for production datasets. Effective archiving should preserve capacity and improve server performance.

- **% of successful restores**—The ultimate success of the digital data archive process is dependent on data being successfully restored from archived media when required. The percentage of restoration requests that are successfully met should be tracked and reported. (This is also a Backup KPI and a Retention and Disposal KPI)
- **Cost of the archive process**—a primary objective of archiving is to reduce or avoid costs.

Retention Strategy

The issue of retention and disposal of archived digital data is a records management issue and should be subject to the Ministry's records management policies and procedures.

The Ministry establishes physical control of records by ensuring they are identified, documented, located, retrieved, and protected from loss, physical damage or inappropriate access. The Ministry establishes intellectual control over its records by ensuring they are classified, retained and disposed of in accordance with their values.

Records management in the BC Government is governed by legislation and policy, including:

- The **Document Disposal Act**—governs the final disposition of government documents by specifying the approvals required before they may be destroyed, transferred to the custody of the government archives or alienated from the Crown provincial;
- The **General Management Operating Policy (GMOP)** and the Financial Administration Operating Policy (FAOP);
- Classification and scheduling systems, such as the government-wide **Administrative Records Classification System (ARCS)** and program-specific **Operational Records Classification System (ORCS)**;
- **BC Archives** establishes government-wide records management policy through specific guidelines, policies and standards; and,
- **Ministry Records Officers** establish ministry policies and procedures in compliance with government-wide policies.

Vision

The vision for the Ministry's Digital Data Retention Strategy is as follows:

By December 1, 2002, MSRM will have in place Ministry-wide policies and procedures that support a consistent and effective process for the retention and final disposition of digital data that meets the Ministry's requirements for the identification and preservation of information with long-term value and the destruction of valueless records in a timely and secure manner.

Strategic Objectives

The objectives of the retention strategy are as follows:

To preserve archived digital data for as long as the Ministry requires, to meet its operational, legal, financial and historical responsibilities.

To dispose of digital data consistent with the Ministry's operational and legal responsibilities.

Scope

The retention strategy is applicable to all digital data stored on Ministry-managed storage media.

Critical Success Factors

In order to realize the objective of the retention strategy, the Ministry must achieve the following critical success factors:

- Archived digital data must be maintained in a way that protects it from loss, damage, degradation, loss of information, and other threats to its physical integrity:
 - Archived digital data must be maintained on media and in formats that ensure it is readable and accessible for the duration of its semi-active retention period;
 - Archived digital data must be housed in environmental conditions that meet its preservation, retrieval and security requirements;
 - The archive format or media must not compromise the Ministry's responsibilities or ability to use the information it contains.

-
- The context and authenticity of archived digital data must be preserved for as long as the Ministry has responsibility for the information:
 - Profile information (metadata) must be linked to archived digital data in a way that ensures that it is identifiable and authentic, and the context of its creation and use is maintained; and,
 - Archived digital data that is moved to different media must maintain its context and authenticity.
 - Archived digital data should be classified as a file and the file should be linked to a retention and disposition schedule through its classification:
 - Digital data should be archived as an ARCS or ORCS 'file'. Each 'file' should be classified;
 - When a file is classified, an appropriate retention and disposition schedule should be applied;
 - The classification function should define the Office of Primary Responsibility (OPR) for each 'file';
 - The date upon which digital data is archived becomes the "scheduling date". This date should be used to calculate how long the file is retained and when it is eligible for final disposition.
 - Archived digital data should be retained for the entirety of its semi-active retention period;
 - Archived digital data should not be retained past the end of its semi-active retention period; and,
 - All digital data archived to a 'file' should have the same scheduling status and be retained for the same period of time.
 - Scheduled final disposition should be applied to all digital data in a 'file'. Final disposition should be applied to files when eligible (once it has reached inactive status), unless a halt or hold to final disposition action is required.
 - A designated individual should authorize final disposition; and,
 - When files are stored in fixed containers (e.g., a DLTtape), final disposition should be applied to the entire container.

Key Performance Indicators

In order to assess the ongoing effectiveness of the Ministry's retention strategy, the following performance indicators should be adopted. Targets should be established for each performance indicator and actual performance should be monitored and reported against the targets.

- **Percentage utilization of server capacity**—the ultimate objective of archiving is to preserve server capacity for production datasets. Effective archiving should preserve capacity and improve server performance. (This is also an archiving KPI).
- **% of successful restores**—The ultimate success of the digital data archive process is that data can be restored successfully from archived media. The % of restoration requests that are successfully met should be tracked and reported.

Implementation Plan

The following Implementation Plan presents the high-level actions that must be undertaken by the Ministry to implement the strategy. The Implementation Plan presents two project plans: the implementation plan for the Digital Data Backup Strategy; and, the implementation plan for the Digital Data Archive and Retention Strategy. These two projects are closely related with some variances in requirements and approach. The Ministry has the option of managing the implementation of the strategies as an integrated project, or as two separate projects. In an integrated project, tasks such as the assignment of project sponsor, project manager, development of the project plan, and monitoring and reporting would be combined.

The Implementation Plan is presented in tabular format and a Gantt Chart for each plan is attached in Appendix D. Long-term tasks in the Digital Data Archive and Retention Strategy have been treated as separate sub-projects and are not included in the project Gantt Chart.

MSRM recognizes the importance of moving forward with the implementation of the Digital Data Backup, Archive and Retention Strategy. Ministry staff have already begun to translate this high-level Implementation Plan into manageable project components that can be quickly implemented. A draft “Interim Regional Strategy” document has been developed and is attached as Appendix E.

Digital Data Backup Strategy

Task Name/Description	Start	Finish	Dependencies	Resource Requirements
<p>1. Assign a Project Sponsor</p> <p>Assign a Project Sponsor for the rollout of the new Digital Data Backup and Recovery process. All projects have an increased chance of success if an individual with the appropriate authority and organizational prestige sponsors them.</p>	June 1/02	June 30/02	None.	<ul style="list-style-type: none"> Staff time.
<p>2. Assign a Project Manager</p> <p>Responsibility for the implementation of the Digital Data Backup and Recovery Strategy implementation project should be assigned.</p>	June 1/02	June 30/02	Task 1.	<ul style="list-style-type: none"> Staff time.
<p>3. Develop a Project Plan</p> <p>The Project Manager should develop a draft project plan for implementation of the Digital Data Backup and Recovery Strategy for submission to and approval by the Project Sponsor.</p>	July 1/02	July 15/02	Task 2.	<ul style="list-style-type: none"> Staff time – estimate of 40 hours.
<p>4. Document the Digital Data Backup Policy</p> <p>The Digital Data Backup Strategy should be documented in appropriate form and be formally adopted as Ministry policy.</p>	July 16/02	July 31/02	Task 3.	<ul style="list-style-type: none"> Staff time – estimate of 25 hours.

