FINAL

Printed on 10/10/2001 9:21 AM

Report to

Geographic Data BC

Digital Image Management

Current Assessment—Future Direction

September, 2001

Sierra Systems
880 Douglas Street, Suite 500
Victoria, BC V8W 2B7 Canada
www.SierraSystems.com

Contact: Terry Tarle
Phone: 250.385.1535
Fax: 250.385.4761
Email: TerryTarle@SierraSystems.com

Date: June, 2001

Sierra
New thinking.

Change Record

Version	Date	Description
V3.0	July 04, 2001	Release of Document to GDBC for working committee review
V3.1	July 16, 2001	Integration of feedback from GDBC working committee
V4.0	August 08, 2001	Integration of feedback from GDBC working committee

TABLE OF CONTENTS

Appendices

EXECUTIVE SUMMARY

Geographic Data B.C. (GDBC) contracted Sierra Systems to assist with its' Digital Image Management (DIM) project, by assessing current business processes and recommending new processes and technology to increase digital image management efficiency within GDBC. The project was carried out in two stages: Current Situation Assessment and Future Direction. The findings and recommendations from this project are contained in this report.

Key Findings from Current Situation Assessment

The key findings from the current situation assessment are as follows:

- Key stakeholders expect GDBC to provide a reliable, efficient, and easy to access central warehouse of geo-spatial digital and hardcopy imagery products.
- Key stakeholders (digital image users) are not fully aware of what images are available at GDBC, or how to obtain these images.
- The demand for digital imagery products and "scan on demand" services will continue to increase, particularly for MOF staff following the recent decision to rollout softcopy stereo image viewing technology to all regions and districts.
- Current production levels for creating digital imagery at the GDBC Lab will **not** be able to keep pace with the expected increase in "scan on demand" services for key stakeholders.
- There is no formal policy on providing "scan on demand" services for other Government agencies (i.e. pricing, priority with archival work, turn around time, etc.). This service tends to be on a "squeaky wheel" basis.
- The current business processes involve many manual processes that could benefit from automation through the use of new technology.
- Image Scanners at the GDBC Lab are not efficiently utilized when compared to industry use of similar equipment (e.g. only one shift, idle at times, etc.).
- There is no Enterprise wide formal data management standards or policies in place for digital imagery (e.g. data custodian and stewardship, backup and recovery processes, data model, etc.)
- Scanned images are currently being archived at the GDBC Lab on CD.
- Due to the absence of a high-speed connection between GDBC and its' image providers (e.g. GDBC Lab, contractors, etc), images are currently being copied to CDs and other storage devices and couriered between sites.

Proposed Solution & Recommendations

The solution proposed in this report involves a significant change in *business processes* within GDBC, and implementation of new technology for digital image management.

Business Process Changes

The key recommendations for changes to the GDBC *business processes* are as follows:

1) Catalogue all imagery and make it available online via the Intra / Internet in three resolutions: full (raw), compressed and "thumbnails" (for online viewing).
2) Digitize (scan) all new aerial photography on acquisition either by making this a deliverable of the aerial photography acquisition contracts or through a separate contract.
3) Adopt U.S. Federal Geographic Data Committee (FGDC) Standards for Digital Geospatial Metadata. This standard is widely used nationally and internationally for the exchange of metadata, is considered best practises and is expandable for specific GDBC requirements. The Canadian federal government requires the adoption of this standard to qualify for some types of funding.
4) Fully review the current Land Data BC website functionality and determine the future direction of this facility, and plan integration with the solution proposed herein.
5) Implement a communication strategy to inform users what digital imagery is available, ownership and restriction of use, and access method, to encourage them to utilize the GDBC Image Data Store.
6) Establish a single (one window) method of processing all Government and non-Government customer requests for digital imagery (i.e. no back door or exception processing).
7) Establish a policy for value added products such as Orthophotos and specialty services such as "Scan on Demand" provided by GDBC staff. At a minimum, this policy should address the products and services to be provided, price, and priority with respect to scanning of archival aerial photography.
8) Investigate options to increase the utilization of the existing scanners at the photo lab to accelerate the digital capture and archival of aerial photography film and to increase the return on capital investment.
9) Investigate a private sector partnership for image scanning.
10) Investigate private sector initiatives for the creation, marketing and distribution of value added digital image products and specialty services.
11) Establish a data custodian and data steward for all digital image products.
12) Make all thumbnail digital imagery owned by the BC Government available to all users free of cost. This would apply only to imagery that is not encumbered by 3rd party licensing agreements or copyright.

New Technology Proposed

The new technology proposed in this report has been selected to adhere to the requirements and technology infrastructure recommendations contained in the *Data Management Architecture Review and Recommendations* report recently completed by Holonics Data Management Group Ltd. Highlights of the proposed technical solution are as follows:

- Establishment of a high-speed connection to imagery originators (e.g. contractors, GDBC photo lab, etc) where possible.
- Implementation of an Image Data Store technical infrastructure that will support the efficient and secure storage and online retrieval of all digital images.
- Implementation of Standards-based Commercial Off The Shelf (SCOTS) software product(s) that will meet the mandatory requirements of the four primary business processes (digital imagery Acquisition, Image Data Store Management, Discovery and Distribution).

Companion Reports

There are two reports that accompany this report.

The *Software Evaluation Criteria* report provides software standards, requirements and evaluation criteria to be used for the selection of SCOTS software product(s) for image management at GDBC.

The *Migration Strategy* report provides a "roadmap" for implementation of the proposed business process and technology. The proposed strategy provides a multi-phased approach structured to provide GDBC with tangible benefits and improvements in efficiency at the completion of each phase.

1. INTRODUCTION

Geographic Data B.C. (GDBC) currently manages a large number of digital images of various types including; aerial photographs scanned for archive and mapping purposes, scanned or "rasterized" maps, satellite imagery, and derived digital image products such as digital orthophotos.

GDBC has identified several inefficiencies in the way these images are acquired, managed and stored, accessed and distributed to both internal and external customers. The Digital Image Management (DIM) project was initiated to investigate ways of improving the efficiency of managing digital imagery through changes to business processes and new technology.

GDBC at this point has moved through at least two major activities related to DIM. A "Data Management Architecture Review" has been completed and now "Current Assessment and Future Directions" (this project) is in progress.

The *Other Related Issues* section describe a possible higher level organization and goal set for DIM than is focused on in this current project. The *Current Assessment & Future Directions Project* section describes the project for which this report is a deliverable.

1.1. Current Assessment & Future Directions Project

GDBC contracted Sierra Systems to assist the DIM project by:

- Assessing current business processes and infrastructure;
- Recommending new processes and technology;
- Assisting with planning for new technology to increase digital image management efficiency within GDBC.

This project was carried out in two stages – *Current Situation Assessment*, and *Future Directions Planning*. The results of this project were this document and two others (*Software Evaluation Criteria* and *Migration Strategy*) that together provide a tactical plan of action to effect change.

1.1.1. Objectives

The objectives of the *Current Assessment-Future Directions* DIM project are:

- Identify and document Geographic Data BC, digital image management procedures and requirements,
- Incorporate MELP/MoF HQ and Region digital image management procedures and requirements where they do not conflict with GDBC requirements,

- Incorporate the image management process within the Data Management Review Strategic Plan and technical architecture,
- Prepare a 'Road Map' or migration strategy outlining the steps required to achieve the implementation and goals of the *Current Assessment-Future Directions* project. This road map will establish a list of staged projects for migration from the current business structure, strategy and infrastructure to a business and technical architecture that meets the goals established within the *Current Assessment-Future Directions* project and ensures compliance with the Data Management Review Strategic Plan.
- Prepare an evaluation plan to test, compare, and rank candidate vended products and services against *Current Assessment-Future Directions* requirements.

1.1.2. Scope

For the purposes of the current DIM project, Digital Image Management comprises all management activities related to the creation, acquisition, assembly, updating, verification, storage, manipulation (including conversion and transformation), access control, representation, distribution, and archiving or removal of the following digital image products:

- Scanned aerial photos for archive- Black and White, Colour and Infrared
- Scanned aerial photos (stereo-models) for TRIM or other softcopy photogrammetric mapping (includes "scan on demand")
- Digital aerial imagery – Monochromatic, Colour or Multi-spectral
- Digital Orthophotos – Black and White, Colour or Multi-spectral
- Digital Mosaics
- Raster versions of vector maps – (e.g. TRIM vector maps merged and exported to TIFF format).
- Scanned index maps or other rasterized vector data.
- Digital Satellite Images (various sources and formats)
- Baseline Thematic Mapping products
- Scanned maps (e.g. NTS Tiles)

The scope of the *Current Assessment & Future Directions Project* does not include the following:

- Non-digital imagery and maps
- Non-technical distribution issues such as pricing policy or e-commerce partnerships.

1.1.3. Purpose of Report

The purpose of this report is to document the findings and recommendations of the *Current Assessment-Future Directions* project. This document progresses from a current situation

assessment through business requirements to new business processes and requirements and to new system requirements.

1.1.4. Intended Audience

This report will be of interest to all GDBC staff involved in the collection, management and distribution of digital images, and also to interested key stakeholders.

The intended audience for the *Proposed Solution – Technical Architecture* is GDBC Management and GDBC Information Systems Management.

Please note that Appendix D, *Stakeholder Requirements Summary Table* is not recommended for general circulation.

1.1.5. Analysis Approach

The project was approached in 2 stages – *Current Situation Assessment* and *Future Direction*.

1.1.5.1. Current Situation Assessment

The current situation assessment stage of the project involved the following activities:

Project initiation

A Project Charter and plan was created. The Project Charter details the Objectives and Scope of the project and also determines methods for resolving changes, mitigating risks and an overview of the project.

Information Gathering

Two workshops were held during the project. One workshop was held with key GDBC staff to determine current processes at a high level and investigation of desirable vision requirements. A series of one on one interviews were then conducted to confirm workshop findings and to create 'as-is' business process models.

A second workshop was held with the key stakeholder representatives of Ministry of Environment, Lands and Parks (MELP) and the Ministry of Forests. The participants of the workshop provided feedback of the workshop findings.

See appendices: C - Client Workshop Notes, D – Stakeholder Requirements Summary Table, E – GDBC Workshop Notes

Document Current Business Processes

The current business process model diagrams were created and reviewed by GDBC staff. Changes to the business process models were made as required.

See appendices: A – Current Business Process Models

Report on Findings

The Current Assessment section of the final report was created.

Steering Committee Checkpoint

A presentation to the Steering Committee was conducted. The presentation consisted of a review of the project deliverables, findings to date, proposed approach and next steps.

1.1.5.2. Future Direction

The future direction stage of the project involved the following activities:

Research

A review of the *Data Management Architecture Review and Recommendations* report was conducted to understand the future direction of GDBC with respect to data management and technology architecture, and to ensure that the proposed solution was in line with the recommendations contained in this report.

Sierra Systems also conducted research on industry trends, geo-spatial data standards, and GIS technology through Internet searches, and gathered information from recent best practices research and other relevant Spatial IT projects carried out for other customers with similar Geo-Spatial data management problems.

See appendix: G – Storage Media and Infrastructure Research

Analysis of information

The information from the current situation assessment and the research was then analyzed to identify opportunities of improvement or efficiencies. Some quantitative analysis on process was conducted on image storage procedures and the impact on the business processes. A qualitative analysis was also conducted to determine and rank requirements regarding Digital Image Management.

See appendix: F – Digital Imagery Storage Comparison Analysis

Develop Proposed Solution & Recommendations

A Proposed Solution was developed. This solution included the development of "To Be" Business Process Models and recommendations, and Proposed Technical Architecture and "Vision" for a new Digital Image Management system.

See appendix: B – Proposed or "To Be" Business Process Models (Microsoft Visio Document), H – Summary of the FGDC Standards, I – Overview Details of the Z39.50 Protocol, J – Overview Details of OGC Standards for Interoperability

Migration Strategy

A supporting Migration Strategy report was created to provide a high-level conceptual "road map" for a phased migration from the current situation to the "Vision" system.

Software Selection Criteria report

A report was created outlining the standards, requirements and evaluation criteria to be used for the selection of SCOTS software product(s) for the target digital image management system.

Final Reports

Final drafts of the three reports (this report plus the Migration Strategy and Software Selection Criteria reports) were created after review and feedback from GDBC staff.

1.1.6. Companion Reports

Two reports accompany this document:

- The *Software Evaluation Criteria* report provides software standards, requirements and evaluation criteria to be used for the selection of SCOTS software product(s) for image management at GDBC.
- The *Migration Strategy* report provides a "roadmap" for implementation of the proposed business process and technology. The proposed strategy provides a multi-phased approach structured to provide GDBC with tangible benefits and improvements in efficiency at the completion of each phase.

2. CURRENT SITUATION ASSESSMENT

2.1. Introduction

This section of the report documents the image management processes currently in place at GDBC.

The objective of this stage of the project was to gain an understanding of the existing processes that may impact current and future management of digital imagery. The documentation of these current business processes is also required to communicate the impact of change.

All in scope existing business processes were assessed and documented. Detailed business process models for all of the current image management processes were created using MS Visio, and are included in Appendix "A" of this report. The models that do not impact digital imagery management are included in the appendix, but are not referred to again in this document.

2.2. Current Assessment Approach

Information required for the current situation assessment stage of the project was gathered through a workshop and interviews with staff currently involved with digital imagery within GDBC. These processes were documented through a series of work flow diagrams that were confirmed by the respective GDBC staff.

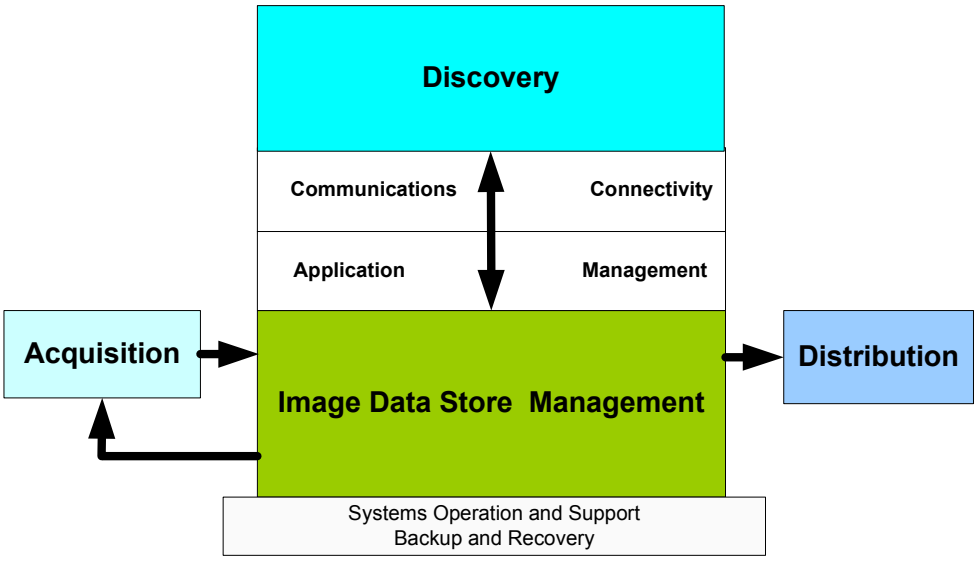
An additional workshop was held with Ministry of Environment, Lands and Parks and Ministry of Forests stakeholders (digital imagery users) from Victoria and several regional offices, to determine their requirements, issues and concerns. These requirements, issues and concerns were documented and categorized under four functional components or modules.

2.2.1. Logical Components

To address the objectives of the project in a systematic manner, the overall digital imagery management function is divided into 4 logical components or modules.

The following four functional modules are used throughout both stages of this project, to address the current situation assessment and also to address the proposed solution.

These modules are illustrated in the following diagram and further described below:



Imagery Acquisition

Imagery Acquisition encompasses all processes involved in the receipt and preparation of images for storage within the Image Data Store. The preparation of the imagery includes workflow procedures to load the imagery into the Image Data Store and the metadata into the metadata database. It also includes the creation of products derived from the raw or base imagery. The Acquisition process tracks all steps of the development and ensures that identified interim images will be stored in the Image Data Store, and be available for future requirements.

The acquisition process module follows a project from its initiation to the creation of the final image products. Acquisition also recognizes that due to technology limitations the images may be received at GDBC in a variety of storage media. The Acquisition process module populates the Image Data Store with the images, metadata and supporting information.

Image Data Store Management

Image Data Store Management involves the storage, safekeeping, and tracking of digital images, metadata and supporting information. The Image Data Store Management component contains

and enforces business rules concerning access, processing, security, modifications and offline archiving. It also includes Image Data Store backup and recovery procedures.

Imagery Discovery

Imagery Discovery denotes the processes and technology required for internal and external customers to search for specific images or groups of images via predetermined and ad-hoc search criteria. The Discovery process ultimately depends upon the existence of metadata to support such searches. Current Discovery facility does not support on-line geographic "area of interest" search but does support some textual based searches through LDBC facility. Manual geographic "area of interest" search is enabled through flight line perusal.

The Discovery module interfaces with the Distribution module by triggering the process for Distribution. Also triggered may be an Acquisition event requiring the creation of a derived product or service (e.g. the scanning of an analogue image currently on file). The Discovery module is not used for the initiation of contract acquisition of new aerial photography.

Imagery Distribution

Imagery Distribution refers to the processes to deliver selected products to a customer. Distribution of specific images and supporting files has multiple variances and is dependent upon existing technology format, image file size requirements and the ability to 'transport' the image and metadata to the user. Transport is impacted by technology infrastructure and may be online (currently this is rare or non-existent, for compressed images only) or via a portable storage device, primarily CDs.

2.3. Current Architecture and Asset Inventory

The *Data Management Architecture Review & Recommendations*, Holonics Data Management Group Ltd., 22 March, 2001 inventories the existing infrastructure.

Readers are encouraged to review the "As-Is" inventory in the above named document. What follows is a precise of the inventory.

Information Architecture

- Various metadata is held in LDBC, APS, Excel spreadsheets and within MASCOT.
- TRIM and BTM workflow is held in Excel. APS and MASCOT track workflow relevant to their own data management activities.
- TRIM, DRA, EBM and gridded DEM are managed as files. Analogue and hardcopy material are managed offline, as are scanned air photos and satellite images. MASCOT, BCNGIS and TWA manage content within the applications. Most other assets are managed as files, offline.

Storage

- Oracle is used to hold LDBC metadata, BCNGIS, MASCOT(new), and APS.
- Informix is used to store TWA and MS-Access holds Watersheds BC.
- Excel files are used to hold TRIM scanned photo catalogue, TRIM contract and process management and the Ortho catalogue.
- Air Photos and Reprographic mylars are primarily hardcopy at this time.
- All other data (TRIM, BTM, toponymy, Air photo index, etc) are in files.

Servers

- SolRep & Gamma are two Sun ES3500 servers used to house production Oracle databases and have a total of 650 GB of file storage, all of which is allocated to current needs.
- Three other sun ES2000 and ES3000 servers and several Compaq servers are in place to serve Oracle CASE Repository, research, web server and file server needs.

Software

- Oracle DBMS, middle layer components, application development tools and CASE.
- ESRI ArcView, MapInfo, SAIF/FME, CARIS, PAMAP
- Trimble URS/Office, PCI, Flight Line Design, Micro Station, MAPS 3D
- ZI Photoscan TD
- LDBC Interface

2.4. Current Business Process Models

The current business process models were developed as a result of direct interviews with GDBC staff. All process models have been reviewed and approved by the relevant GDBC staff members.

The business processes that deal with digital imagery are outlined in the following sections:

2.4.1. Scan on Demand

The detailed graphical model of this process is included in Appendix "A". An overview of the process follows:

13) Customer places an order for a digital product that requires scanning of aerial photography.

14) GDBC validates the order and sends copy to the Reprographics lab
(* Note orders placed via LDBC are not validated).

15) Reprographics lab locates film roll/frame numbers specified and prepares film to be scanned.

16) The film negative is then scanned by the following steps

    a. Preview is performed to view and adjust density parameters.

    b. Scanning Parameters are set.

    c. Black and White photography is scanned at 14 microns; Colour photography is scanned at 21 microns.

17) The scanned image is saved on the workstation's hard drive and quality assured to specifications.

18) The scanned file name is entered. The convention is roll number/frame number: (i.e.) BCC736_124.TIF.

19) Quality assurance steps are performed.

20) Metadata collected includes workstation name, resolution, roll/frame# and date of scan. This information is recorded on an Image Tracking paper form.

21) The scanned image is copied to the "burner" workstation where it's copied to CD.

22) CD image is then verified on two separate workstations: one for image verification, the second for data verification.

    a) If the image fails either of the quality assurance stages, the images are copied on another CD, which then undergoes the same Q/A procedure.

    b) If the image passes the quality assurance the date and name of the Q/A person is entered in the Image Tracking paper form.

23) Image is deleted from scanning and 'burner' workstations.

24) CD labels are generated indicating the roll number and frame numbers of the scanned images (i.e.: BCC736 124-125).

25) Digital images are couriered as follows:

    a) Raw image orders are sent directly to the customer.

    b) Scans for Orthophotos or stereo models are sent to GDBC head office for further processing.

26) The order is marked as 'closed'.

    a) A hardcopy of the order is kept for lab records.

    b) LDBC is updated so that order becomes closed.

2.4.2. Archival Scanning

GDBC has a mandate to scan the older five-inch aerial photography frames, and this is currently being done in a chronological manner. Occasionally an older film roll is part of a customer order,

and the deterioration of the film is such that the roll is then scanned. This is the only exception to the chronological method.

The detailed graphical model of this process is included Appendix "A". An overview of the process follows:

1. The next film roll/frame number sequence to be scanned is identified and retrieved from the film library.
2. If no film roll exists, the existing prints will be used.
3. The media is prepared for scanning (this includes a visible check of the frame or print and laying it out on the scanner).
4. A scan preview is performed to view and adjust density parameters.
5. Other parameters are set. The technical scanning procedure involves tiling and other set up modifications before the image is finally scanned.
6. Black and White (all older 5" film) scans are done at 14 microns.
7. The scanned image is saved on the workstation's hard drive where it's quality assured to specifications. File naming convention is the roll number/frame number – (i.e. BCB41_078.TIF).
8. The Image Tracking paper forms are completed and stored in a binder. The information collected includes workstation name, resolution, roll/frame#, date of scan and any comments on print/roll/frame quality.
9. Quality assurance procedures are then used to certify the scanned imagery.
    a) If the image fails Q/A, the frame or print is scanned again.
    b) If it passes Q/A and the number of saved scanned images on the scanner workstation's drive has reached the predetermined quantity, the images are copied to the "burner" workstation where the CDs are created.
10. The images are copied to 2 separate CDs (a working copy and an archive copy).
11. Both copies are verified on two separate workstations: 1 for image verification, the second for data verification.
12. Quality assurance testing is performed on the created CDs.
    a) If the image fails either of the quality assurance stages, the images are copied on another CD, which then undergoes the same Q/A procedure.
    b) If it passes the Q/A stage, the date and name of the Q/A person is entered in the Image Tracking paper form.
13. The images residing on both the scanning and CD burner workstations are deleted.

14. CD Labels are generated, indicating the roll number and frame numbers of the scanned images (i.e.: BCB41 078-084).

15. Both copies are stored separately in the vault at the reprographics lab.

2.4.3. Orthophotos

The detailed graphical model of this process is included Appendix "A". An overview of the process follows:

1) A customer (usually a Ministry of Forests region) requests an Orthophoto that currently does not exist at GDBC.

2) If the order is validated, the Reprographics lab receives it, and then either generates the scans of specified roll/frame numbers (based on the Scan on Demand procedure) or produces a set of hardcopy diapositives for the Image/Mapping Contractor. (For all TRIM 2 photography, the Reprographics lab will have produced 1 set of diapositives.)

3) If diapositives are generated, these are Quality Assured at the lab, then delivered to GDBC head office.

4) GDBC receives the hardcopy diapositives or scanned imagery, then sends these to image contractor for further processing. If the lab delivers hardcopy diapositives, the image contractor generates the scans.

5) If scans are created and delivered by the Reprographics lab, metadata (project#, roll/frame#, region, etc.) is captured and entered into an Excel spreadsheet by GDBC staff.

6) Image contractor generates the orthophotos and support files, and delivers four copies of each on CD to GDBC.

7) The four files are copied to each CD: COT, TIF, and the ASCII IRP and TFW files. Naming convention used is based on the map sheet they represent (i.e.: 93B054.TIF).

8) If the deliverables from the image contractor include scans from diapositives, metadata is entered into an Excel spreadsheet. Fields include contractor name, project #, roll/frame numbers, AT codes, media type, format, etc.

9) Orthophotos are loaded on a workstation and checked for accurate orthorectification and map sheet coverage. The working copy is affixed with a blue dot signifying it has been verified.

10) Metadata concerning the orthophotos are entered into an Excel spreadsheet. Fields include branch inspection information, contract name, project number, roll/frame numbers of photos, etc.

11) If the CDs pass the Quality Assurance process, 3 additional copies are distributed (1 to distributor, another to the respective Forest Regional Office and one to the Regional distributor). The remaining copy is stored at GDBC in CD file cabinets.

12) Metadata concerning the distribution is entered into an Excel spreadsheet. Fields include agency and contact person, location, status, distribution #, file names, etc.

13) If the image contractor used hardcopy diapositives, they are returned to GDBC and stored at the Blanchard office.

14) LDBC is then updated as to the availability to new Orthophoto(s) – an Excel spreadsheet is used to maintain catalogue information, fields include map sheet name and quantity.

2.4.4. Softcopy Stereo Models

The detailed graphical model of this process is included Appendix "A". An overview of the process follows:

1) A customer requests a new softcopy stereo model.

2) The order is received by the Reprographics lab, which then either generates the scans of specified roll/frame numbers (based on the Scan on Demand procedure) or produces a set of diapositives for the Image/Mapping Contractor.

3) Scans/diapositives are Quality Assured at the lab, then delivered to GDBC Head Office.

4) GDBC sends the air photography scans or Diapositives to the image contractor for further processing. If the lab delivers Diapositives, the image contractor generates the scans.

5) The image contractor generates the stereo models and support files from the scanned photos, delivers 2 copies of each CD to GDBC (one for storing, the other for distribution). Up to four files may exist for each photo model on CD: a SIS, SDT, MOD and a DGN file.

6) GDBC receives the stereo model(s) from the contractor on CD, the original scanned images (either lab or contractor generated) and, if they were generated, the diapositives. Metadata is captured and maintained in an Excel spreadsheet by GDBC staff (describing the digital stereo models). Data fields include contractor name, project #, roll/frame numbers, image size, sheet#, format, etc.

7) Both CD copies of the stereo models are quality assured by image loading on a workstation to ensure accurate orthorectification and map sheet coverage. The working copy is affixed with a blue dot signifying it has been verified and stored in CD file cabinets.

8) The other copy is distributed to the requesting agency.

9) LDBC is updated of the availability of the new stereo models.

2.4.5. Satellite Imagery

The detailed graphical model of this process is included Appendix "A". An overview of the process follows:

1) A customer requests a product derived from satellite imagery.

2) The raw satellite imagery is ordered. GDBC Digital Image Applications Unit contacts the Image Contractor.

3) The ‘ceonet/nrcan’ (now 'GeoConnections') Website is queried for available imagery (large scale - IRS/SPOT or small scale - Landsat TM7).

4) The Image Contractor obtains the raw satellite imagery on behalf of GDBC.

5) The raw satellite imagery is quality assured and processed accordingly by the Image Contractor.

6) The processed satellite imagery is quality assured by the Image Contractor.

7) Quality assured digital files (the derived product, source imagery and any support files) are delivered on CD or tape to GDBC.

8) GDBC receives the digital files on CD or tape and carries out their own Q/A process on the data. The imagery is checked for file corruption, accurate georeferencing and colour balancing.

9) If the imagery fails quality assurance, the image contractor is notified accordingly.

10) If imagery passes quality assurance, metadata is captured about the image and entered into an Excel spreadsheet. Data fields include track and frame number, scene date, bands, centre lat/long, storage media type, etc.

11) If further imagery processing is required, the required data is sent out to Image Contractor and/or processed in house. This newly processed data is quality assured as above.

12) The satellite imagery derived product(s) are copied to a Solaris file server. The original tape(s) or CD(s) containing raw satellite imagery and the processed product(s) with support files are stored in file cabinets or tape boxes.

13) The product is distributed to the customer.

14) If the customer requests the product on CD, a copy is made, quality assured then delivered.

15) If the product is not requested on CD, the image is delivered via FTP.

16) MELP is notified of new derived image product and raw satellite imagery availability.

3. PLANNING FOR CHANGE

3.1. Motives, Priorities and Opportunities

GDBC is substantially changing the way it manages digital images, and also is substantially changing the way it interacts with GDBC customers.

This change is motivated by increasing GDBC customer demand for simple but effective facility for discovering, ordering, paying for and obtaining GDBC products. GDBC as an organization also has a strong desire to more efficiently manage increasingly diverse products and services.

The constantly decreasing cost of computer and communications hardware is another motivator. Investment now or in a one to two year time frame into systems infrastructure has clear benefits, and is expected to have reasonable return on investment.

GDBC must guard itself against changes in the international software marketplace, to reduce risk of lack of software vendor support and to reduce potentially high costs of systems and application maintenance and change. A period of significant change such as GDBC is now going through is an appropriate time to adopt standards based components available from a non-monopolistic marketplace, thus increasing component longevity and decreasing purchase and support costs.

Management expects to meet in the short term 80% of GDBC's and GDBC's customers' needs and requirements for optimized systems. It is recognised that implementing new processes and systems to meet the remaining 20% will be time consuming and difficult. This need will not be ignored, but the concentration of limited resources will be on activities with a clear and immediate benefit. An analysis of critical success factors should be incorporated to determine the 80/20 rule.

GDBC management has stipulated that the first priorities and focus for resource investment will be the management of and customer interface for aerial photos and orthophotos. Second will be raster maps and satellite imagery. Third will be orthomosaics and stereo models.

3.2. Stakeholders Workshop

GDBC determined that involvement of its major Stakeholders or customers would be critical to the success of the DIM project. These stakeholders were involved through participation in a one-day workshop. The primary intent of the workshop was to provide a forum for the stakeholders to:

- Gain an understanding of the DIM project and goals.
- Provide the DIM project with an understanding of the customer requirements.

- Inform the DIM project of related initiatives that may be ongoing within the customers business and allow GDBC to make appropriate considerations of these initiatives.
- Discuss Digital Imagery specifications across Ministries for compression, standards and usability.

3.2.1. Workshop Participants

The participants of the Stakeholders' workshop and the agency they represented are shown in the following table:

Stakeholder Name	Agency
Brian Howden	MOF - Victoria
Ann Morrison	MOF - Smithers
Graeme Weir	MOF - Victoria
Mike Worgan	MOF - Invermere
Bill Cumming	MOF - Castlegar
Enrique Sanchez	MOF - Williams Lake
Ole Joergensen	GDBC (MELP)
Bill Anderson	GDBC (MELP)
Brian Cavanagh	MELP – Nanaimo
Jennifer Ballantine	MELP – Williams Lake
Bruce Mackenzie	MELP – Victoria
Dave Clark	MELP – Victoria
Evert Kenk	LUCO

3.2.2. Feedback from Stakeholders' Workshop

The workshop notes and summary were submitted in a separate report. The report is included in Appendices "C" and "D". A draft report was sent to the workshop participants for review, and their feedback was incorporated into the final report. Feedback from the workshop was documented in this report, and initially categorized under the following headings: (The bracketed numbers following each heading indicates the number of issues identified for this category.)