Task Name/Description	Start	Finish	Dependencies	Resource Requirements
<p>5. Identify Roles and Responsibilities</p> <p>Roles and responsibilities for backup and recovery functions should be defined.</p>	July 16/02	July 31/02	Task 3.	<ul style="list-style-type: none"> ■ Staff time.
<p>6. Assign Roles and Responsibilities</p> <p>Roles and responsibilities for backup and recovery functions should be assigned.</p>	Aug 1/02	Aug 31/02	Task 5.	<ul style="list-style-type: none"> ■ Staff time.
<p>7. Develop the Backup and Recovery procedures and Communicate to all stakeholders</p> <p>The Ministry needs to develop and communicate the complete backup and recovery procedures to all stakeholders:</p> <ul style="list-style-type: none"> ■ Develop and implement negotiated service level agreements with the MSRM client base. These service level agreements need to identify roles and responsibilities, response times, and recovery requirements; ■ Develop a backup and recovery procedures manual that outlines roles, responsibilities, procedures and standards for both regional and headquarters staff; and, ■ Develop and implement a Communications 	Aug 1/02	Sept 1/02	Task 3.	<ul style="list-style-type: none"> ■ Staff Time – estimate of 100 hours.

Task Name/Description	Start	Finish	Dependencies	Resource Requirements
Plan.				
<p>8. Identify metadata requirements and collect metadata</p> <p>The complete metadata of the backup data must be preserved for the entire retention period for the backup:</p> <ul style="list-style-type: none"> ■ An application must be available to store and track the backup metadata; ■ Metadata must be linked to backup data in a way that ensures that it is identifiable and recoverable, and the context of its creation and use is maintained; ■ Backup data that is moved to different media must maintain its context and authenticity; and, ■ Physical tapes must be labelled to sufficiently identify them with the corresponding metadata identifier. <p>This metadata application must be accessible by all staff involved in backup and recovery procedures.</p>	July 16/02	Aug 31/02	Task 4.	<ul style="list-style-type: none"> ■ Staff Time – estimate of 120 hours ■ Database to house metadata. ■ Application server to run metadata management application.
<p>9. Maintain backup data</p> <ul style="list-style-type: none"> ■ Backup data must be maintained on media and in formats that ensure it is readable and accessible for the duration of the retention 	Immediate	Ongoing	None.	<ul style="list-style-type: none"> ■ Staff time (ongoing).

Task Name/Description	Start	Finish	Dependencies	Resource Requirements
period; <ul style="list-style-type: none"> ■ Backup data must be housed in environmental conditions that meet its preservation, retrieval and security requirements; 				
<p>10. Implement a unified backup and recovery infrastructure for both Unix and Windows NT/2000</p> <ul style="list-style-type: none"> ■ A single backup and recovery tool is required in order for consistent backup and recovery procedures. The Information Management Branch (IMB) is planning to implement a new technical infrastructure using Veritas NetBackup software to perform the backups and Veritas Global Data Manager to remotely control backup operations at regional sites. ■ Existing DLTtape drives will be moved from the UNIX server to the Windows 2000 server. The UNIX backups will then be executed using a network connection between the Windows 2000 server and the UNIX server, controlled through a backup client program running on UNIX. ■ Single DLTtape drives may have to be moved from the Windows 2000 servers to the UNIX 	July 16/02	Aug 31/02	Task 4.	<ul style="list-style-type: none"> ■ Veritas Consultant – estimate cost to be between \$10,000 and \$15,000. ■ Purchase required jukebox technology. ■ Purchase Veritas for Unix Agent. ■ Potential upgrade to network connections between Windows and Unix servers.

Task Name/Description	Start	Finish	Dependencies	Resource Requirements
servers to aid in the recovery from, or conversion of, existing tapes.				
<p>11. Develop and Deliver a Training Program</p> <p>Develop and deliver a training program for all key staff involved in the backup and recovery of digital data. Consideration should be given to ensuring that sufficient numbers of staff are trained to provide service coverage throughout the year:</p> <ul style="list-style-type: none"> ■ Develop training materials; ■ Determine key staff to be trained; and, ■ Deliver training. 	Sept 1/02	Ongoing	Task 4.	<ul style="list-style-type: none"> ■ Veritas Trainer ■ Training materials ■ Training locations
<p>12. Migrate old format backup tapes (Legato or Novell)</p> <ul style="list-style-type: none"> ■ All existing monthly Legato backup tapes must be migrated to a Veritas format. ■ Existing Novell-based backups must be converted to a Windows compatible format. 	Sept 1/02	Oct 31/02	Tasks 4 & 5.	<ul style="list-style-type: none"> ■ Backup technician ■ Tape drive on the Unix server ■ Novell environment
<p>13. Assess Onsite Media Storage Facilities</p> <p>Identify all storage facilities used for onsite backup media and assess the appropriateness of those facilities for protecting the backup media.</p>	July 1/02	Aug 31/02	Task 4.	<ul style="list-style-type: none"> ■ Staff time.

Task Name/Description	Start	Finish	Dependencies	Resource Requirements
<p>14. Upgrade Media Storage Facilities</p> <p>We expect that the assessment of media storage facilities will confirm the need to upgrade those facilities. While most regional offices use secure, fireproof cabinets, we are aware of several regions that do not have adequate storage facilities.</p>	Sept 1/02	Oct 31/02	Task 13.	<ul style="list-style-type: none"> ■ Staff time. ■ Secure fireproof cabinets.
<p>15. Track key performance indicators</p> <ul style="list-style-type: none"> ■ % of backups performed within the backup window; ■ % of successful restores; and, ■ Average time to restore data from daily, weekly and monthly backup tapes. 	Sept 1/02	Ongoing	Task 4 (Service Level Agreements).	<ul style="list-style-type: none"> ■ Staff Time
<p>16. Monitoring and Reporting</p> <p>Progress in implementing the Digital Data Backup Strategy should be monitored and reported to the project sponsor. Progress reports should be provided periodically over the project's life.</p> <p>Performance measures and performance targets should be developed, monitored and reported on.</p>	June 1/02	Oct 31/02	Task 3.	<ul style="list-style-type: none"> ■ Staff time.

Digital Data Archive and Retention Strategy

Short-Term Project Tasks				
<p>1. Create an overall media and data management infrastructure</p> <p>The Ministry's backup, archive and retention practices are a sub-set of the overall media and data management processes. To ensure that all media and data management requirements are addressed, an overall media and data management infrastructure should be established to integrate the Ministry's approach to the management of hardcopy and digital records.</p>	June 1/02	Ongoing	Must be coordinated with other data management projects underway in the Ministry.	<ul style="list-style-type: none"> ■ Management sponsorship. ■ Staff time.
<p>2. Assign a Project Sponsor</p> <p>Assign a Project Sponsor to the rollout of the new digital data archive process. All projects have an increased chance of success if an individual with the appropriate authority and organizational prestige sponsors them.</p>	June 1/02	June 30/02	None.	<ul style="list-style-type: none"> ■ Staff time.
<p>3. Assign a Project Manager</p> <p>Responsibility for the implementation of the Digital Data Archive Strategy should be assigned.</p>	June 1/02	June 30/02	Task 2.	<ul style="list-style-type: none"> ■ Staff time.

<p>4. Develop a Project Plan</p> <p>The Project Manager should develop a draft project plan for implementation of the Digital Data Archive Strategy for submission to and approval by the Project Sponsor.</p>	July 1/02	July 15/02	Task 3.	<ul style="list-style-type: none"> ■ Staff time – estimate of 40 hours.
<p>5. Document the Digital Data Archive Policy</p> <p>The Digital Data Archive Strategy should be documented in appropriate form and be formally adopted as Ministry policy.</p> <p>The policy should be supported by clear procedures for the creation, maintenance, recovery and deletion of archives.</p>	July 16/02	July 31/02	Task 4.	<ul style="list-style-type: none"> ■ Staff time – estimate of 25 hours.
<p>6. Identify Roles and Responsibilities</p> <p>Responsibilities for identification, classification, archiving, storage, restoration and disposal of archived digital data must be defined. Possible role definitions may be as follows:</p> <p>User Role, including:</p> <ul style="list-style-type: none"> ■ Identifying the opportunity to archive data; ■ Migrating data to an archive directory; ■ Requesting an archive; ■ Providing required metadata; ■ Identifying the need to recover data from an 	July 16/02	July 31/02	Task 5.	<ul style="list-style-type: none"> ■ Staff time – estimate of 40 hours.