- Communications (2)
- Business Requirements (8)
- Technical Requirements (26)
- User Requirements (16)
- Out of Scope (18)
- Other Observations (15)

From the workshop discussions it was clear that the stakeholders presently view GDBC as a service organization that provides digital and analogue imagery. In fact, the stakeholders have a clear expectation that GDBC provide a reliable, efficient, and easy to access warehouse of geo-spatial digital imagery that removes some of the burden of management of these assets from the customer. The stakeholders also clearly appreciated the technical difficulties associated with spatial imagery management, and did not voice unachievable expectations.

GDBC's' customers stated that their business requirements focus on the ability to easily acquire information about products and processes (metadata). MOF staff participants indicated that their demand for digital imagery was expected to increase with the recent decision to rollout softcopy stereo image viewing technology to all regions and districts.

3.2.3. Stakeholders' Requirements

Following the workshop, the requirements identified by the stakeholders were categorized under communications issues plus the four primary component modules (i.e. Acquisition, Image Data Store Management, Discovery and Distribution). These requirements are as follows:

Communications Issues:

- Clearly define GDBC's mandate and responsibilities.
- Clearly define the rights and usage limitations regarding digital imagery.
- Clearly demonstrate how GDBC is going to meet the needs of their stakeholders at present and in the future.

Acquisition:

- Customers will accept lower resolution images as suitable for many intended uses.
- The customers and GDBC need to determine an optimal scanning resolution (quality vs. image size and related transmission and system performance).
- The customers want an online, searchable inventory of all images – digital and hardcopy.
- The customers want access to an online, single point-of-contact "Scan on Demand" service through GDBC.

Image Data Store Management:

- Determine and publish all metadata, associated with the catalogue of products.
- Distribute all relevant metadata bundled with the image.
- There is an anticipated increase in demand for digital imagery due to then MoF decision to implement diapositive viewers at district and regional offices (SIS format).

Discovery

- An online discovery facility is required.
- The discovery tool needs to be intuitive.
- Ability to find all images (digital and hardcopy).
- Ability to find images via flight indices.
- Ability to use image 'foot prints'.
- Ability to search by geographic 'bounding box' to find specific images.
- Ability to use textual searches/filters for images (e.g. Year, Scale, etc.).
- Ability to use geographic names or toponymy for search criteria.
- Ability to view/order clips of images at re-sampled resolutions.
- Ability to order multiple images at one time.

Distribution

- Ability to provide images in several industry-standard formats.
- Ability to provide online display of certain classifications of images for online use and reference.
- Ability to distribute online or via CD, DVD, portable hard drives, etc.
- Ability to compress images with either lossless or acceptable loss of image quality and integrity.
- Ability for the customer to track orders for images by order and customer name.

3.3. Project Team Observations & Findings

3.3.1. Observations

Currently, GDBC has established a set of business procedures and methods to support the production, archiving and distribution of digital imagery. The current processes are constrained by technology. The digital products that GDBC deals with are by their very nature dependent on technology.

The decision to utilize write-able Compact Disks was a very valid and cost effective decision at the time. The only practical way to archive and distribute very large digital imagery was either in tape format or CD. The tape format is proprietary and would lead to logistical problems for the various consumers of the imagery, the CD format is 'open system' and was the best format to achieve the distribution and archiving requirements.

The GDBC staff is very customer oriented and each member interviewed was quality and service oriented. Most departments are engaged in these activities without sufficient budgetary support

to deliver the level of service they would like. This leads to a high stress level within GDBC for both personnel and equipment, but is a common aspect and a fact of life within today's government.

Based on a review of the current digital image management business processes, the project team has made the following observations:

- The current business process model involves many manual processes that could benefit from automation through the use of new technology.
- Current record keeping, filing and tracking of digital image products is done manually, and is susceptible to human error or loss of data.
- Key stakeholders expect GDBC to provide a reliable, efficient, and easy to access warehouse of geo-spatial digital and hardcopy imagery products.
- Key stakeholders (digital image users) are not fully aware of what images are available at GDBC, or how to obtain these images.
- The demand for digital imagery products and "scan on demand" services will continue to increase, particularly for MOF staff following the recent decision to rollout softcopy stereo image viewing technology to all regions and districts.
- There is no formal policy or procedure for providing "scan on demand" services for other Government agencies (i.e. pricing, priority with archival work, turn around time, etc.). This service tends to be on a "squeaky wheel" basis.
- Current production levels for creating digital imagery at the GDBC Lab will **not** be able to keep pace with the expected increase in "scan on demand" services for key stakeholders.
- Image Scanners at the GDBC Lab. are not efficiently utilized when compared to industry use of similar equipment (e.g. only one shift, idle at times, etc.).
- There is no GDBC or MELP wide formal data management standards or policies in place for digital imagery (e.g. data custodian and stewardship, backup and recovery processes, etc.).
- Scanned aerial photo images are currently being archived at the GDBC Lab. onto CD.
- Due to the absence of a high-speed connection between GDBC and its' image providers (e.g. GDBC Lab, contractors, etc), images are currently being copied to CDs and other storage devices and couriered between sites.

3.3.2. Project Team Recommendations – Business Processes

Based on a review of the current digital image management business processes, the project team has made the following recommendations:

1) Catalogue all imagery, and make it available online via the Intra / Internet in three resolutions – full (raw), compressed and "thumbnails" (for online viewing).

2) Digitize (scan) all new aerial photography at the acquisition phase of the business by making this a deliverable of the aerial photography acquisition contracts.

3) Adopt U.S. Federal Geographic Data Committee (FGDC) Standards for Digital Geospatial Metadata. This standard is widely used nationally and internationally for the exchange of metadata, is considered best practises and is expandable for specific GDBC requirements. The Canadian federal government requires the adoption of this standard to qualify for some types of funding.

4) Fully review the current Land Data BC website functionality and determine the future direction of this facility, and plan integration with the solution proposed herein.

5) Implement a communication strategy to inform users what digital imagery is available, what the ownership and restriction of use stipulations are and what the access method is, and to encourage them to utilize the GDBC Image Data Store.

6) Establish a single (one window) method of processing all Government and non-Government customer requests for digital imagery (i.e. no back door or exception processing).

7) Establish a policy for value added products such as Orthophotos and specialty services such as "Scan on Demand" provided by GDBC staff. At a minimum, this policy should address the products and services to be provided, price, and priority with respect to scanning of archival aerial photography.

8) Investigate options to increase the utilization of the existing scanners at the photo lab to accelerate the digital capture and archival of aerial photography film, and increase the return on capital investment.

9) Investigate a private sector partnership for image scanning.

10) Investigate private sector initiatives for the creation, marketing and distribution of value added digital image products and specialty services.

11) Establish a data custodian and data steward for all digital image products.

12) All thumbnail digital imagery owned by the BC Government should be made available to all users free of cost. This would apply only to imagery that is not encumbered by 3rd party licensing agreements or copyright.

3.3.3. Project Team Recommendations – Technical Infrastructure

Based on a review of the current digital image management business processes, the project team has also made the following recommendations regarding technical infrastructure:

- Establish a high-speed connection to imagery originators (e.g. contractors, GDBC photo lab, etc) where possible.
- Implement an Image Data Store technical architecture that will support the efficient and secure storage and online retrieval of all digital images.
- Purchase and implement Standards-Based Commercial Off The Shelf (SCOTS) software product(s) that will meet the mandatory requirements of the four workflow (Digital Imagery Acquisition, Image Data Store Management, Discovery and Distribution).

3.4. High–level DIM Requirements

The requirements listed here are high-level and are expressed solely to provide a basis and justification for Technical Architecture formulation, which in turn enables software and hardware selection. This is not intended or required to be a complete or detailed requirements list.

Where decisions regarding business or technical aspects of the system have not yet been made, the requirements are written generally to support all prospective alternatives.

Further detailed definition of requirements for user interfaces, applications, storage, system functionality, metadata and other components of a digital image management system will be made in later phases and GDBC projects.

BC In View, (HTTP://home.gdbc.gov.bc.ca/BCInView/Default.htm)GDBC 01/06/21 provides a detailed list of requirements for an Air Photo Ordering System. Air photos are related to over 90% of the product and service to be provided by the Digital Image Management systems, and so this document provides a foundation for DIM Product and Service Requirements.

Refer to Appendix C, Client *Workshop Notes* and Appendix D, Stakeholder *Requirements Summary Table* for further detailed requirements of aspects of DIM.

3.4.1. Business and Technical Axioms

Note: Axioms are stipulations made by or through GDBC that are accepted as true and necessary, without further analysis or justification.

- At least 80% of the data management and data delivery requirements are to be met using Standards-based Commercial off the Shelf (SCOTS) products.
- The technical architecture will be modeled as an n-tier system with components rather than as a monolithic system.
- Customer interfaces are to be implemented through HTML/Java web-style browsers, and are to be generally as thin as possible.
- The Oracle DBMS is to be used for underlying data storage and management by all application and server level components. The actual image data, however, may be stored as files in either a NTFS or UFS, under the management of an application that uses Oracle DBMS services to control access and input/output to files.
- Microsoft and Sun Microsystems operating systems are to be used exclusively and in a cooperative environment.
- GDBC technical systems and GDBC customer systems must communicate using Canadian or internationally accepted standards for communication, content and format. This removes barriers to the Internet world in general when interfacing with GDBC.

- Purchased or developed GDBC technical systems components must communicate (internally at GDBC) using Canadian or internationally accepted standards for communication, content and format. This does not preclude packaged or "one-source" solutions as long as package components can be shown to be internally conformal. This removes barriers to general component purchasing and unties GDBC from specific vendors.
- Systems facility, software tools and subject databases must be available to all GDBC and MELP staff limited only by GDBC security policy, intra/internet accessibility and by vendor license agreements.

3.4.2. Product and Service Requirements

The business goal of the Digital Image Management systems is to facilitate the provision of product and service to GDBC customers related to the following:

- Scanned aerial photos for archive- Black and White, Colour and Infrared
- Scanned aerial photos (stereo-models) for TRIM or other softcopy photogrammetric mapping
- Digital aerial imagery – Monochromatic, Colour or Multi-spectral
- Digital Orthophotos – Black and White, Colour or Multi-spectral
- Digital Mosaics
- Raster versions of vector maps
- Scanned index maps or other rasterized vector data
- Digital Satellite Images
- Baseline Thematic Mapping products
- Scanned maps

Access to and distribution of products must be recorded, for both external and internal users. Distribution channels must be controlled, data must be secure and access policies must be enforceable.

The systems must provide independent GDBC facility for the collection, storage, access and search of metadata for all products, both digital and hardcopy. Metadata facility must be capable of supporting highly customized metadata content and of supporting Canadian and international standards for metadata content and for metadata search and delivery. Metadata facility must also be tolerant of situations where metadata is missing or incomplete, such as for old 5" B&W photos.

The systems must provide independent GDBC facility for the collection, storage, access and search of a complete inventory of all products, both digital and hardcopy.

The systems must support independent GDBC on-line and offline ordering and payment of product. Offline ordering is a necessity of incremental change, but will not be supported in later

phases. Targeted customer notification of newly available product is desirable for the e-commerce components.

Customers must be able to track orders.

The systems must support independent GDBC on-line and offline delivery of product. On-line delivery is not required in the near term but over five years of GDBC activity WAN capacity will increase and may be able to support realistic on-line delivery times. For the middle time frames, on-line delivery might be specified to allow for 24-hour delivery time frames and thus must be schedulable.

Metadata must be bundled for delivery with all products. The mandatory content of bundled metadata is subject to further analysis and decision-making.

Customer specific specialized on-line delivery mechanisms (direct transmission from store to end point or client-server data services) may need to be supported. Interfaces such as these will be put in place only if the interfaces for communication, content and format conform to Canadian or international standards and if control and record keeping for access and service is retained by GDBC.

A browser based viewing/delivery facility is required that does not require downloading digital images, except for interactive screen-sized images which are generalized images or raw data at full resolution but limited extent.

3.4.3. Technical Infrastructure Requirements

The primary requirement of the Digital Image Management systems is to enable management activities related to the creation, acquisition, assembly, updating, verification, secure and recoverable storage, manipulation, access control, representation, distribution, and archiving or removal of digital image products.

Vector data must also be managed and useable within DIM systems. This type of data is typically used during product search and indexing. An example is aerial photo indexes.

Customer access to discovery and distribution facility must be on a 7/24 basis, with scheduled outages of any reasonable duration allowed. All components must be capable of handling situations where communication to other components degrades or is abruptly terminated. All components and the system in general must be capable of disaster recovery within the time frame and to the degree stipulated by GDBC and government policy for providing service.

The systems must exist within and use the Metropolitan Area Network (MAN) and the GDBC-NET. This does not preclude stipulating changes or additions to GDBC-NET.

Backup and recovery, and disaster provisions must be integrated with GDBC facility and standards. This does not preclude stipulating changes or additions to these facilities as long as heterogeneous methods are reduced in the process.

The systems must enable single, non-redundant storage of data, including raw data and final forms of derived products, and of intermediate products such as digital mosaics needed to optimize Customer searches for products.

Systems components for storage, discovery and delivery should be "dedicated" and separate from those of acquisition and product creation, so that customer response time is not affected by internal operations.

Systems components must be scalable, and in keeping with standards based communication must be exchangeable by category. Capacity growth will be incremental and will be ongoing.

The systems must provide a single interface for access to data (portal concept).

GDBC customer interfaces must support customer registration, authentication and secure connections as part of e-commerce and as part of on-line delivery.

E-commerce components of the system must support both charge-for-time and charge-for-product accounting. Both Customer billing groups (e.g. MoF, MELP) and named Customers must be supported, with individual pricing structures. The history of ordering needs to be taken into account when calculating charges for re-requesting delivery or service.

The systems must support production workflows, project management and contract management. Internal product production management and workflow control systems must support "workspaces" and long transactions, including multiple user check-in/check-out and QC facility.

It is desirable to be able to offer arbitrary clips and generalizations of base imagery as derived products. These would be generated by the system on customer request without manual intervention. For example, a customer may want to order only a detail clip of a recreational facility and may not want an entire image.

On average, 50,000 photos are taken annually. There are 2.5 million archived hard copy negatives. Currently 4,000 images are scanned annually and the current in-house scanning infrastructure has a capacity of about double that. Currently, 1% of the total potential image data store is already digital in nature (satellite data + existing air photo scans on CD).

A component of the systems are image compression, both lossy and non-lossy, and Stereo Image formats. The standards for compression algorithms and compressed image storage are allowed to be proprietary, and currently include TIFF, GeoTIFF, SIS and others. It is desirable to reduce the number of internally required formats and also to reduce the number of externally supported formats to a feasible minimum.

The system must support attributes of images. Some important attributes are flight index data, footprint, bounding box, contained toponymy and specific metadata such as year, type, film, scale, centre, etc.

The system must support both spatial and attribute image indices, and provide for ubiquitous unique object identifiers for images.

From the analysis of the ordering interface, it is apparent that the system must be able to quickly provide thumbnails of images and to provide image footprints.

Delivery by CD is not an option for customers who purchase in volume, but is an option for occasional customers such as public interest groups. Returnable, portable media is acceptable.

The system must quickly transfer huge volumes of data from image scanning sources (which are few but may not be in-house) to store, and must quickly transfer huge volumes of data from store through both on-line and off-line distribution channels.

3.4.4. Organizational/Process Requirements

The systems and organization must support workflows in four primary business processes (Acquisition, Discovery, Image Data Store Management and Distribution). Several allied workflows must also be supported, including contract management and project planning and management.

Continuous improvement processes must be supported by the systems. One aspect of this is the Ministry standardized (single facility) creation and maintenance of data models for product metadata and for digital products.

GDBC or MELP must formulate Ministry wide formal data management standards and policies for digital imagery management (e.g. data custodian and stewardship, backup and recovery processes, etc.)

The systems must be capable of supporting completely separated development, testing and production environments, and a managed and auditable change control process.

The systems must be supported and managed by the GDBC organization, with respect to application management (planning, standards, security, change control), host management (availability monitoring, performance monitoring, capacity plan, security), network management (availability monitoring, performance monitoring, capacity plan, security), DBMS management (standards, monitor, capacity plan, security) and data management (standards, change control, QC).

4. PROPOSED SOLUTION

4.1. Background

Through consulting assignments with various government and non-government customers, Sierra Systems Spatial IT consultants have observed that the demand for high-resolution imagery is expected to continue to increase for the foreseeable future. This increased demand has been fuelled by improvements in compression technology, higher speed Internet connections and new Web Based GIS vendor products that support the display and manipulation of high-resolution digital imagery. Areas where demand for high-resolution imagery is increasing are:

- Orthophoto base (backdrop) for Geo-Spatial web portals
- Interpretation and/or update of forest inventory
- Mapping terrestrial ecosystems
- Mapping terrain stability
- Mapping large scale municipal areas
- Calculation of volumes for open pit mines and similar activities
- Archiving of historical imagery
- General public usage for tourism, land development, business and personal use
- Environmental change detection

Feedback from the Stakeholders' workshop carried out for the Current Situation Assessment has confirmed this trend and enforces the timeliness of the current GDBC activities.

GDBC is preparing for technological and business change and recognizes the importance and impact the increasing demand for digital imagery will have on both itself and its customers. This change impacts both the technology used by GDBC and the workflow procedures and methods carried out by its employees.

4.2. Need for Incremental Change

4.2.1. Overview

GDBC staff are currently working at maximum capacity and it is determined that the continuation of this work must be continued during any changes that may be imposed. The migration to new hardware, software and procedures must be coordinated to allow for the inclusion of the GDBC staff for subject matter expertise. The impact of the changes and any additional training that may be required must be performed in a non-intrusive manner. These implementation requirements

dictate that the changes and their impact must be controlled. Sierra System's experiences in integrating new systems have proven that planning for incremental releases is highly successful. The benefits of this strategy allows GDBC to realize immediate benefits and return on the investment as well as allowing the daily business to continue while employees are being trained and familiarized with the new applications and procedures.

In summary the strategy of incremental change provides:

- User 'buy in' by observing results and improvements early and often
- Quicker utilization of the modules as users only need training on released components
- Manageable phases with contained objectives and resource requirements
- Ability to adapt the start of each phase to the changes that may have occurred
- Incorporation and utilization of accomplishments from previous phases

4.2.2. Proposed Priorities

The review of the current Business Process Models clearly identified that the current evolution of digital image management within GDBC has resulted in a dispirit set of methods of dealing with digital images. These various methods must be combined or harmonized prior to implementation of a final target DIM system.

The integration of technology in any business mandates that technology cannot change without an impact on the process. A realignment of the activities and knowledge of employees must be managed to optimize any changes in technology.

All Imagery Available Online

Currently a significant amount of work time is consumed in the tracking, storing, retrieving and accessing digital images (see appendix "F" – Digital Imagery Storage Comparison Analysis). Utilizing portable data storage for images, (i.e. compact disks) created within various departments for internal use, customers or other business requirements have resulted in inefficiency when compared to the use of online technology. The inefficiency lies within the staff time to copy data to a CD, ensure the data has been copied completely, record the activities and store the CD. There is also customer and staff inefficiency within the efforts to locate and use data on CDs.

It is recommended that all images are captured, catalogued and made available online for access.

In the immediate time frame, all currently digitized images should be loaded onto a central server within a file and directory structure. This brings the various images being managed by GDBC together in a central, online, file system based Image Data Store.

As time goes on, all hardcopy images should be loaded into a steadily maturing Image Data Store.

Data Custodians and Stewards

A data Custodian is the "owner" of the data and is responsible for setting standards and for ensuring the data meets the business needs of the organization.

A data Steward, on the other hand, is responsible for the maintenance of the Image Data Store and ensuring that all information and images meet the Custodian established specifications and quality control. The data Steward is also responsible for safeguarding the Image Data Store from an access control and disaster recovery requirements. The Steward is responsible for ensuring that images are available to the user community in a timely manner.

It is recommended that a Data Custodian and a Data Steward be established for all digital image products managed by GDBC.

Establishment of a data Custodian and a data Steward for the various digital image products is important. Doing so will meet both the Customer and internal need to address confusion over heterogeneous data, will address the need for timely and continuous planning and improvement within GDBC and will facilitate communication externally and internally.

Technical Centralization and File Management

This activity is a transitory step that needs to be carried out before further technical systems enabling new business processes can be put in place. The result of this activity will be based on the preliminary Image Data Store and will provide metadata and catalogue management systems. This will result in optimized accessibility to images within GDBC.

Existing metadata will be concurrently loaded with images to a DBMS managed metadata store. Also, a simple Central Electronic Catalogue (CEC) will be populated during this phase.

The CEC will be based on the Film Roll and Frame Number for digital photo images, and the 1:20000 Map sheet ID for orthophotos and other map products. Satellite images will be indexed in the catalogue by a convention of 'track-Frame + 'monthyear' + sensor (i.e. 48-22-0920001-L7).

Digital Image Management and Distribution Specification

This activity will detail the business and technical requirements of the final target environment and technology. The specifications will describe each process module at a technical level, and will specify the underlying technology, architecture and infrastructure to support the Digital Image Management modules.

4.3. Proposed Solution - Business Processes

This section of the report describes the proposed or "To Be" business processes for digital image management within GDBC based on the observations and recommendations from the *Current Situation Assessment*. The proposed solution adheres to relevant recommendations contained in the *Data Management Architecture Review and Recommendations* report mentioned previously.

The proposed business solution includes process models but does not include functional models. The functions that the departments of GDBC currently perform were not analysed with the intention of re-engineering, but were modeled to ensure that all relevant processes were identified, so that a technical architecture could be formulated.

This section describes new business processes, and is not intended to provide detailed information on technical implementation of the DIM system. The distribution of this section to stakeholders will facilitate communications and help them understand the changes and possible impact on service.

4.3.1. Approach

To address the objectives of the project in a systematic manner the overall digital imagery management function was divided into four logical components or modules (i.e. Acquisition, Image Data Store Management, Discovery, and Distribution as described previously). Each of these process modules is considered as distinct and discreet in order to enable design activities. However, each module must support and interact with other processes from a customer, management, technical and user perspective. This modular approach allows for the integration of "best of breed" commercial software solutions. The discreet module concept also allows phased implementation.

The proposed business processes described below concentrate on the business processes that require analysis, re-engineering or rationalization.

4.3.2. High Level Business Process Model

The following diagram shows a high level graphical representation of the proposed business process model, grouped by the appropriate functional component or module (i.e. Acquisition, Image Data Store Management, Discovery and Distribution).

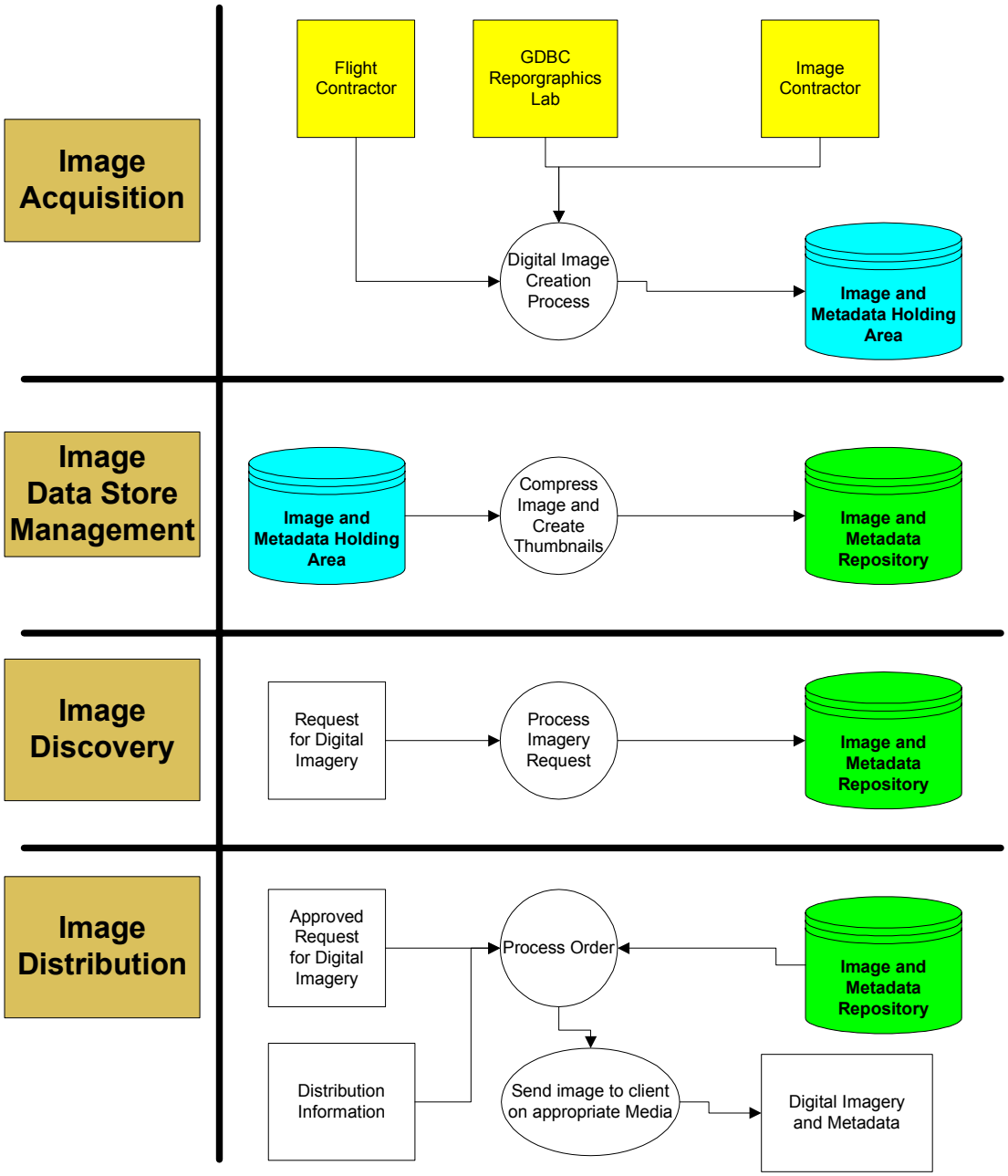
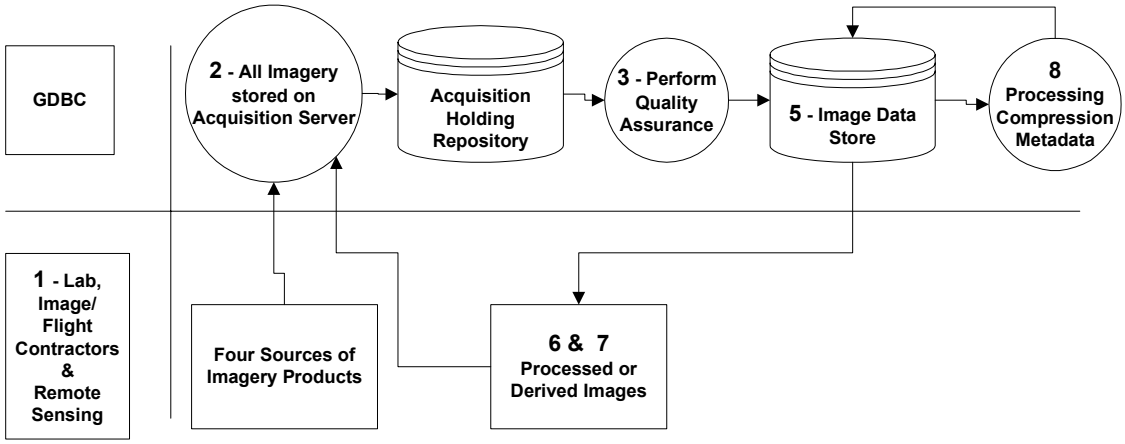


Image Acquisition
Flight Contractor
GDBC Reporgraphics Lab
Image Contractor
Digital Image Creation Process
Image and Metadata Holding Area
Image Data Store Management
Image and Metadata Holding Area
Compress Image and Create Thumbnails
Image and Metadata Repository
Image Discovery
Request for Digital Imagery
Process Imagery Request
Image and Metadata Repository
Image Distribution
Approved Request for Digital Imagery
Process Order
Image and Metadata Repository
Distribution Information
Send image to client on appropriate Media
Digital Imagery and Metadata

4.3.3. Proposed Business Process – Image Acquisition

The following high-level diagram is intended to provide an overview of the overall process. The diagram is followed by a point form description of the general steps in the process. A detailed model diagram is provided in Appendix "B".



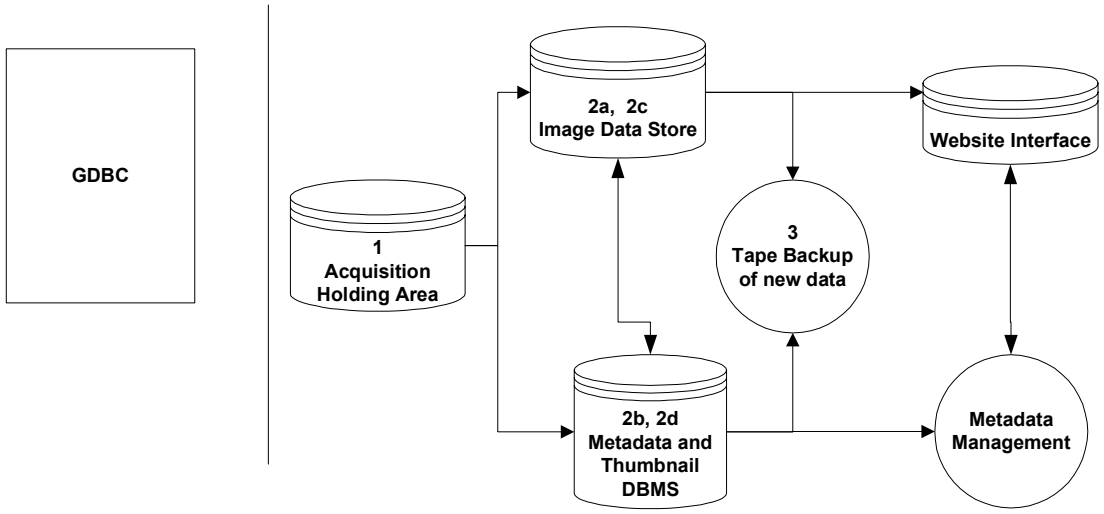
The high-level process steps are:

1. Five potential sources of imagery were identified– 1) Image contractor, 2) Flight contractor, 3) GDBC Reprographics Lab, 4) Canada Centre for Remote Sensing and 5) GDBC Image Data Store Manager. The fifth source consists of images already available from the Image Data Store; that are used to create derived products.
2. All digital imagery arriving from these first four sources (be it from CD/DVD/tape, portable hard drive or on-line transfer) is stored on the Acquisition Server.
3. Digital imagery on the Acquisition Server is reviewed and inspected, which results in two possible outcomes: 1) image is rejected or 2) image is accepted.
4. Images that fail the quality assurance stage result in the contributor being contacted with the rejected data information.
5. Imagery passing quality assurance is immediately copied to the Image Data Store (including digital image, metadata and supporting files)
6. Imagery requiring further processing will be distributed to third parties for processing. The subsequent product would be submitted back to and treated as new imagery (i.e. start at step 1 above).
7. Imagery requiring in-house processing is conducted and quality assurance testing. Upon satisfactory completion the images are copied to the Image Data Store and linked to the appropriate metadata.

8. Compression processes to produce the thumbnail and compressed images are performed. Quality assurance tests are conducted and upon satisfactory completion the images are copied to the Image Data Store and linked to the appropriate metadata.