<p>archive;</p> <ul style="list-style-type: none"> ■ Requesting the recovery of data from an archive; and, ■ Identifying changes in the metadata associated with an archive, such as a change to the record classification. <p>Regional Support Role, including:</p> <ul style="list-style-type: none"> ■ Generating the archive; ■ Maintaining the archive metadata, including identifying factors that could affect the availability of an archive such as hardware or software obsolescence; ■ Ensuring the appropriate storage and safekeeping of archive media (local copy); ■ Restoring data from the archive (local copy); ■ Refreshing archive media as required by policy; ■ Migrating storage media from one generation to the next; ■ Applying final disposition to archived materials; and, ■ Integrating old archives or archives coming from other organizations that do not follow the Ministry's standards. 				
--	--	--	--	--

<p>Headquarters Support Role, including:</p> <ul style="list-style-type: none"> ■ Maintaining the Ministry’s archive policies and procedures; ■ Ensuring the appropriate storage and safekeeping of archive media (central copy); ■ Restoring data from the archive when the region is unable (central copy); ■ Refreshing archive media as required by policy; ■ Identifying conflicts between changes to the Ministry’s technical architecture and the continuing ability to recover information from archived storage media; ■ Migrating storage media from one generation to the next; ■ Applying final disposition to archived materials; and, <p>Providing and maintaining a Ministry-wide tracking system for archived data.</p>				
<p>7. Assign Roles and Responsibilities</p> <p>Roles and responsibilities for digital data archiving should be assigned.</p>	<p>Aug 1/02</p>	<p>Aug 31/02</p>	<ul style="list-style-type: none"> ■ Task 6 — roles and responsibilities must be defined. 	<ul style="list-style-type: none"> ■ Staff time.

<p>8. Develop the Archive and Retention procedures and communicate to all stakeholders</p> <p>The Ministry needs to develop and communicate the complete archive and retention procedures, providing detailed instructions for the workflow and technical procedures for the creation, maintenance, recovery and deletion of archives, to all stakeholders:</p> <ul style="list-style-type: none"> ■ Develop and implement negotiated service level agreements with the MSRM client base. These service level agreements need to identify roles and responsibilities, response times, and recovery requirements; ■ Develop an archive and retention procedures manual that outlines roles, responsibilities, procedures and standards for both regional and headquarters staff; and, ■ Develop and implement a Communications Plan. 	Aug 1/02	Sept 30/02	Task 5.	<ul style="list-style-type: none"> ■ Staff Time – estimate of 80 hours.
<p>9. Develop a Classification Scheme</p> <p>In cooperation with the Ministry’s Records Officer, develop a classification scheme for project datasets.</p> <p>The classification scheme should address the</p>	Aug 1/02	Sept 30/02	■ Task 4.	<ul style="list-style-type: none"> ■ Staff time – estimate of 50 hours.

requirements for project datasets, but should ultimately encompass all forms of digital data stored on primary and archival media.				
<p>10. Develop Final Disposition Procedures</p> <p>As part of the Archives Procedures Manual, develop detailed procedures for final disposition of archived materials.</p> <p>The final disposition procedures should provide for the transfer of archived materials to BC Archives or the destruction of the archived media. The procedures should also address the authorization and documentation of the final disposition actions.</p>	Aug 1/02	Sept 30/02	■ Task 4.	■ Staff time – estimate of 40 hours.
<p>11. Develop Naming Convention</p> <p>A standard naming convention should be developed for each archived dataset. The naming convention should support the ability of both users and systems support staff to be able to quickly locate a requested archived dataset.</p> <p>The naming convention should consider the following:</p> <ul style="list-style-type: none"> ■ Dataset number; ■ Ministry; ■ Responsibility Centre or location; 	July 16/02	Aug 31/02	■ Task 4.	■ Staff time – estimate of 40 hours.

<ul style="list-style-type: none"> ■ Type of dataset (project type or other); and, ■ Dataset name (project name, etc.). 				
<p>12. Identify metadata requirements</p> <p>Identify metadata requirements for the archive tracking system. This process should review the existing metadata that is tracked in each region and consider the broader requirements identified by the strategy document.</p> <p>The metadata requirements for datasets may include the following:</p> <ul style="list-style-type: none"> ■ Dataset name (name or catalogue #); ■ Dataset date (i.e., the date of the project report or other milestone date); ■ Data classification (i.e., for purposes of retention and disposal); ■ Disposal date; ■ Migration history (i.e., records of where the dataset has been migrated from one archive media to another); ■ Compression format (if any); ■ Archive software (i.e., Veritas Version X); 	July 16/02	Aug 31/02	<ul style="list-style-type: none"> ■ Task 4. 	<ul style="list-style-type: none"> ■ Staff time – estimate of 40 hours.

<ul style="list-style-type: none"> ■ Data custodian (who authorized the archive); ■ Archivist (person performing the archiving); ■ Date archived (mm/dd/yyyy); and, ■ Archive size (megabytes). <p>The metadata requirements for files within datasets may include the following:</p> <ul style="list-style-type: none"> ■ Filename; ■ File format (i.e., Word 2000, Excel 2000, ArcInfo 8.02, Microstation, Mapinfo, Oracle 10i, etc.); ■ Original (native format) or Preservation Master (non-proprietary format) checkbox; and, ■ Operating system environment (i.e., Unix AIX 8.1, Windows NT/2000, Novell Netware, etc.). 				
<p>13. Implement a Short-Term Tracking System</p> <p>A Ministry-wide tracking system is required for archived digital data. The Enterprise Document and Records Management System (EDRMS) being developed by BC Archives may provide a long-term solution. As EDRMS is only in the pilot stage, it will not be available as a solution for at least one year. Until such time as the Ministry can evaluate EDRMS</p>	Sept 1/02	Sept 30/02	<ul style="list-style-type: none"> ■ Task 4. 	<ul style="list-style-type: none"> ■ Staff time – estimate of 40 – 100 hours. ■ Expected use of Excel or MS Access.

<p>as a longer-term solution, it will need a tracking system to meet the need in the short term.</p> <p>The Ministry should be cautious about making a significant investment in a tracking system until more is known about the capabilities of EDRMS to support digital data archiving.</p> <p>There are a number of regional tracking systems in place that may form the basis of a shorter-term solution.</p>				
<p>14. Implement Pilot Project</p> <p>Select a regional office to act as a pilot project for the archive and retention strategy. The pilot project should be structured to validate the effectiveness of the following:</p> <ul style="list-style-type: none"> ■ The hardware and software infrastructure for both Unix and Windows NT/2000; ■ The archive and retention policy; ■ The archive and retention procedures; ■ Recovery procedures; ■ The data classification scheme; ■ The final disposition procedures; ■ The metadata requirements; 	<p>Sept 1/02</p>	<p>Oct 31/02</p>	<ul style="list-style-type: none"> ■ Tasks 5, 8, 9 10, 11, 12, and 13. 	<ul style="list-style-type: none"> ■ Staff time – estimate of 100 hours. ■ 1 Tape Drive. ■ 1 Veritas License – estimated to cost \$3,500 and \$630 per year for support.