4.3.4. Proposed Business Process – Image Data Store Management

The following high-level diagram is intended to provide an overview of the overall process. The diagram is followed by a point form description of the general steps in the process. A detailed model diagram is provided in Appendix "B".

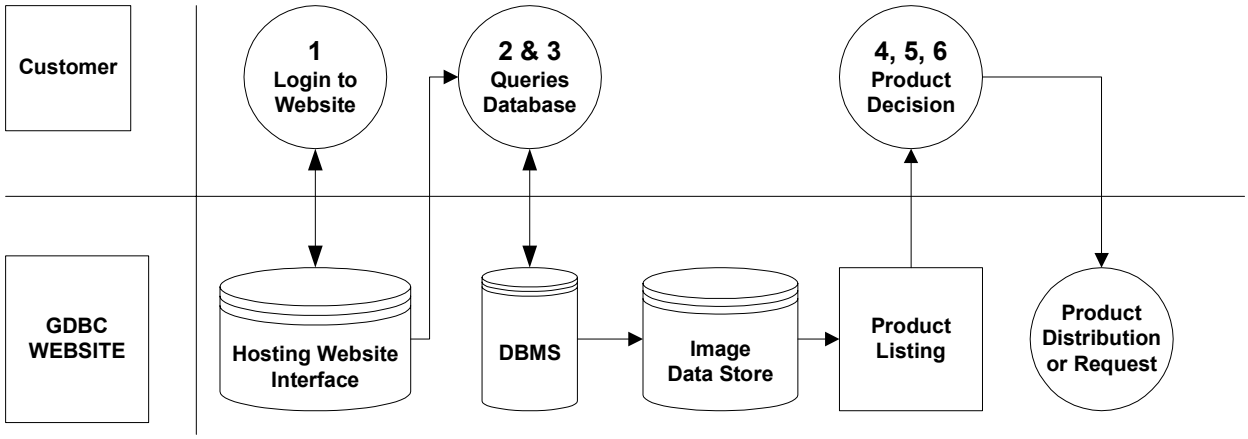


The Image Data Store Management process is not a linear activity, but a list of activities that will ensure that the imagery and data is available, efficiently accessible and secure. The high-level process steps include:

1. All imagery is copied or populated into the Image Data Store via the Acquisition process (Image Acquisition process).
2. Digital images in staging area are copied to Image Data Store (Image Acquisition process)
   a Raw images and compressed images are copied to Disk array
   b Metadata is loaded into the DMS via the Metadata Manager application
   c Flight, Vector and Raster indices loaded to DBMS and Internet Map Server application
   d Thumbnail images loaded into DBMS for subsequent access via the Internet Image Server
3. Backup to tape for offsite storage of new data
4. Ongoing Data Base Administrator activities for security, efficiency and optimization

4.3.5. Proposed Business Process – Image Discovery

The following high-level diagram is intended to provide an overview of the overall process. The diagram is followed by a point form description of the general steps in the process. A detailed model diagram is provided in Appendix "B".

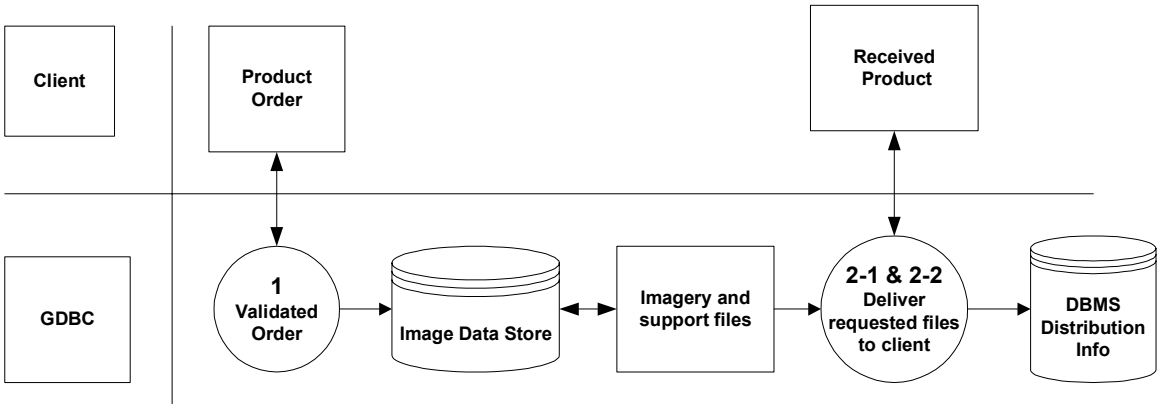


The high-level process steps are:

1) Customer accesses Website
2) Two discovery streams are presented: i) using map viewer (map navigation) or ii) direct text based queries of metadata fields.
    a) Map Viewer method will allow for a 'zoom-in to area of interest' discovery process, where footprints of photos/orthophotos/etc will be presented depended on display scale. Flight indexes (flight lines/roll numbers/photo numbers) will also be visible at appropriate scales.
    b) The textual method will allow users to specify their search criteria by inputting data into appropriate search fields, then querying the database directly for matches.
3) Either method's result will be a listing of available products (including analogue images)
4) The user would then order the selected product(s) found to exist in the Image Data Store, and the Image Distribution process would continue processing the transaction.
5) If the specified product was not found in the Image Data Store, but data existed that would allow creation, then a request to have it acquired could be submitted.
6) If the specified product does not exist and no data was found that could create it, then the user would contact Customer Support for further actions.

4.3.6. Proposed Business Process – Image Distribution

The following high-level diagram is intended to provide an overview of the overall process. The diagram is followed by a point form description of the general steps in the process. A detailed model diagram is provided in Appendix "B".



The high-level process steps are:

1) Verify imagery order for appropriate account or credit approval and imagery availability
2) Determine distribution method (Option 1 or Option 2)

Option 1 – CD, DVD or portable hard drive

a) Create copy to media instructions
b) Update order status information
c) Receive 'copy to media' instructions
d) Copy ordered imagery and data to media
e) Quality assure imagery
f) Send media to packaging and shipping department
g) Update order status information with shipping information
h) Update image ordered history

Option 2 – Online delivery

a) Download image via the internet
b) Update order status information
c) Update image ordered history

4.4. Proposed Solution – Technical Architecture

4.4.1. Overview

This section of the report describes the proposed technical architecture for the Digital Image Management system. It describes known constraints, recommended solutions and areas where further action and analysis are needed to complete a detailed architecture plan.

The technical architecture proposed has been presented at a high level. A more detailed technical architecture plan with decisions related to the system architecture can only be developed after the selection of software packages and operating systems.

Sierra Systems' technical architecture recommendations are focused on the immediate infrastructure requirements of the Digital Image Management system. They are not intended as a general infrastructure recommendation for all of GDBC's operations and do not, for example, directly address MASCOT's or other program needs.

4.4.2. GDBC Framework Architecture

Figure 1 represents the high level "to be" architecture diagram as presented in the *Data Management Architecture Review and Recommendations* report.

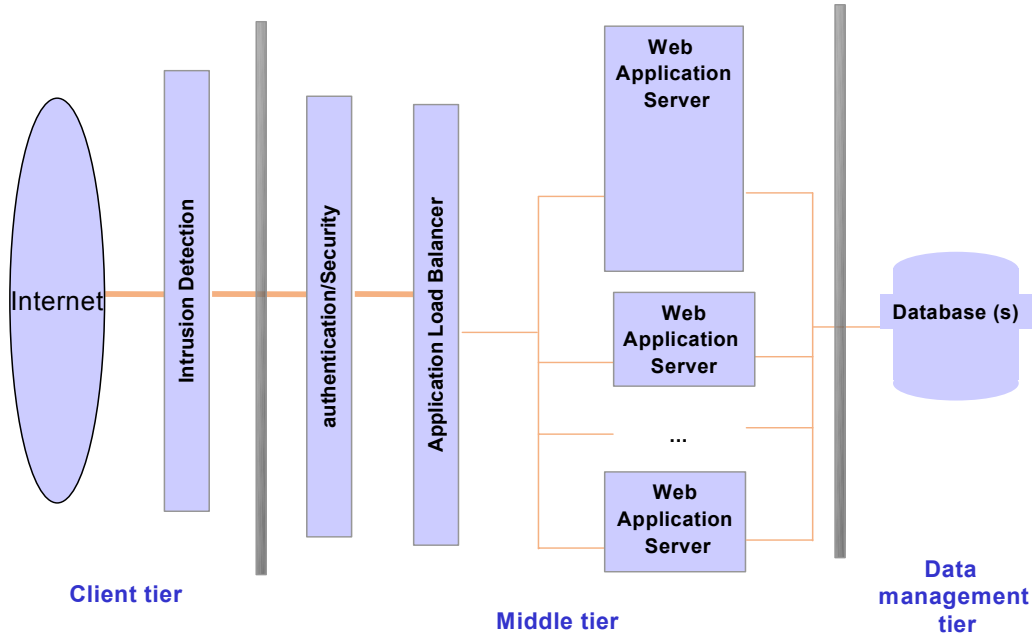


Figure 1 – High level "to be" n-tier architecture

Recommendations of the *Data Management Architecture Review and Recommendations* report relevant to the recommendations made in this report are quoted or listed below:

SCOTS solutions: GDBC is in the information management business, not the technology development business. Commercial-Off-the-Shelf products can meet more than 85% of the data management and data delivery requirements of GDBC today. Implementing solutions based on commercial products with a positive track record should minimize risk and investment in technology.

ISO and OGC standards*:* ISO TC211 and the OGC have signed a 'Class A' liaison, and will be working together to fast-track several standards for commercial use. This means that each of these standards will be endorsed in over 40 countries and by 200 commercial vendors. This minimizes risk for implementation and ensures that any system built by GDBC will be compatible with commercial and public sector stakeholders in the long term. [1]

Ministry standards for network configuration*,* e.g., firewalls, routers, authentication servers, etc., are critical to ensure that GDBC can share benefits of infrastructure supplied by the Ministry.

Oracle as the database for core assets*:* The Oracle data management environment is axiomatic for this project.

4.4.3. Overview of Proposed Technical Architecture

The proposed Technology Architecture has been designed to meet the unique needs of a Digital Image Management system for the Acquisition, Image Data Store Management, Discovery and Distribution of resource based digital imagery. The major consideration for this design is the high I/O throughput and high storage capacity that is required by digital imagery. Other contributing factors include data acquisition, data distribution, metadata management and system security.

[1] The three major spatial data management agencies within the US all have stated that they will not procure software that does not meet these spatial standards (U. S Geological Survey, National Image and Mapping Agency, and Army).

Logical Architecture

The Logical Architecture diagram (figure 2) describes the *n*-tier architecture that is the basis for the recommended solution.

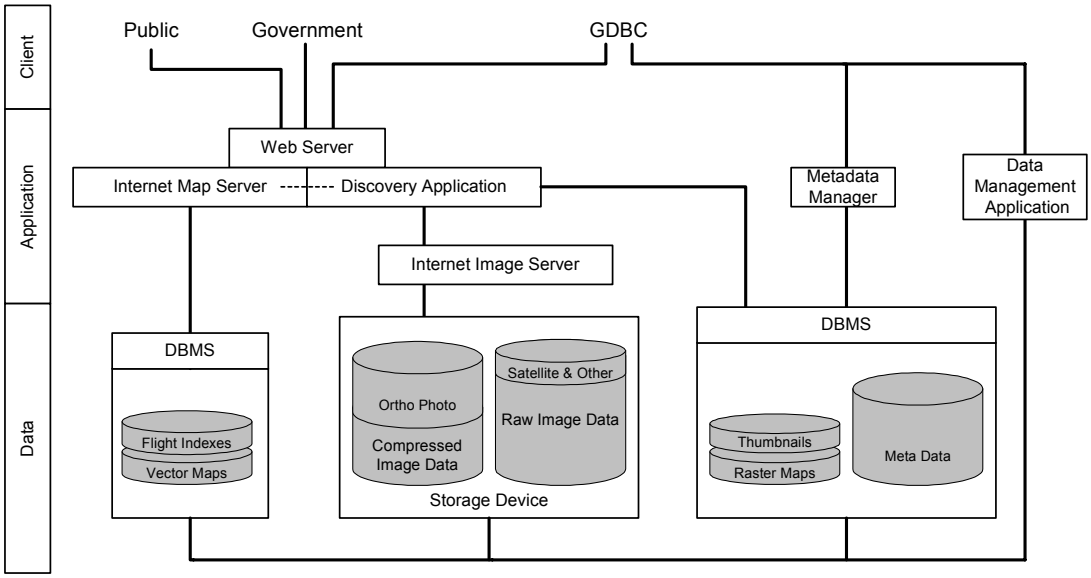


Figure 2 – Logical Architecture

Technology Architecture

The recommended Technology Architecture (figure 3) consists of an n-tier systems design encompassing all enterprise aspects of the Digital Image Management System.

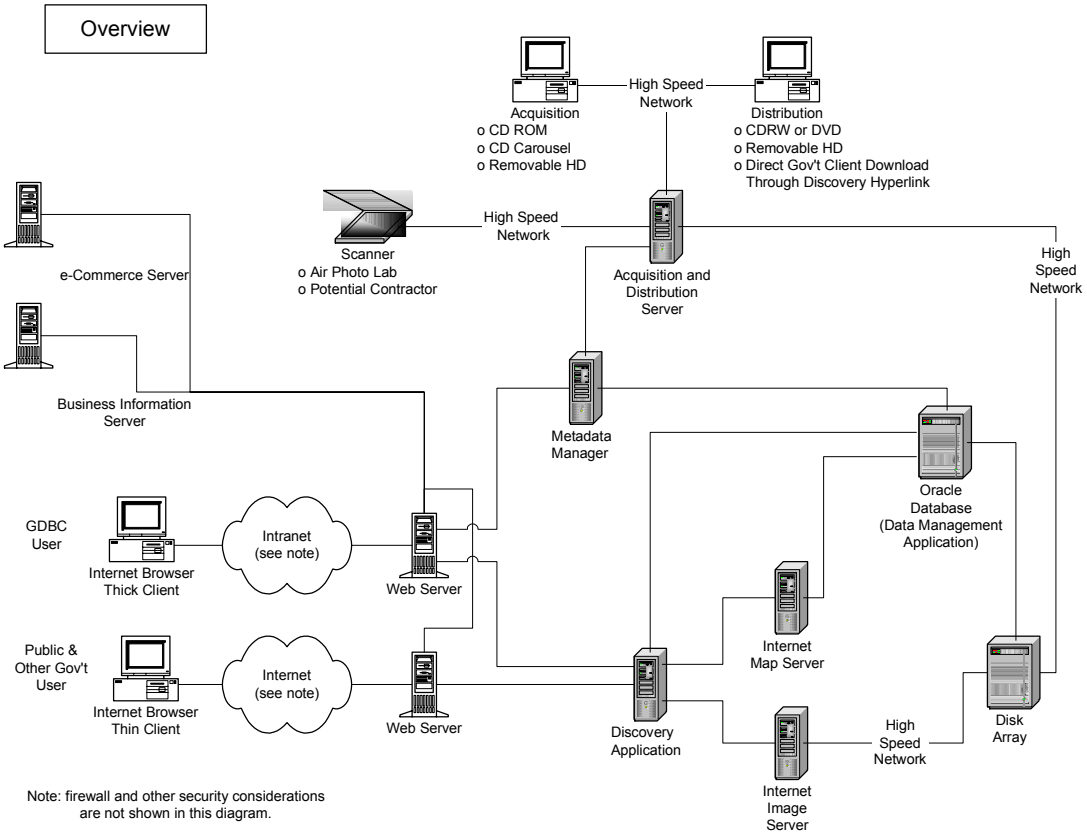


Figure 3 – Recommended Technology Architecture

4.4.4. Proposed Architecture

The overall Digital Image Management system has been separated into four functional components or modules: Acquisition, Image Data Store Management, Discovery and Distribution. Each of the module areas will require distinct technology solutions while sharing in the common infrastructure. The following sections provide a high level view of the recommended Technology Architecture as it applies to each of these modules.

When deployed onto servers, it is proposed that both the Acquisition and the Distribution application module sets will be combined onto a single server. Acquisition estimates must be

combined with similar statistics for Distribution to size the combined Acquisition and Distribution server.

4.4.4.1. Proposed Architecture – Acquisition

Acquisition modules provide facilities to populate the digital image catalogue with metadata and image products.

The Technical Architecture involved with Acquisition is shown within the greyed area in figure 4.

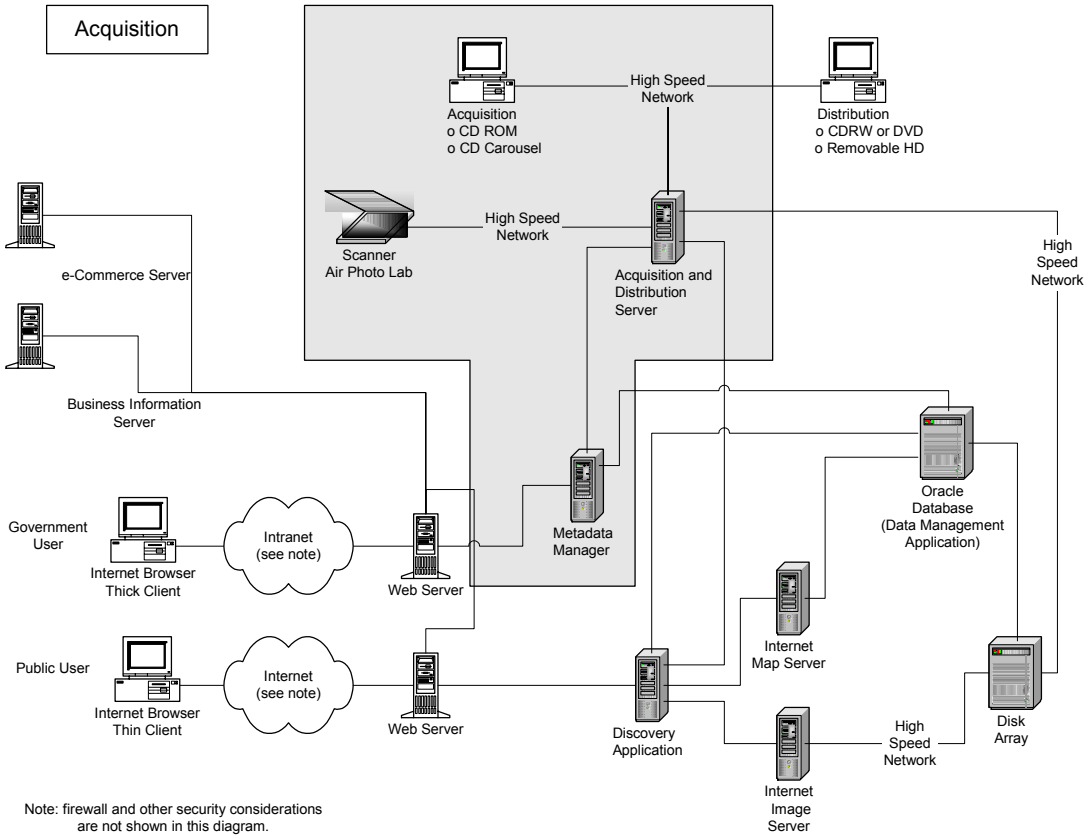


Figure 4 – Technology Architecture – Acquisition

Acquisition Services

Acquisition modules provide the services to receive and store imagery temporarily from various acquisition sources including:

- Direct upload from Air Photo Lab.
- External contractor delivery to GDBC on CD, DVD or removable hard drive.

In addition to data storage, the services will host image compression software to create compressed versions.

The acquisition of digital images into the GDBC Image Data Store will primarily be from archival and scan on demand activities at the Air Photo Lab, and value added products creation by the private sector under contract to GDBC.

Metadata Manager

Data loading will require access to Metadata Management software for entering metadata details of each image uploaded to the main Image Data Store.

Acquisition Data Management Module

There is a requirement for a data management module to manage the upload synchronization of digital images and image metadata to Acquisition storage. If this requirement cannot be met within packaged solutions, the Data Management Module will require custom application development.

Data Compression Application

As part of the data acquisition process, raw image data will be compressed at pre-defined ratios and loaded to Acquisition storage along with the raw image data.

Functional requirements for compression software are detailed in the TRIM II Data Compression Software Evaluation report prepared for GDBC by Sierra Systems.

Thumbnail Creation Application

A function or application will be required to generate thumbnail images for quick reference in the Discovery stage.

Oracle Application Database

Oracle DBMS services will be used for all Acquisition persistent storage.

Storage Capacity Planning

Storage capacity must be dedicated to Acquisition services sufficient to manage the volume of data that will be uploaded to the main server on a daily basis, and to provide a buffer for upload storage of 5 working days. This buffer will be used when large shipments are made from contractors, when QA services are not available for an extended period or when the Image Data Store is not available.

It is estimated that the current capacity and timetable for archival scanning activities at the Air Photo Lab could produce two 500 Mb images every twenty minutes. The proposed technical architecture must be able to initially accommodate a minimum of two concurrent uploads to the main disk array every ½ hour during regular business hours (9:00 – 4:30). This capacity requirement must be reviewed whenever changes to GDBC scanning facility are contemplated.

Data volumes from Acquisition activities for other products such as orthophoto and satellite imagery will be required to determine specific architecture requirements. This will be obtained through further analysis during the initial phases.

The volume represented by this data will depend on a policy decision by GDBC as to what additional services and products it will provide beyond the archival of 5" photography. Also, basic services and configuration for a server hosting Acquisition modules will be dependent on the operating system selected.

Based on the information above, sizing can be initially characterized as follows:

- High Volume
- Large transaction size
- Moderate digital image storage requirements

The following table illustrates the initial estimates of actual production scanning and potential volatile disk space required on the Acquisition server.

Image Type	Individual File Size	Total per day (Qty of Scans)	Daily load to Acquisition Server
5 inch (archival)	80 - 100 Mb	70 - 80	8 Gb
5 inch Compressed 10:1	10 Mb	70 -80	80 Mb
10 inch (scan on demand)	300 Mb	2	600 Mb
10 inch Compressed 10:1	30 Mb	2	60 Mb
Ortho Photo	450 Mb	1	450 Mb
Ortho Photo Compressed 10:1	45 Mb	1	45 Mb
10 inch per ortho (from ortho production)	500 Mb	4 - 6	3 Gb
10 inch Compressed 10:1	50 Mb	4 – 6	300 Mb
Satellite Imagery (sub sample)	80 Mb	1	80 Mb
Satellite Imagery (raw)	300 Mb	3	900 Mb
Controlled 10 inch	400 Mb	6	2.4 Gb
		Potential Volatile Disk Requirement	**31.8 Gb per day**

Notes:

1. One scan on demand operation would replace two 5" scan operations.
2. Estimates for Ortho photo delivery represent potential daily requirements for a single 1:20,000 Orthophoto.
3. Approximately four to six raw 10" images are delivered with each orthophoto.
4. Estimates for Satellite imagery represent potential daily requirements for a single 1:250,000 satellite image.
5. The values presented in the above table would represent a “worst-case” daily data volume, since the production and loading of 10" photography, orthophotos and satellite imagery happens on an ad-hoc basis, and this type of operation offsets the production of 5" imagery.

The conclusion is that the Acquisition services will require 32GB * 5 working days = **160 GB** of disk storage.

Network Capacity Planning

It is estimated that the capacity for archival scanning activities at the Air Photo Lab could be two 500 Mb images every twenty minutes. It is mandatory that there be no bottleneck in the hand-off of this data to Acquisition services, and the nature of the transmission is such that this data is sent in bursts of 500MB every 10 minutes, which are expected to complete within 60 seconds.

Therefore, a standard transfer rate of 10 MBPS would be sufficient to provide timely uploads of image data to a combined Acquisition & Distribution server.

However, given the low cost difference of a standard 100 MBPS connection from the Air Photo Lab to Acquisition services, and the fact that this connection could also be used for other data transmission by GDBC, a dedicated 100MBPS connection should be considered as the minimum acceptable.

4.4.4.2. Proposed Architecture – Image Data Store Management

Image Data Store management monetarily represents the most significant component of the required technology architecture for the Digital Image Management system due to the high storage capacity required. This area will require a data base server to house an Oracle data management system and a disk array for the archival storage of digital image data.

The Technical Architecture involved with Image Data Store Management is shown within the greyed area in figure 5.

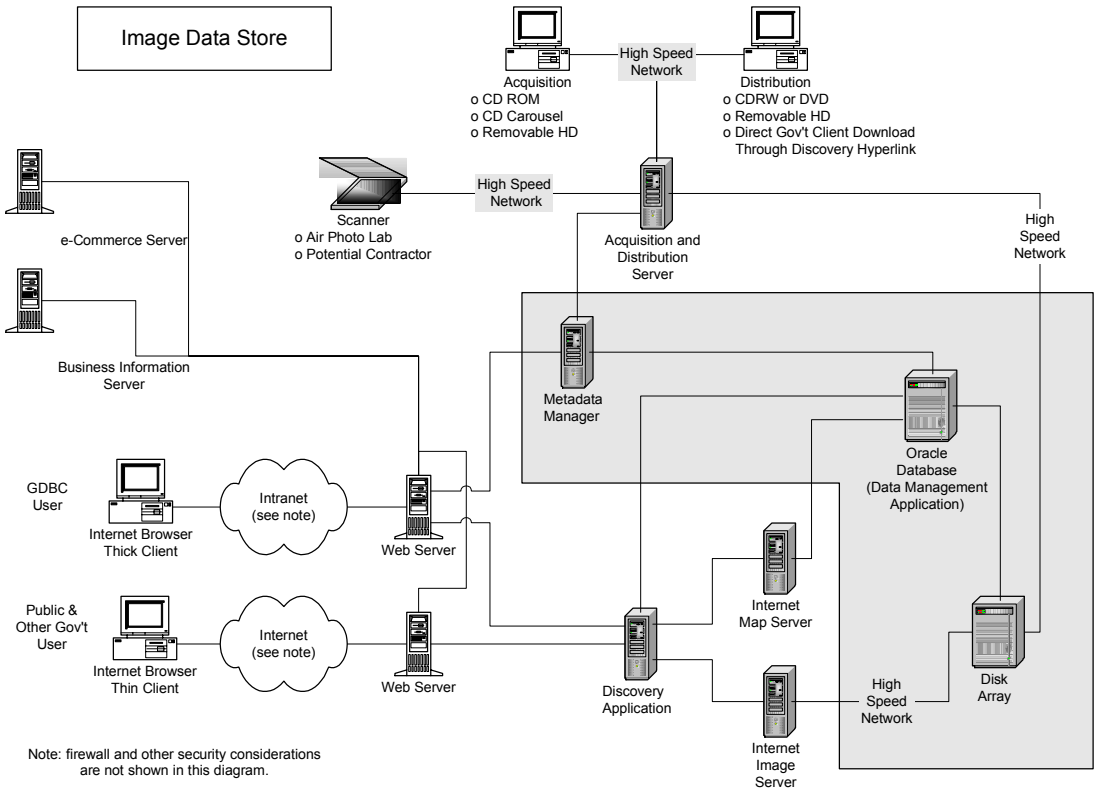


Figure 5 – Technology Architecture – Image Data Store Management

Image Data Store Management Services

Image Data Store Management services provide the underlying facility of a DBMS (transactions, security, indexing, integrity, storage, client-server communication) plus basic data storage, backup/recovery and availability facility. To this are added GDBC specific data management services supporting data I/O workflows, some QA facility and data/metadata management integration.

Metadata Management Software

Although used by Acquisition, Discovery and Distribution, the metadata management software is best described under Data Management. Metadata management will provide the facilities to create and update digital image metadata for all archived digital images and products. The selected software must be FGDC compliant and support the following database and operating system platforms:

- Oracle DBMS
- Unix and Windows NT/2000

Additionally, the metadata management software must fully implement user extensions, for data entry, data storage, and for search and retrieval.

Oracle Application Database

As identified previously Oracle is axiomatic as the DBMS for GDBC's core assets.

Sierra Systems furthers this recommendation by specifically recommending Oracle Spatial to support the storage of vector base map and flight indexes within the Oracle environment for the Digital Image Management system. This recommendation is made for the following reasons:

1) Oracle provides open, standards-based spatial data and location services.
2) Oracle provides management of spatial data on all hardware platforms supported by the Oracle Database Enterprise Edition.
3) Oracle states that it has passed the Open GIS Consortium conformance test for SQL access to geographic information stored in a database. Regardless of the veracity of this statement, all of the leading GIS, mapping and location services tools can now directly access Oracle Spatial, and amongst these are many that explicitly conform to OGC standards.
4) Oracle is an integral member of the important national and international spatial standards organizations, including OGC. This puts Oracle at the forefront of standards based product development. Oracle has had membership in the Open GIS Consortium for the past two years and sits on the Board of Directors.
5) Oracle is also an active member of the International Organization for Standardization Technical Committee, Geographic Information and Geomatics (ISO/TC 211), and the American National Standards Institute (ANSI) Technical Committee X3L1, Geographic Information Systems (GIS) standards groups. Although the ISO/ANSI SQL3/MultiMedia (MM) - Part 3, Spatial Extensions to SQL data standard and the OGIS are not yet complete and are controversial, Oracle's implementation of technology for spatial information management is in line with the current direction of these standards.
6) Oracle is also actively involved in the emerging Location Based Services standards arena, actively participating in the Location Interoperability Forum (LIF) and monitoring related initiatives in the telematics arena (e.g. MAGIC, AMIC and others). Oracle's broad commitment and participation in these initiatives ensures consistency among the various standards groups and ensures Oracle's product lines adhere to the standards and their direction.

The Oracle DBMS will be used as a basic infrastructure to manage all persistent storage aspects of the Digital Image Management system including; metadata, map based digital image indexes and image thumbnails.

Availability and Reliability

The Image Data Store Management area underlies all the other modules, some of which are exposed to external users through the Internet, and therefore 7x24 operation is desirable. Availability has been deemed critical only in environmental or disaster situations. Form this, it appears that the systems must be fault-tolerant but need not be highly available.

Guaranteed non loss of data is also a requirement of Image Data Store Management, and so Image Data Store Data Management will include back up of all managed data and will include second site or off-site storage or replication.

Oracle DBMS Services

Additional analysis will be required to determine actual metadata and other functional area storage requirements, but Oracle DBMS directly controlled resources (as opposed to file system storage) can be initially defined by the following anticipated characteristics:

Image Attribute Data Management functions:

- Low to medium transaction volume, but bursty (short, intermittent periods of high resource use) while searching
- Small transaction size
- Medium storage requirements

Metadata Data Management functions:

- Low transaction volume
- Small transaction size
- Moderate storage requirements

Data Change Management functions:

- Low transaction volume
- Small transaction size (assuming file based image storage)
- Low storage requirements

E-Commerce Data Management functions:

- Low transaction volume
- Small transaction size
- Low storage requirements

Disk Array Services

The storage of raw and compressed image data, either managed directly by Oracle DBMS or managed by Oracle through the file system will require dedicated disk arrays of sufficient initial capacity and scalability to satisfy expected storage requirements. Over the five-year duration of GDBC activity the disk arrays must be scaled up to meet demand on a schedule determined by available investment and by current and anticipated costs of disk.

Final storage size can be initially defined by the following anticipated characteristics:

- Low transaction volume
- Very large transaction size
- Very large storage requirements

Estimated I/O Capacity Requirements

Initial I/O capacity plans are based on the assumption that nominal maximum I/O needs will occur in the circumstance that concurrently:

- A single image is being delivered
- A single image is being acquired
- Metadata is being input and searched

Since these flows are limited by individual 100MBPS network links, the indication is that 300MBPS I/O would be needed. This nominal maximum could easily be accommodated by a single Fibre Channel link (1 GBPS), which could also be expected to accommodate most instantaneous maximums (300 * 3 = 900MBPS) as well.