<ul style="list-style-type: none"> ■ The short-term metadata tracking system; ■ The storage facilities; and, ■ The file naming convention. 				
<p>15. Assess Media Storage Facilities</p> <p>The Ministry should identify all storage facilities used for archive media and assess the appropriateness of those facilities for protecting the archive media.</p>	July 1/02	Aug 31/02	<ul style="list-style-type: none"> ■ Task 4. 	<ul style="list-style-type: none"> ■ Staff time.
<p>16. Upgrade Media Storage Facilities</p> <p>While most archive libraries use secure, fireproof cabinets, we are aware of several regions that do not have adequate storage facilities. We expect that the assessment of media storage facilities will confirm the need to upgrade those facilities. In addition, as backups will likely be moved into the same secure storage facilities as a result of the Digital Data Backup Strategy, the overall capacity of the secure storage facilities may be exceeded.</p>	Sept 1/02	Oct 31/02	<ul style="list-style-type: none"> ■ Task 15. 	<ul style="list-style-type: none"> ■ Staff time. ■ Secure fireproof cabinets.
<p>17. Implement a unified Archive infrastructure for both Unix and Windows NT/2000</p> <p>Standalone hardware and software must be implemented at each site requiring an archive capability. Our initial assessment has determined that the Ministry should not attempt to utilize the</p>	Nov 1/02	Nov 30/02	<ul style="list-style-type: none"> ■ Task 14. 	<ul style="list-style-type: none"> ■ Tape drives (7 – 10). ■ Staff time – estimated to be 40 – 60 hours. ■ Veritas licenses are estimated to cost \$3,500 per tape drive and \$630

<p>backup infrastructure to perform archiving, as the potential to negatively impact the backup process is high.</p> <p>A tape drive with one or more tape slots using Veritas Netbackup is the likely configuration required. As the Ministry is currently installing new backup tape drives, it is possible that the replaced drives can be reused to support archiving.</p>				<p>per year for support.</p>
<p>18. Develop & Deliver a Training Program</p> <p>Develop and deliver a training program for all key staff involved in the archiving of digital data. Consideration should be given to ensuring that sufficient numbers of staff are trained to provide service coverage throughout the year.</p>	<p>Nov 1/02</p>	<p>Dec 31/02</p>	<ul style="list-style-type: none"> ■ Task 17. 	<ul style="list-style-type: none"> ■ Staff time – estimate of 50 hours to develop, and 100 hours of instructor time to implement. ■ Training materials ■ Training locations
<p>19. Develop a Communications Plan</p> <p>A Communications Plan should be developed with the objective of creating awareness of the existence of a new approach to the archiving of digital data. The Communications Plan should:</p> <ul style="list-style-type: none"> ■ Identify the key communication messages to be conveyed; ■ Identify the appropriate means for conveying the communication messages; and, 	<p>Dec 1/02</p>	<p>Dec 15/02</p>	<ul style="list-style-type: none"> ■ Tasks 17. 	<ul style="list-style-type: none"> ■ Staff time – estimate of 20 hours.

<ul style="list-style-type: none"> Identify the appropriate audience(s) for the communication messages. 				
<p>20. Implement the Communications Plan</p> <p>Implement the Communications Plan to ensure that all Ministry staff involved in the management of digital data are aware of the new approach to the archiving of digital data.</p>	Dec 16/02	Dec 31/02	<ul style="list-style-type: none"> Task 19. 	<ul style="list-style-type: none"> Staff time – estimate of 20 hours.
<p>21. Track key performance indicators</p> <p>Possible kpi's include:</p> <ul style="list-style-type: none"> Total size of the Digital Data Archive Volume of data archived and number of new archives created over the year Percentage of Archived Digital Data that adheres to Archive Standard Number of restores Average restore time Percentage utilization of server capacity % of successful restores Cost of the archive process 	Dec 1/02	Ongoing.	<ul style="list-style-type: none"> Task 17. 	<ul style="list-style-type: none"> Staff time (ongoing).
<p>22. Monitoring and Reporting</p> <p>Progress in implementing the Digital Data Archive</p>	June 1/02	Dec 31/02	<ul style="list-style-type: none"> Task 4. 	<ul style="list-style-type: none"> Staff time.

<p>Strategy should be monitored and reported to the project sponsor. Progress reports should be provided periodically over the project's life.</p> <p>Performance measures and performance targets should be developed, monitored and reported on.</p>				
<p>Longer-Term Project Tasks</p>				
<p>23. Migrate Existing Archived Data</p> <p>Procedures should be developed and communicated to aid in the migration of existing archives (e.g. DLTapes, CD-ROMs).</p> <p>Systematically identify all archived digital data with continuing value to the Ministry. For archived materials of continuing value, migrate the archived material to the new archive standard. For archived materials with no continuing value, apply final disposition.</p> <p>The first activity in this task is to locate and take control of all existing archive media. This activity can begin immediately, without waiting for the implementation of archiving equipment or the finalization of the procedures. This activity should be assigned a high priority to ensure that archived materials are not lost as a result of possible staff or Ministry reorganization.</p>	<p>Jan 1/03</p>	<p>Mar 31/03</p>	<ul style="list-style-type: none"> ■ Tasks 17 and 18. 	<ul style="list-style-type: none"> ■ Staff time – estimate of 150 hours per region. ■ Hardware and software capable of processing the old archives (and the datasets archived if that is no longer supported). ■ Sufficient tapes to hold the existing archives (some existing tapes currently used for archives may be freed up).

<p>24. Implement a Long-Term Tracking System</p> <p>The Ministry should implement a tracking system for its archived digital data. EDRMS is expected to be available to Ministries sometime in the next year or two. The Ministry should consider ways in which it can evaluate EDRMS prior to its official rollout. EDRMS creates the potential to better integrate the management of digital and hardcopy data.</p>	Jan 1/03	Dec 31/03	<ul style="list-style-type: none"> ■ Must wait for the rollout of the ERDMS in order to assess feasibility of that solution. 	<ul style="list-style-type: none"> ■ Staff time. ■ Should ERDMS not be appropriate, the Ministry would have to acquire or develop an alternative solution. No basis for estimating this cost at this time.
<p>25. Develop a Migration Trigger Mechanism</p> <p>Develop a mechanism within the tracking system that will signal to the Ministry when it is necessary to migrate digital data when the format of the digital data can no longer be supported for restoration.</p> <p>Digital data is created in a technology context that limits the ability to access the information. The context includes the hardware, operating system software, archive software and application software used to create and store the data. As changes occur in the Ministry's technical architecture, it is possible that the Ministry will lose the ability to access the archived materials.</p> <p>The trigger mechanism should identify when changes to the Ministry's technical architecture impact the ability to access archived data. To do</p>	Jan 1/03	Dec 31/03	<ul style="list-style-type: none"> ■ Must wait for the rollout of the ERDMS in order to assess feasibility of that solution. 	<ul style="list-style-type: none"> ■ Staff time. ■ Possible development costs depending on requirements and capability of EDRMS.

this, the trigger mechanism must track data formats, data formats supported by the Ministry and situations where the formats can no longer be supported.				
--	--	--	--	--

Implementation Issues

In the course of developing the Digital Data Backup, Archive and Retention Strategy, we identified several issues that will impact the Implementation Plan, including the following:

Veritas Rollout

We understand that the rollout of the Veritas NetBackup and Global Data Manager has encountered a number of technical challenges which are still being addressed. The Implementation Plan is dependent upon the successful rollout of the Veritas products.

Non-Project Data

As identified in the project scope, we were asked to focus on the issue of archiving large project datasets. This direction recognized that the size of project datasets offers the Ministry the greatest opportunity for archiving and reducing storage costs. Our strategy document focused on the archiving and retention of these large project datasets at the time their status changes from 'active' to 'semi-active'. A key requirement of the strategy is the classification of digital data files at the time they are archived from primary storage. The strategy did not address the need to classify digital data that is still 'active' and stored online. It must be recognized that digital data stored online is generally not classified by the Ministry and issues of retention and final disposition are not being addressed. This issue was considered to be out of scope in the development of the strategy document.

Archive Media Capacity

In most cases, archive media will have the capacity to store multiple archived datasets. For retention purposes, the Ministry may need to ensure that each archive media stores datasets that have the same classification and retention period. This would allow the Ministry to dispose of the archive media without having to delete only specific datasets on the media. If the archive and retention procedures restrict putting only archived datasets with the same classification and retention period on archive media, the Ministry will likely not be optimising the utilization of the media capacity.

Evolving Roles and Responsibilities

There are a number of government initiatives underway that may result in organizational change and shift roles and responsibilities. With respect to responsibilities for backup and archiving of digital data, it is likely that there will be some shift of responsibilities from MSRM to CITS. These changes will have to be taken into consideration in the implementation of the backup, archive and retention strategy.

Appendix A – Interview List

Name	Position
Alex Halkett	Spatial Data Analyst, Data Services, Information Management Branch
Anne Murphy	Regional GIS Coordinator, Fort St. John
Brandt Bady	Manager, Emerging and Applied Information Technologies, BC Archives, Ministry of Management Services
Carol Ogborne	A/Head, GIS Analysis, Decision Support Services Branch
Charles Middleton	Senior Network Analyst, Network Services Unit
Chris Scarff	Data Applications Officer, Nelson
Colleen Coccola	Head, Database Administration & Application Q&A Services, Information Management Branch
Drew Smyth	Ministry Records Officer, Information, Privacy and Records Services Section
Graham Hawkins	Forest and SRM Planning Specialist, Resource Planning Branch
Greg Roy	Head, Regional Systems, Nanaimo
Jason Birch	Spatial Data Analyst, Nanaimo
Jennifer Ballentine	Senior GIS Analyst, Williams Lake
Jennifer Dawes	A/Manager, Provincial Baseline Atlas Service, Base Mapping and Geomatic Services Branch
Jerry Renwick	Systems Officer, Kamloops
Johanna Pfalz	Section Head, GIS/Data Management, Smithers
Joshua Chan	A/GIS Coordinator, Surrey
Linn Anderson	Systems Officer, Nelson
Lynne Tibbitt	Team Leader, Electronic Information Management Initiatives, BC Archives, Ministry of Management Services
Mike Hanson	Systems Officer, Williams Lake
Mike Howard	Resource Information Section Head, William Lake
Mike Ross	Senior Land Data Systems Analyst, Base Mapping and Geomatic Services Branch
Mike Seper	Regional Coordinator, Information Technology, Surrey

Name	Position
Nancy Liesch	Business Area Coordinator, Decision Support Services Branch
Paul Middleton	Senior Database Administrator, Information Management Branch
Paul Quakenbush	Supervisor, Standards and Data Exchange, Base Mapping and Geomatic Services Branch
Per Wallenius	GIS Data Management, Nelson
Rick Thomas	Extension Specialist, Registries and Resource Information Division
Sean Barry	GIS Analyst, Prince George
Stan Hoffmann	Systems Administrator, Sub-Surface Tenures, Registries and Resource Information Division
Terry Chan	Application Delivery Specialist, Information Management Branch
Tom Fulton	A/Head, Data Admin Unit, Information Management Branch
Tyrone Lindley	Systems Officer, Fort St. John
Wendy Herring	Systems Officer, Prince George

Appendix B – Glossary of Terms

Term	Definition
Archive	<p>There are several accepted definitions of archiving:</p> <ul style="list-style-type: none"> ■ As a noun, an archive is “a place in which public records or historical documents are preserved, or the material preserved.” ■ As a verb, archiving means “to file or collect (as records or documents) in or as if in an archive”. ■ Archiving involves selective preservation of electronic information that has long-term value (“official” records) over an extended period of time. In the British Columbia government context, “archiving” does not occur until the Ministry records of long-term or permanent value are transferred to BC Archives. ■ Archiving is the process of systematically making copies of your most current data and storing those copies in a safe, off-site location where they can be readily accessed if needed for restoration. Backup tapes should be regularly and quickly sent to an off-site archive for protection. Archiving ensures a chance of recovery from a data disaster if on-site copies of data are lost. Also, archiving is used for long-term data retention to safeguard operational needs and meet legal or regulatory requirements.⁴ ■ Archiving and backup are two distinct processes. Backing up is essential to systems management in that it involves creating copies of hard disk files on tape (or other storage media) for ready retrieval, if needed. Archiving involves a detailed system for securely and systematically storing backup files in a safe, climate-controlled alternate location. It also includes a detailed cataloguing and retrieval process for data recovery and restoration when needed. And, archiving takes into account media

⁴ <http://www.dlftape.com/Practices/Archive/Default.htm#>.

Term	Definition
	<p>migration issues so that legacy data remains accessible as media device standards change over time.</p> <ul style="list-style-type: none"> ■ Digital data archiving is the process of moving digital data of historical value from on-line to off-line storage media, in order to optimize storage costs, and in a way that preserves the data so that it can be accessed and understood in the future.
Backup	Copying files to a second medium (a disk or tape) as a precaution in case the first medium fails.
Cold Backup	A cold backup is performed when an Oracle database is shutdown and all Oracle data files are backed up tape or disk. This type of backup allows for recovery of the complete database or any of the data files associated with the database.
Dataset	An organized group of related information stored on storage media.
Digital data archiving	Digital data archiving is the process of moving digital data from on-line storage media to near-line or off-line storage media in order to preserve the data and to optimize storage costs.
Document	Recorded information, which can be treated as a unit. Any medium that contains information is a document. A document has the status of a record only if it is created or received in the course of government business and if it is maintained as evidence of that business. Most documents created or received by government offices are records.
Export	An Oracle export is a function that will export the complete table structure, data and security configuration for a complete Oracle database or any of the existing schemas in an Oracle database. These export files allow for recovery of complete databases, individual schemas and/or individual tables in an Oracle database.
File	The logical entity used to organize and manage