High-speed connections will be required between the disk array and all components it serves.

Furthermore, high-speed network connections would be cost-effective and justified throughout the internal GDBC DIM server systems. It is recommended that a minimum transfer rate of 200 Mb/s is available to provide timely loads of image data through a combined Acquisition & Distribution server for distribution.

As noted, the volume of I/O is a major consideration to the performance of the Digital Image Management system. A typical compressed Black & white image is approximately 30 megabytes (at 14 micron scan resolution), a compressed colour image is approximately 100 megabytes (at 21 micron scan resolution). The digital image management system will be required to serve the thumbnail and possibly the compressed imagery through the Internet Image Server and the Data Delivery module to the web server. This volume of data must be served efficiently within the system architecture.

It is critical that the network transmittal of image data to the Internet Image Server be as efficient as possible. It is recommended that a minimum transfer rate of 200 Mb/s be available for this connection. Technical options to meet this requirement will require further analysis.

Estimated Storage Capacity Requirements

Initial estimates of file size (transaction size) range from 100 Mb to 2Gb. The selection of an operating system or management system for the disk array will be impacted by the large image file size. Additional analysis will be required to determine final digital image file storage requirements.

The following table illustrates the initial estimates of actual and potential disk space required, staged by year. Projections to year 5 have not been estimated due to uncertainty of production volume.

Image Type	Individual Scan Size	Estimated Initial Load	Yearly image production	Estimated Year 1	Estimated Year 2
5 inch (archival)	80 - 100 Mb	567 Gb	16000	1.6 Tb	1.6 Tb
5 inch Compressed 10:1	8 – 10 Mb	N/A	16000	160 Gb	160 Gb
10 inch (scan on demand)	300 Mb	300 Gb	1000	300 Gb	300 Gb
10 inch Compressed 10:1	30 Mb	N/A	1000	30 Gb	30 Gb
Ortho Photo	435 Mb	1.3 Tb	400	174 Gb	174 Gb
Ortho Photo Compressed 10:1	43.5 Mb	N/A	400	17.4 Gb	17.4 Gb
10 inch (from ortho production)	500 Mb	8.5 Tb	2400	1.2 Tb	1.2 Tb
10 inch Compressed 10:1	50 Mb	N/A	2400	120 Gb	120 Gb
Satellite Imagery (sub sample)	80 Mb	25 Gb	12	960 Mb	960 Mb
Satellite Imagery (raw)	300 Mb	75 Gb	36	10.8 Gb	10.8 Gb
Controlled 10 inch	400 Mb	1.6 Tb	1600	640 Gb	640 Gb
		11.5 Tb		**4.3 Tb**	**4.3 Tb**
Year one initial load		**11.5 Tb**			
Potential Year 1 Disk Requirement				**15.8 Tb**	
Potential Year 2 Disk Requirement					**20.1 Tb**

Notes:

1. All sized are rounded up to the nearest unit.
2. 1 and 2 year projections based on a 200 work days per year (single shift).

3. Estimates for Ortho photo delivery represent potential yearly requirements based on current estimates that show production at approximately 400 images per year.
4. Estimates for Satellite imagery represents potential yearly requirement are based on current estimates that show production at approximately 12 images per year.
5. The values presented in the above table represent "maximum-case" annual data volumes based on the daily data volumes from the previous table. These actual volumes could be greater or lesser depending on a policy decision by GDBGC as to what additional services it will provide.

The Image Data Store management module must be scalable to meet the estimated initial, year1, year 2 and ongoing storage requirements. In the longer term the system must ultimately be scalable to meet by year five the total estimates for pre 1961 archival scanning projections of **673 TB** and also the on going yearly production. The ongoing yearly production is expected to increase by year five.

4.4.4.3. Proposed Architecture – Discovery

Discovery will provide the user an efficient method to determine whether an image product exists within a desired area of interest. The hosting system software must support both a map-based visual search and a direct text based field query of the metadata attributes.

The Technical Architecture involved with Discovery is shown within the greyed area in figure 6.

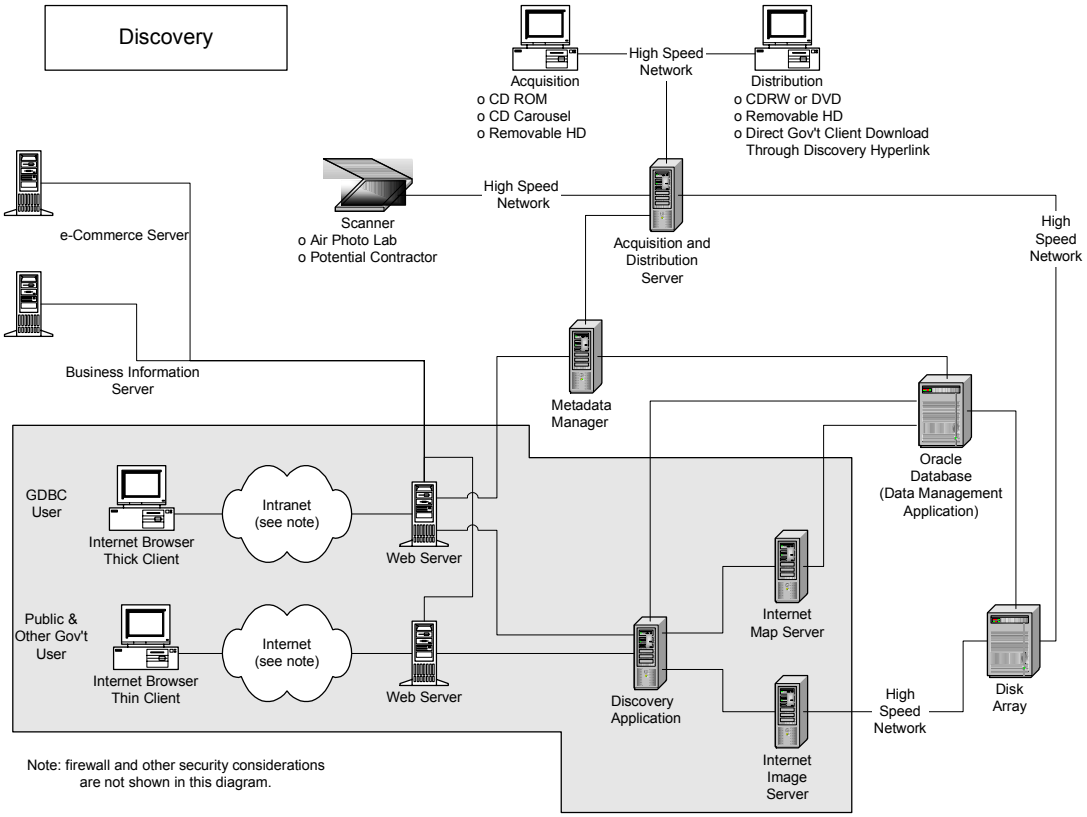


Figure 6 – Technology Architecture – Discovery

Discovery Services

A web enabled discovery application will be required to provide textual and graphical querying of the metadata and various visual indexes to access to the digital image archive.

The logical architecture diagram identifies separate application layers for a Discovery Application, an Internet Map Server and an Internet Image server. Evaluation, selection and final configuration of these components may result in the use of one main application enhanced by a compatible add on or plug-in from a variety of vendors. This section details the mandatory requirements for each identified component regardless of the final configuration.

In the event that the final selection does encompass a variety of software modules from different vendors, custom integration of the map viewer, the image viewer and the discovery software may be required to provide a seamless toolset for the discovery engine.

Discovery Application

A web enabled discovery application will be required to provide textual and graphical search functionality to the metadata database for access to the digital image archive.

The functional requirements of a Discovery Application are as follows:

- Configurable textual search capability
- Configurable geographic search capability
- Thumbnail image display
- Image ordering (single or multiple)
- Account or payment processing (e-commerce)
- Compressed image download
- Raw image download

The Discovery Application must be compliant with ANSI/NISO Z39.50 protocol and OGC standards for interoperability where required.

Appendix "I" contains overview details of the Z39.50 protocol.

Appendix "J" contains overview details of OGC standards for interoperability.

The OGIS Introduction to Interoperable geoprocessing can be viewed at
http://www.opengis.org/techno/guide.htm

Internet Map Server

An Internet Map Server will be required to display vector and raster data to support graphical searching through various visual indexes to access the digital image archive.

The functional requirements of an Internet Map Server component are as follows:

- Scale dependant vector display
- Raster display (with translucent shading)
- Attribute link functionality including: Fly over flags to display attributes or metadata.
- Pan, zoom and window area.
- Scale sensitive zoom (drill down).

The Internet Map Server must be compliant with OGC standards for interoperability. Appendix "J" contains overview details of OGC standards for interoperability. The OGIS Introduction to Interoperable geo-processing can be viewed at
http://www.opengis.org/techno/guide

Internet Image Server

An Internet Image Server will be required to provide for the display of raw and compressed image data served from the disk array.

The functional requirements of an Internet Image server component are as follows:

- Raw Image Display
- Compressed Image Display
- Pan, re-centre, zoom and restore
- Functions related to the "seamless" movement from one orthoimage to an adjoining image
- Function related to the generalization of images
- Functions related to the creation and display of mosaics

Oracle Application Database

The Oracle DBMS will be used to manage all aspects of the Discovery persistent storage.

Human Interface Types

There are two human interface types defined for this recommendation: 'Thin client' and 'Thick client'.

Thin Client

The thin client will host the public (external) user to provide discovery, selection and retrieval of thumbnail image (and possibly the compressed image depending on the size and connection band width) and for raw image product ordering.

In it's basic form, the thin client will support browser-based access driven primarily by HTML content.

The final selected Discovery Application and Internet Map Server software may add the required loading of client side plug-ins to support map data and image display, in which case this is no longer strictly a thin client.

The client will allow support for browser-based access driven by the following:

- HTML content
- Discovery Application and map server software plug-ins

Thick Client

The thick client will host the GDBC (internal) user to provide discovery, selection and retrieval of thumbnail and compressed images (assuming a high-speed connection is in place) and access to DIM management services including the metadata loading application.

The thick client will allow support for browser-based access driven by the following:

- HTML content
- Discovery Application and map server software plug-ins
- Application tools for DIM management services
- Possibly as yet undefined image applications

Web Servers

Web server specifications are out of scope for this report as they are assumed to be part of the general Infrastructure Standards for GDBC. Please refer to the *Data Management Architecture Review and Recommendations* report for general high-level Technical Architecture standards for middle tier server requirements.

Application Servers

The Recommended Technology Architecture diagram depicts the Internet Map Server, Internet Image Server and Discovery Application residing on their own respective physical servers. However, the proposed architecture allows for flexibility in deployment onto servers. Applications could be hosted on a single server, on separate servers or an array of multiple servers with multiple instances of each application.

Additional analysis for performance and load balancing will be required to determine the optimum number of servers and configuration to handle the system traffic, once the various applications have been selected by GDBC.

4.4.4.4. Proposed Architecture — Distribution

Distribution will provide the facilities to disseminate requested digital imagery on line or by transportable media. It is recommended that a single shared server be used to deploy both Acquisition and Distribution services.

The Technical Architecture involved with Distribution is shown within the greyed area in figure 7.

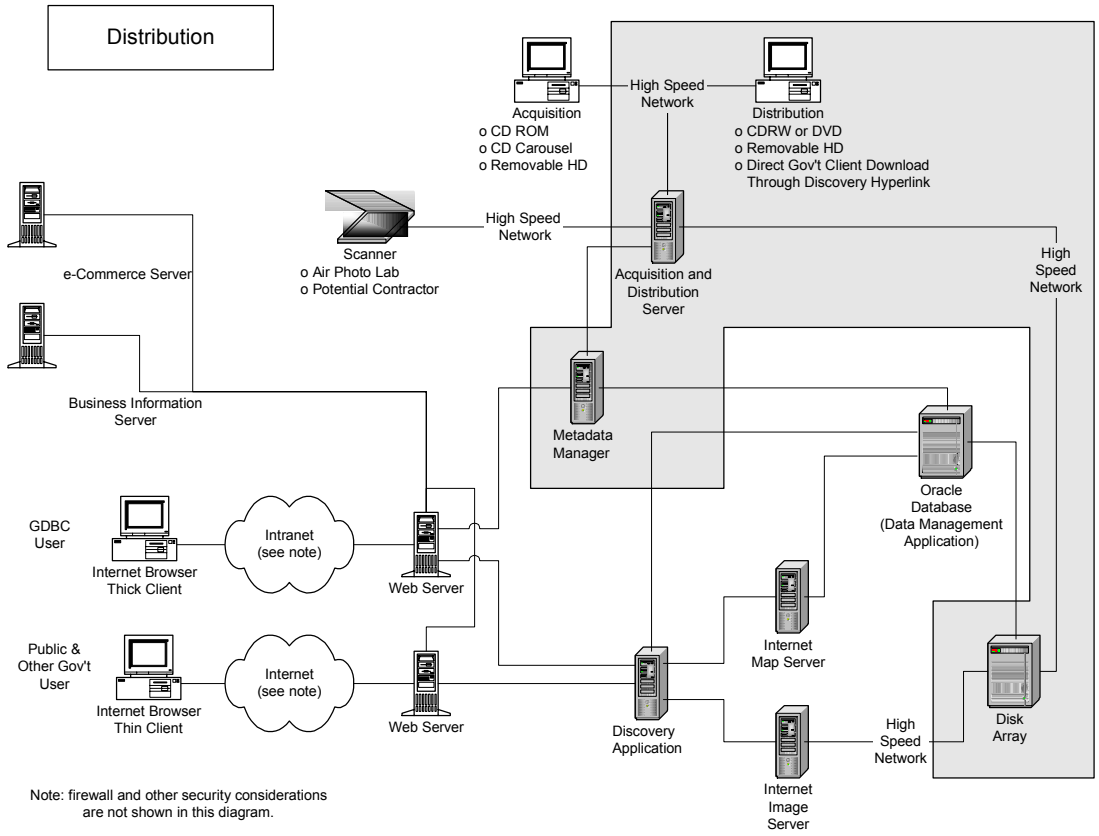


Figure 7 – Technology Architecture – Distribution

Distribution Services

Distribution services provide facility to:

- Receive and store temporarily imagery from the disk array for distribution on portable media or network.
- State based delivery software that initiates, monitors and completes network based large object transfers.
- Interface with e-commerce and control modules.

Distribution requires the use of peripheral devices to facilitate the transfer of image data to portable media. The following devices have been identified as viable for the distribution of image data:

- Commercial CD R/W device
- DVD RW device
- Removable hard drive.

The Distribution Technology Architecture diagram depicts these devices in relation to a Distribution workstation. Further analysis is required to determine the ideal host for these devices.

Basic services and configuration for this server will be determined by detailed I/O analysis.

Distribution Data Management Module

There will be a need for a data management module to manage the distribution of digital images and image metadata to and from Distribution storage.

If this requirement cannot be met through packaged software, the Data Management Module will require custom application development.

Oracle Application Database

The Oracle DBMS will be used to manage all aspects of the image Distribution persistent storage.

Capacity Planning

No metrics have been established to date for estimation of potential distribution throughput however; server size can be initially defined by the following criteria:

- Periodic High I/O Volumes
- Large transaction size
- Moderate digital image storage requirements

Network Considerations

The transfer of scanned image data from the disk array to the acquisition & distribution server will require a dedicated, high-speed network connection. A recommended minimum transfer rate of 200 Mb/s will be required to provide timely uploads of image data to the acquisition & distribution server.

4.4.5. Proposed Data Architecture

At a high level there are two main considerations for the data architecture:

- An attribute data model for the storage of image metadata.
- The relationships between metadata, the stored index maps, and the image data.

Once again the proposed data architecture is addressed under the 4 functional components of the system – Acquisition, Image Store Management, Discovery and Distribution.