Term	Definition
	records. A file manages a group of records that together provide evidence of a complete transaction or a collection of reference material.
Gigabyte (GB)	Approximately 1 billion bytes of data.
Hot Backup	A hot backup is a mechanism where every data change against a database (update, insert, delete) is recorded in special files called archive logs. In a hot backup mechanism, the archive logs are automatically backed up to tape, disk or another database (called a Recovery Catalogue) to allow for point-in-time recoveries of databases in the event of media failure or incorrect user actions.
Incremental or Differential Backup	The backup of all data files that have been modified since the last incremental backup.
Lossless compression	Refers to data compression techniques in which no data is lost.
Lossy compression	Refers to data compression techniques in which some amount of data is lost. Lossy compression technologies attempt to eliminate redundant or unnecessary information.
Megabyte (MB)	Approximately 1 million bytes of data.
Metadata	Data about data. Metadata describes how and when and by whom a particular set of data was collected, and how the data is formatted.
Migration	Migration is a set of organized tasks designed to achieve the periodic transfer of digital materials from one generation of computer technology to a subsequent generation.
Near-line storage	Near-line storage is used as an inexpensive, scalable way to store large volumes of data. Near-line storage devices include DAT and DLT tapes (sequential access), CD-ROMs, DVD-ROMs DVD-RAMs and Magneto-Optical (random access) devices. Retrieval of data can be slow, but the type of data (archives, past records, etc.) dictates that the information will not be accessed often. While near-line storage uses the same technology as off-line storage, the difference is that

Term	Definition
	disks and tapes are not removed from the storage media and can be accessed electronically without manual intervention.
Off-line storage	Off-line storage is storage media that is not connected to the computing environment. Removable disks and tapes are examples of off-line storage media.
On-line storage	On-line storage is storage media that is connected to the computing environment. File servers and database servers are examples of on-line storage media.
Preservation Management	Relates to the physical and intellectual maintenance of records in order to preserve the information they contain.
Record	<ul style="list-style-type: none"> ■ A document created or received in the course of government business and maintained for action or reference by an agency as evidence of that business. ■ Recorded information, in any form, created or received and maintained by an organization in the transaction of business and kept as evidence of such business.
Records Management	The exercise of physical and intellectual control over records to ensure their integrity in support of government accountabilities and actions.
Retention Period	The length of time a file is retained, and is determined by a retention and disposition schedule.
Schema	An Oracle schema is defined as an Oracle user that owns objects (tables, indexes, etc.). The schema is usually synonymous with an application database. Multiple schemas may be required, however, to completely recover an application database.
Terabyte (TB)	Approximately 1 trillion bytes of information.

Appendix C – Available Storage Medias

To support the backup and archive requirements of the Ministry, the following storage medias are available.

Digital Tape⁵

Until about 1996, the dominant technology for tape backup of mid-range systems (mini computers, department servers, high-end workstations) was 8mm tape. The leading supplier of that technology was Exabyte. Their product was 8mm helical scan tape, a product descended from consumer videotape technology. In the desktop market, the dominant technology was 4mm helical scan tape. In the mainframe space, the dominant technology was IBM's 3590 "square tape." It was widely thought that these tape technologies would be fairly quickly superseded by optical disk technology.

Another technology on the scene was Digital Equipment Corporation's DLT, which was introduced in 1985 to support Digital's MicroVAX II workstation product line. It incorporated a square-tape design similar to IBM 3590, and a linear recording technology, whereby data is recorded in a linear pattern along the length of the tape, rather than in short pieces across the width of the tape. DLT was a hit with Digital's customers, but was not used by other system builders (Digital's competitors!). That all changed in 1994, when DLT technology was bought by Quantum Corporation, a major player in the disk drive industry. Under Quantum, DLT quickly went from obscurity to prominence. All the leading system manufacturers (Compaq, Dell, Digital, Hewlett-Packard, IBM, and Sun) adopted DLT as a standard peripheral option. By 1997, unit sales of DLT had surpassed those of 8mm tape. Within a few years, DLT technology (re-named DLTtape technology by Quantum) had captured over 80% of the mid-range tape market, taking share away from 8mm helical scan tape. In the desktop space, 4mm continued to dominate, and StorageTek's 9840 has taken over the biggest share of the mainframe market. Optical disk technology has not become a significant factor; the relatively low capacity of each optical disk means that it takes a lot of disks to back up gigabytes of data, and keeping track of lots of disks presents a serious management overhead burden.

While DLTtape technology and library automation were gaining prominence, an even more important trend has emerged: the tremendous explosion in demand for greater storage capacity at every level of the computing infrastructure. Internet, e-commerce, and image-based applications, including video, film, and audio, demand enormous amounts of storage. Some will use tape technology as

⁵ Excerpted from: *Quantum DLT tape Technology Versus Sony AIT-2 Technology in Automated Library Applications*, a StoreTek White Paper, April 9, 2001.

a primary storage medium. Most will use RAID and other disk technologies as primary storage. But all will require backup, and tape continues to be the dominant technology for backup.

As new and old applications drive the well-documented explosion in storage capacity, more and more situations will demand more tape storage capacity than is available on a single, stand-alone tape backup unit. More and more systems will be equipped with autoloaders and libraries. Instead of a standalone DLTtape drive with a single cartridge and native capacity of 20 or 40 GB, a video editing workstation is likely to have an autoloader with a single drive and a 7-cartridge magazine - and a native capacity of 280 GB or more. And so on up the line, to libraries with dozens of drives, hundreds of cartridges, and multi-terabytes of capacity, supporting applications like medical image archiving (tens of thousands of on-line X-ray images) or backup for Internet search engines. In very short order, libraries - and bigger and bigger libraries - will become the standard for tape storage across a very wide range of applications.

While DLTtape has emerged as the dominant tape technology and the dominant tape automation technology, it is not without challengers. They include:

- **Ultrium**—a new half-inch tape product announced by a consortium of leading system and storage suppliers. Ultrium drives have not been delivered in quantity at this writing.
- **Mammoth2**—a next-generation 8mm product from Exabyte. Mammoth2 has been delayed in getting to market, but is available.
- **Sony AIT-2**—a new 8mm technology from Sony. AIT-2 offers attractive specifications and has had some success in the automation market.

Capacity

Today, AIT-2 compares favourably with the most widely used DLTtape drives in terms of capacity, with a top capacity of 50GB per cartridge, versus 40GB per cartridge for the DLT 8000 drive. (All figures are for native (uncompressed) capacity. However, that is about to change. The new **Super DLTtape** media cartridge has an uncompressed capacity of 110GB running on the first of the new family of Super DLTtape drives, the **SDLT 220**. Future SDLT drives already in the product roadmap will drive that capacity up to 500GB, uncompressed. In contrast, the AIT roadmap extends only to 100 GB with **AIT-3** and 200 GB with **AIT-4**.

Speed

The speed comparison follows the same pattern: current AIT-2 speed (6MB/s) is the same as that of the current-generation DLT 8000, but well below that of the new SDLT 220 (11MB/s).

Reliability

Tape system reliability is usually measured using four specifications:

- Media life, based on end-to-end tape passes;
- Mean Time Between Failure, based on power-on hours at a specified duty cycle, excluding head life and excluding media life;
- Head life, based on tape contact hours; and,
- Duty cycle, defined as tape motion time over power-on hours.

In general, reliability specs for AIT-2 and DLT drives are comparable.

Scalability

Currently, AIT-2 offers a capacity of 50GB per cartridge, versus 40GB for DLT 8000. However, Quantum is now rolling out the first of its next generation Super DLTtape drives, which leapfrog the AIT-2. The first Super DLTtape drive to be launched, the SDLT 220 offers native capacity of 110GB per cartridge. Over the course of the currently published Super DLTtape product roadmap, that figure will extend to 500GB and ultimately to over one Terabyte over three generations of drives. In contrast, the AIT-3 product road map extends only to 100 GB per cartridge. That capacity edge will translate into a huge advantage when scaled across a larger library with several drives and dozens of cartridges.

Cost

Unit costs for AIT-2 and DLT 8000 drives are similar. Similarly, unit costs for DLTtape and AIT-2 media are currently comparable. This will change with the introduction of Super DLTtape. The highest capacity media usually carries the lowest cost per gigabyte. With double the capacity, the SDLT 220 with Super DLTtape will offer a per gigabyte cost well below that of AIT-2.

Summary

DLTtape offers a number of advantages over AIT-2, including:

- Most vendors are committed to DLTtape, while only a few are committed to AIT-2;

- DLTtape is the industry standard for automation, with support from all the major vendors;
- DLTtape drives and media are available from second sources. AIT media is second-sourced, but drives are not;
- Every significant backup and storage management software product supports DLTtape technology; and,
- 8mm helical scan technology does not have the technology future available with DLTtape.

Optical Disk

An optical disk is a plastic (or glass) substrate with one or more thin-film layers coated upon its surface(s). Data may be pre-recorded on the surface of the substrate by the disk manufacturer, or the user may record it on one or more of the thin film layers.

CD-Recordable/Rewritable Media⁶

Cost

CD-RW drives can write to both CD-R (write-once) and CD-RW (rewriteable) discs. The drives are inexpensive and software is readily available that will not only write your files to the CD media but will keep an inventory on-line of the data's location. Depending upon the CD media you purchase, discs cost between \$0.25 - \$10 each.

Capacity

The capacity of a CD is 650 MBs.

Reliability

While the life of good quality CDs has been demonstrated to be over 100 years, real world experience has shown that there are significant reliability issues which question using CD technology as an archival media for valuable information.

Sunlight and temperature, as well as humidity can be major factors in the degradation of CD-R and CD-RW discs, regardless of dye type. The more extreme the exposure, the faster the deterioration.

Drive compatibility is another critical factor. Variations in recorders and readers can cause media compatibility problems. These difficulties can be as subtle as slow read times and as serious as the inability to read at all. Problems can also be caused by differences in the write strategy (the way in which the drive's laser is modulated) from one drive manufacturer to the next. In addition, CD-R and CD-RW discs from different manufacturers have variations that can affect the tracking of the reader or writer.

Scalability

At 650 MB capacity, CD technology is limited in its ability to store larger datasets.

DVD-Recordable Media

DVD Recordable (DVD-R) technology allows anyone to create DVD discs at the desktop. Similar in concept to Compact Disc Recordable (CD-R), DVD-R is a write-once medium that can contain any type of information normally stored on mass produced DVD discs – video, audio, images, data files, multimedia programs, and so on. Depending on the type of information recorded, DVD-R discs are usable on virtually any compatible DVD playback device, including DVD-ROM drives and DVD Video players.

Capacity

A DVD-R disc is able to contain a maximum of either 4 – 4.7 gigabytes of information on each side, depending on the type of blank media used. Since the DVD format supports double-sided media, up to 9.4 Gbytes can be stored on a single double-sided DVD-R disc.

Speed

Data can be written to a disc at a DVD “1X” equivalent of 11.08 megabits per second (Mbps), which is roughly equivalent to nine times the transfer rate of CD-ROM’s “1X” speed. These transfer rates, coupled with DVD-R’s capacity and conformance to worldwide DVD standards, makes it a viable and cost effective storage medium.

Reliability

Life expectancy is a key issue when considering the use of DVD-R for applications such as document imaging and other archival applications. Although each disc media manufacturer has its own life expectancy rating, Pioneer DVD-R media is currently rated at better than 100 years.

Storage and Archival

DVD-R media provides archival lifetimes that are equal to or better than CD-R; media is rated at greater than 100 years. For this reason, the format may be suitable for long-term archival of any information that can be stored digitally. DVD-R's capacity makes it suitable for large image files that do not fit onto a single CD-R volume.

Since DVD discs are dimensionally identical to the CD family of discs, they have the advantage of being compatible with existing CD-based jukebox and changer mechanisms. This allows automated retrieval of recorded DVD-R volumes in networked environments, with a six to seven-fold increase in storage density as compared with CD-R technology.

As an example of how DVD-R can reduce overall archival system costs, a 100-disc DVD-ROM jukebox can contain a total of 470 gigabytes, or nearly a half-terabyte of data. Using CD-R, seven of the same jukebox mechanisms would be required to maintain the same disc-to-drive ratio, which significantly adds to the system cost.

Appendix D – Project Gantt Charts

Appendix E – Interim Regional Strategy (Draft)

Note: In this document the term ‘project’ refers to the Digital Data Backup, Archive and Retention Strategy project and the term ‘project dataset’ refers to working or operation datasets required for analysis and planning purposes such as environmental studies and Land Use Plans.

A. Requirement

The deliverable for ITQ IMB024 – Digital Data Backup, Archive and Retention Strategy is intended to address MSRM project data archive issues. IMB considers that the scope of a full data archive solution includes two components: a comprehensive strategy that must satisfy mandated MSRM requirements into the future; and, an interim solution to meet immediate MSRM operational requirements. This discussion presents the proposed interim solution.

a. Long-term

- The DDAS Strategy Document currently defines a comprehensive long-term, high-level strategy for addressing the archive component of SRM digital data management policy.
- The DDAS Implementation Plan presents the process by which the long-term strategy can be successfully delivered.
- The recommendation is an incremental implementation in manageable steps to successfully meet project requirements as implementation issues are addressed.

b. Short-term

- Neither deliverable document currently addresses the implementation requirements of an immediate short-term data archive solution.
- The first step in incremental implementation is a short-term solution to provide an interim remedy to regional operational problems.
- A short-term solution may be considered more tactical than strategic, but as such must contain a depth of detail sufficient to guide immediate implementation of a workable project data archive process.
- The need is to add this short-term solution as a sub-set of the long-term plan to meet project requirements and IMB and regional expectations with the caveat that implementation not stall at this interim step.
- The interim solution must deliver a process uniform amongst regions.

- Immediate benefits are a remedy to regional server capacity issues and a mechanism for knowledge retention, where the danger is loss of inconsistent project data and/or data not catalogued because of lack of process, and enhanced best practices.

B. Immediate Implementation

a. Mandatory Requirements

- The delivery of SRM backup, archive and retention/disposal strategies/policies relate to one part of data management policy. Procedurally they differ such that risk management requires backup tasks be performed separately from archive tasks.
- Critical to this project are planned server upgrades, Veritas backup hardware and software, DLT/SDLT multi-tape drives, and SRM / MOF SAN access.
- The intention is to deploy the Veritas NetBackup enterprise backup system from the SAN. Operational/configuration difficulties between UNIX and NT systems resulted in partial deployment of Veritas Backup Exec for work groups. Deployment of the enterprise system requires that the SAN be rebuilt. This is in progress and will be followed by a de-bug pilot prior to full deployment in mid-June; complete Veritas NetBackup functionality is planned for September 02, 2002.
- Three tape drives are deployed in Nanaimo, Prince George and Kamloops. The remaining three are scheduled for the week of May 06-10 in Nelson, Williams Lake and Smithers. Nanaimo and Surrey have DLTtape drives; all others have SDLTtape drives
- The MOF SAN is full; the SRM SAN is being rebuild. The possibility of regional SAN access for staging and/or temporary storage can only be considered after re-deployment.
- The interim regional project data archive solution is dependent upon and must align with the CSD completion schedule for Capacity Review II hardware and software deployment to regions. The current CSD plan schedules regional server and tape drive deployment for completion June 30, 2002.
- Where CSD Capacity Review implementation is constrained by technical problems, archive implementation will attempt to utilize current technology where possible.