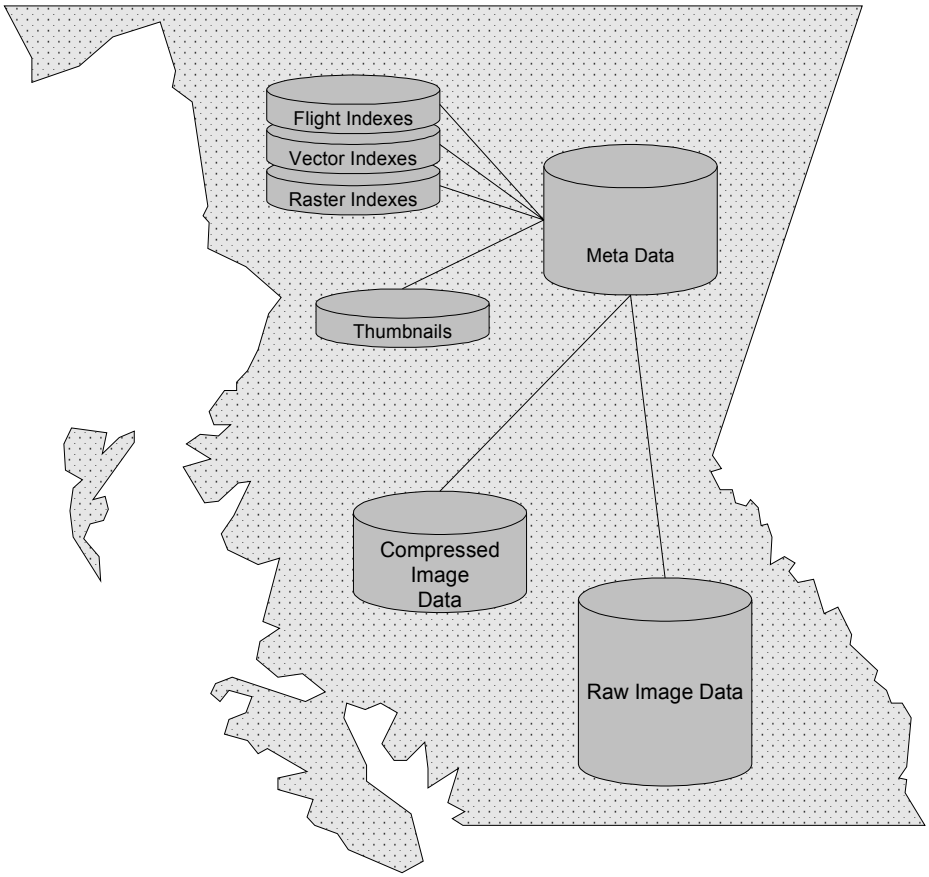


Figure 8 – Data Relationships

Proposed Data Architecture – Acquisition

Acquisition will provide the digital Image Management system with it's content. This will require the definition of business rules to define the naming and storage conventions for data that is being received and loaded to the system.

Metadata

Metadata will be loaded to the system during acquisition. See the subsequent section on Proposed Data Architecture – Image Data Store Management for metadata details.

Digital Imagery

Digital image products are the core of this system and represent essentially the entire anticipated system overhead. The subsequent section on Proposed Technology Architecture - Image Data Store Management details the projected volumes of all digital imagery for the initial load and subsequent 2-year period.

At present, digital image acquisition deals with the following image categories:

- 5” B&W aerial photography
- 10" B&W aerial photography (including Stereo Models)
- 10" Colour aerial photography
- Orthophoto
- Satellite imagery
- Controlled 10" aerial photography (B&W and Colour)

In addition to the above image products, control files or other additional non-metadata information files may be stored with each image.

Thumbnails

Each digital image product available in the Digital Image Management system will require a thumbnail image for cataloguing and discovery. Ideally these thumbnails would be created automatically during acquisition and update. Software selection for acquisition and/or Image Data Store management will determine the feasibility of this functionality. Additional analysis is required to determine the best approach and specifications for creating the thumbnails.

Sierra Systems proposed solution recommends that all thumbnail images be treated as image attributes and be stored in the Oracle RDBMS.

Proposed Data Architecture – Image Data Store Management

Once data has been loaded to the system Image Data Store Management will direct the storage and data transfer through the system. The following data types are critical to function of Image Data Store management.

Metadata

Sierra is recommending the adoption of the U.S. Federal Geographic Data Committee (FGDC) Standards for Digital Geospatial Metadata. This standard will provide the framework for modeling the metadata database.

The Standards for Digital Geospatial Metadata encompasses nine areas that can detail the following information about a digital image:

- Identification Information
- Data quality information
- Spatial data organization information
- Spatial reference information
- Entity and attribute information
- Distribution information and
- Metadata reference information

Note that thumbnails are part of this standard and are termed "browse-graphic".

Our analysis indicates that this standard will accommodate all current and legacy metadata maintained by the GDBC for imagery. This accommodation depends on the use of "user extensions" to enhance the metadata content and to encompass specific GDBC metadata.

Additional modeling analysis will be required to define the database content, structure and dependencies. This analysis will be done partly through examination of LDBC models. During that process investigations should be made into integration with the LDBC metadata repository.

Appendix "H" contains a summary of the FGDC Content Standards for Digital Geospatial metadata prepared by the US Geological Survey. The detailed FGDC Content standards for Digital Geospatial Metadata Specification can be viewed at: http://www.fgdc.gov/metadata/contstan.html Extensions to support remote sensing metadata can be viewed at: http://www.fgdc.gov/standards/status/csdgm_rs_ex.html.

Digital Imagery

In addition to the traditional data model, Image Data Store Management will also require a structure for data storage on the disk array. The implicit metadata found in current file-naming conventions for digital image products at GDBC suggests the use of a standard NTFS or UFS file system to house the image data. Figure 9 provides a sample of a typical file system for digital imagery.

Sierra Systems recommends that the management of the file system for digital image storage be automated through the Data Management services. This implies that the file system should not be

directly accessible to general users, and that access is transparent and controlled by the Data Management services.

One issue the Data Management services will make transparent is any UFS and NTFS file system size or directory file count limitations.

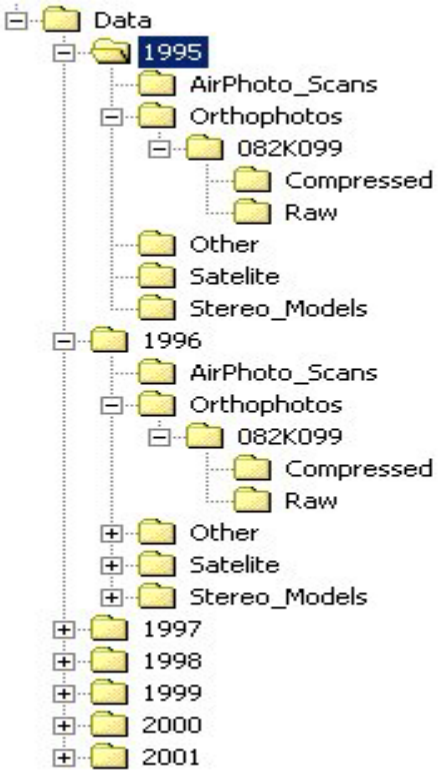


Figure 9 – Recommended NTFS File System for Digital Imagery (typical)

Proposed Data Architecture – Discovery

The Discovery module will rely on the data residing in the Image Data Store Management area to facilitate the search and discovery of images in the Digital Image Management system.

Metadata

The Discovery module will submit textual queries against the metadata database. Once the search criterion finds a hit in the metadata, Image Data Store management will return the related data to the user for selection.

Other metadata data requirements have been described in the previous Image Data Store Management Section.

Graphic Indexes

Discovery will also be capable of graphic queries based on a single geographic coordinate, a bounding box, and adjacency or buffering. The Oracle database will hold the various graphic indexes required for this feature. Sierra Systems is recommending that Oracle Spatial be used to store these indexes because of their static nature, the performance advantages that this could bring to the discovery process, and because of the enhanced spatial algebra offered by Oracle Spatial services. This implies that graphic indexes would be geo-referenced.

Graphic indexes considered include:

- Digital vector base maps and BCGS grid network
- Vector Flight indexes including:
  - Digital flight lines, photo centres and derived photo footprints
- Low resolution (thumbnail) imagery backdrops including:
  - Satellite and orthophoto

Thumbnails

Once Discovery has found a hit through either query method, the Digital Image Management system will optionally offer a thumbnail image of the digital image that has been located. Sierra Systems is recommending that Oracle Spatial be used to store thumbnails.

Proposed Data Architecture – Distribution

Distribution will provide for the delivery of digital image products on line or through transportable media.

Digital Imagery

Once an image has been identified for download or distribution the Distribution Services will deliver the image data either as a direct download to the users site or through transportable media.

Metadata

Distributed images will require accompanying descriptions and statistics that will be found in the stored metadata. This will require the definition of business rules based on image type to define what minimum metadata is extracted from the metadata base for delivery with the image.

Additional design is required to determine the business rules or procedures for metadata extraction.

Other metadata data requirements have been described in the previous Image Data Store Management section.

4.4.6. Proposed Use of Standards

The Architectures Working Group (US Military) C4ISR Architecture Framework stipulates that a technology architecture consists of three primary elements: Operational, Technical and Systems. Although this document is not explicitly structured according to C4ISR, it has addressed the Operational (tasks, activities and workflows) and Systems (servers, network, programs, storage) components.

The Technical view of C4ISR looks at a minimal set of rules governing the arrangement, interaction and interdependence of system parts or elements, whose purpose is to ensure that a conformant system satisfies a specified set of requirements. This *Proposed Use of Standards* section completes the Technical Architecture by discussing standards for communication between the Systems components of DIM and also for communication between DIM and GDBC Customer's systems.

Sierra Systems recommends the following standards be adopted for DIM:

- The *OpenGIS Service Architecture* should be taken as the framework for design and implementation of DIM components. In fact, DIM recommendations to date from Sierra and other parties have remained consistent with this component interaction model. This specification does not contain specific measurable stipulations for service or communication, but does provide a model for overall design. Conformance to this architecture should provide for primary interchange ability amongst vendor supplied products, and should provide for standardized interfaces and outerfaces.
- The *OpenGIS Catalogue Services* should be conformed to when modelling DIM Discovery and ordering services or when purchasing SCOTS for this purpose. This standard provides for an internationally recognisable communications and format standard for metadata and for search services. This specification is not explicit as to content, and so the *Federal Geographic Data Committee Standards for Digital Geospatial Metadata* – Geo Profile should be adopted to provide an internationally recognisable content standard for metadata. Alongside this, the ISO Z39.50 protocol for communication and search services should be adopted.
- The OpenGIS *Simple Features Specification For OLE/COM* and/or *for SQL* should be adopted as standards for the internal request and transmission of vector data between DIM modules. This standard will have to be approached in a pragmatic manner and non-compliant interfaces allowed if justified by cost or performance factors.
- The OpenGIS *Web Map Server Interface* should be adopted as a standard for the design of service and data transfer specifications for the vector map server and web front end applications, including those that serve imagery. It may not be feasible to enforce this emerging standard in the near term due to non-compliance by legacy software applications. However, it should represent a mandatory requirement for all newly implemented applications.
- For the disk arrays the following protocols must be supported: NFS v3 over UDP or TCP; PCNFSD v2 for NFS client authentication; and Microsoft CIFS with NT using Net BIOS over TCP/ IP.

5. OTHER RELATED ISSUES

This section describes a possible higher level organization and goal set for DIM than is focused on in the project for which this report is a deliverable.

This description is presented here in order to assist in understanding where the current project fits in a broader scheme of things, and to assist the reader to understand where issues exist that this document does not address.

5.1. A Strategic DIM Project

5.1.1. DIM Project Objectives

Although a definitive exposition of the DIM Strategic objectives have not yet been published, the following tentative strategic DIM project objectives and scope were obtained by examination of previous DIM project deliverables and through discussions with the Project Sponsor.

- Revise GDBC data management operations and infrastructure to support the current and growing business demand for geographic information.
- Replace the use of in-house systems with commercial off the shelf systems in order to remove barriers to change and to lower risk and costs.
- Set the framework for systems initiatives at GDBC in order to optimize support for its operations and to enhance their effectiveness as managers of spatial data in the Province of British Columbia.
- Identify and document business processes, data flows and associated data management tasks and to define models, processes, tools and roles and responsibilities in order to optimize investment and operational spending on an ongoing basis.
- Take advantage of and be proactive about the renaissance in Geographical Information System (GIS) technologies regarding managing spatial data holdings together with other corporate information. This is projected to increase systems utility and to decrease systems infrastructure costs. Aspects of the renaissance are: newly available vendor offerings for data storage, services and metadata; emergence of general GIS standards; DBMS support for GIS data storage; emergence of GIS specific communication protocols; and formulation of general metadata standards.
- Track imagery, other spatial data, and associated attribute information across the GDBC organization thereby eliminating the administrative challenges and shortfalls that exist in today's management systems.
- Enable discovery, access and commerce systems for access to GDBC through the Internet in order to reduce software distribution and support costs, to optimize interfaces with primary customers and to vastly increase the pool of potential customers.

- GDBC will not allow itself to be limited in the long term by current intergovernmental and private industry agreements, but recognises the need to migrate to new business and systems in a non-disruptive manner while providing continuous service.

5.1.2. DIM Project Scope

Digital Image Management comprises all management activities related to the creation, acquisition, assembly, updating, verification, storage, manipulation (including conversion and transformation), access control, representation, distribution, and archiving or removal of GDBC managed data.

It is expected that 80% of GDBC's needs and requirements for optimized systems will be addressed over a five-year time span by the DIM project.

The following products and data outline the scope of managed data:

- Scanned aerial photos for archive- Black and White, Colour and Infrared
- Scanned aerial photography stereo-models for TRIM or other softcopy photogrammetric mapping
- Digital aerial imagery – Monochromatic, Colour or Multi-spectral
- Aerial photography index maps in raster and vector format
- Digital Orthophotos
- Digital Mosaics
- Raster versions of vector maps
- Scanned index maps or other rasterized vector data.
- Digital Satellite Images
- Baseline Thematic Mapping products
- Scanned maps
- Metadata regarding all of the above, and TRIM, BTM, DRA, EBA and gridded DEM.
- Control point data and management activities (MASCOT and ACS)
- Watershed data and activities (TWA and Watersheds BC)
- Toponymy data (BCGNIS and Enhanced toponymy)

Operations included are:

- Data acquisition projects and project management
- Primary data management, storage, backup and managed change control
- Customer access to product/data catalogues and metadata

- Customer ordering and payment
- Continuous improvement

5.2. Data Management Architecture Review & Recommendations

Two reports produced previously as part of a broader DIM project formed a primary input to the *Current Assessment and Future Direction* project. These are:

- *Data Management Architecture Review & Recommendations*, Holonics Data Management Group Ltd. , 22 March, 2001
- *Geographic Data BC – Transition Strategy and Migration Plan*, Holonics Data Management Group Ltd. , 27 April, 2001

The *Current Assessment and Future Direction* project was directed to take into account and incorporate the image management process and recommendations of the *Data Management Architecture Review & Recommendations*. The *Current Assessment and Future Direction* project was also directed to prepare a migration strategy outlining the steps required to achieve the implementation and goals of the *Current Assessment and Future Direction* project. This road map should comply with the *Geographic Data BC – Transition Strategy and Migration Plan*. (See *Project Charter*).

Appendix A. Current Business Process Models

Appendix B. "To Be" Business Process Models

Appendix C. Client Workshop Notes

Appendix D. Stakeholder Requirements Summary Table

Appendix E. GDBC Workshop Notes

Appendix F. Digital Imagery Storage Comparison Analysis

Appendix G. Storage Media and Infrastructure Research

Appendix H. FGDC Standards

Appendix I. ANSI/NISO Z39.50 Protocol

Appendix J. Open GIS Consortium Standards