-
- Currently evolutionary organizational structures in SRM and CITS will impact Victoria and regional responsibilities for archive implementation such that roles and tasks defined now are likely to change.
 - The tracking process is Veritas NetBackup Data Vault. This must include at least a raw digital data classification procedure, perhaps based on exiting flat-file data retention criteria, to identify digital data candidates, apply consistent and complete metadata, and guide the archive schedule. Depending on Data Vault functionality, there may be a requirement for a central Access metacontent database.
 - The recommended archive process must be consistently applied across all regions, with components extrapolated from regions with functioning regional archive procedures in place.
 - Interim regional strategy implementation includes a test phase in one regional data centre, probably Williams Lake, followed by fine-tuning and regional deployment.

b. Archive Process Model Options

- Distributed – responsibility for archive initiation and process is regional.
- Centralized – responsibility for archive initiation and process is in Victoria.
- Blended – responsibility for archive initiation and process is shared amongst Victoria and the Regional Data Centres.
- Recommendation: Distributed Archive Model.

c. Technical Options

- As a rule-of-thumb, the cost of disk systems doubles with tape; approximately \$36, 000 per year per terabyte of data.
- The business case for tape already exists as an extension of tape backup.
- Regional server space must be allocated as an archive staging/retrieval area.
- Existing backup – 22 Bay Tape Drive – given, all regions, for scheduled backup, allocate 2-3 bays for archive; archive and backup schedules cannot be coincident.
- Parallel system – requires a separate multi-tape drive per region dedicated to archive use (22bays x 120Gig = \$26,000.00 per region). The 7 bay tape drives (DLT 3570) replaced in the backup upgrade plus shared media licenses are dedicated to archive.

- Centralized system – tape drive units sufficient for archive and clone plus a server and shared media license are deployed in Victoria and accessed regionally.
- DLT Tape – type IV, 80 Gigabytes per tape – Nanaimo and Surrey use these.
- SDLT – 120 Gigabytes per tape – all other regions have or will have these – cost is \$200.00 per 120 Gigabyte.
- Usage management – fill tapes with many archive data sets or restrict to just one; data of like retention schedules can be put on the same tape; if all data is on the same retention schedule, then all tapes are always filled.
- Storage cost – on-site or off-site; there must be an archive copy on-site and a clone copy off-site.
- Recovery cost – \$25.00 handling fee per mount and dismount SAN access – cost of the full disk backup cycle (monthly, weekly, daily) is \$2.00 per Gigabyte per month.
- Implementation – the change from Legato Networker (UNIX) to Veritas NetBackup (Windows 2000) backup technology is encountering technical difficulties that will influence archive implementation, which should proceed, if necessary, independently of the Veritas rollout by utilizing the Legato product.
- Blended technology – some tape archive, some SAN archive.
- Recommendation: Blended technology – utilize the available backup product and redeployed 7-bay tape units for archive; SDLT tape, filled with same-retention data; complement with SAN access (when and if available), media type is dependent on data access and retention requirements; two copies – archive (on-site) + clone (off-site) – managed by the RSO under service agreement with CSD/CITS.

d. Staffing Options

- Operator cost – tapes = people; manual process, backup process requires approximately ½ day per week, archive adds to this, more so if multiple archive copies; SO or contract.
- Training cost – the distributed delivery model may involve system training that the centralized model does not; if the Veritas backup rollout includes a regional training component, archive can be incorporated with this.

- Recommendation: use RSO FTE time, deploy archive process parallel to the current backup function; minimize training requirement.
- System permissions; local authority to archive/restore; security.

C. Process Definition

a. File Structure

- Is one currently in use that can be applied across regions?
- Develop or modify existing appropriate to regional and Victoria server directory structure; common naming convention, common directory structure for datasets identified for archive (ArcDone), common directory structure for retrieval.
- Apply across regions.

b. Data Identification

- Where is it? – Server, CD, tape.
- What is it? Data type, format – Spatial, Attribute, GIS (Microstation, ArcView, Arc Info, etc.), database (Access, Excel, etc.).
- How big is it? – Volume.

c. Data Classification

- Raw classification based on data business requirement; 90% of data is spatial with temporal and historical value and as such the life cycle may never go beyond off-line retention.
- Initiates archive process; classification may be assigned at project inception.
- Eligibility – project life cycle, no access (4-6 months), data volume.
- Retention schedule – based on SRM document management criteria; default is retain everything always until process is refined.
- Medium & Access – tape, SAN; online, offline; local, Victoria; automatic, manual.

d. Data Tracking

- Tracking – data catalogue – data and location; metadata component; Global Data Manager; Veritas NetBackup Data Vault.

-
- The data tracking system must provide a catalogue of tape contents and metadata on the dataset content; if Veritas cannot provide this then there is a requirement for a supplemental, centrally managed, regionally populated database for metacontent.
 - Metacontent might include: spatial model, positional data accuracy, attribute data accuracy, data source, functional application & operating system, method of tiling, extent of coverage, data owner, and so on.
 - Complexities include the unknown questions for a long-term query requirement against the base line data – the more thorough the metacontent, the more meaningful the response to potential queries.

e. Procedure

- RIIM / GIS Coordinator initiates archive process based on data eligibility criteria.
- Data is moved to an archive directory structure for either tape or SAN archive.
- SO archives tape data as part of backup process; copies archive tape; stores locally and offsite – this should be longer term, low or no access required.
- System reader in Victoria identifies and moves SAN data – this can be data with an anticipated access requirement.
- Archive data is purged from regional server.

f. Schedule

- As required when target data is identified in archive directory.

g. Roles and Responsibilities

- Data management falls within the RIIM purview; archive data identification and classification procedures are applied the GIS Coordinator.
- Physical data archive processes are tasked to the regional RSO under service agreement with CSD/CITS.
- SAN operations are CSD/CITS.

D. Recommendation

a. Distributed Archive Model

- Responsibility for archive initiation and process is shared amongst Victoria and the Regional Data Centres.

b. Technology Options - Blended technology

- Promote vendor-independence so that the process can succeed with both Legato and Veritas products.
- Utilize a parallel multi-tape jukebox system dedicated to archive and retrieval.
- SDLT tape, filled with same-retention data.
- Complement with SAN access as needed and/or available.
- Medium type dependent on data access and retention requirements.

c. Staffing Options – SRM FTE

- Use regional RSO FTE time.
- Deploy archive process independently of the current backup function.
- Minimize training requirement.
- Negotiate service level and permission/trust agreements.

d. Process Tactics – consistent across regions

- Universal file/directory structure and file naming conventions.
- Regional dataset identification.
- Regional data classification to SRM data retention requirements.
- Mandatory automated/interactive data tracking.

e. Procedures

- As determined by SRM (Regions & IMB) and CSD and documented by IMB.

f. Schedule

- As required.

g. Roles and Responsibilities

- As determined and negotiated by RIIMs, IMB and CSD/CITS.