

Report

Status:

Version:

Date

WP 2

Final

2.0

2000-09-30

Requirements and structure for Storage and Archive

Appendix 9

Requirements and structure for Storage and Archive

Final version 1.0

September 30, 2000

Harry Hietanen

National Land Survey of Sweden

Table of Contents

- 1 Requirements and structure for Storage..... 1**
- 1.1 Background..... 1**
- 1.2 Objectives/Aim 1**
- 1.3 Method and Technique 1**

- 2 Result of investigations 2**
- 2.1 Formats 2**
- 2.2 Storage 2**

- 3 Checklist and Questions..... 3**
- 3.1 Questions for each product..... 3**
- 3.1.1 Storage hierarchical in Computer system 3**
- 3.1.2 Storage in Oracle databases 4**
- 3.2 Test server/domain 6**

Report	Status:	Version:	Date
WP 2	Final	1.0	2000-09-30
Requirements and structure for Storage			Page 1
WP2-App9			

1 Requirements and structure for Storage

1.1 Background

To handle 13 000 files for the production and for the prototype system it's necessary to have some kind of large storage system. The amount and volume of original and compressed files is very large.

The National Land Survey of Sweden treasures a large archive of historical maps and their descriptions. Today, the archive in Gävle contains approximately 300 000 deeds/maps. A part of these maps have been scanned in 1992-1995 and are stored on tapes in several digital raster formats (TIFF, etc).

To keep track of the large number of deeds/maps, metadata have been stored in a MS Access 2.0 database based on older analogue registers and after inventories have been made. The large amount of data in the Access database has increased the response time and caused problems for multi-user use. Within this background the database has been transferred into an Oracle database.

Parallel to the DHM project the NLS has obtained a new computer system (Geoserver) for storage of raster data and databases. The process to build up a state of the art, functioning, secure, scalable and good archive system is on going and in the very beginning.

1.2 Objectives/Aim

The objectives are:

- To find out the structure for storage data in computer system
- To provide an easy way to search and query the database.

The aim of the specification is to describe the structure and content of the database.

1.3 Method and Technique

In view of the project description and the needs from the prototype the task was to find out underlying demands for functionality, availability and quality for production, transferring, long term storage and for use in the prototype.

There are several questions to take care of (see below 3 Checklist and Questions).

All answers are not represented because lack of relevance for this application.

Meetings and interviews have been made with personal from NLS Data service.

Facts about storage, Web-solution, database and communications in each country have been examined.

New processes for transferring data have been developed.

Report	Status:	Version:	Date
WP 2	Final	1.0	2000-09-30
Requirements and structure for Storage			Page 2
WP2-App9			

Together all these activities have affected on the requirements and the structure for the storage solution.

2 Result of investigations

As specified in Image production the file-formats TIFF, TIFF Group IV and MrSid format.

Storage and archive of original files in the Geoserver with slow access.

Compressed files must be stored in the Web server with fast access to fulfil the user needs of high availability.

The prototype should be in same application for each country, there is no defend for build up three separate prototypes with different computer architectures. NLS have already a good basis for the whole system. So the prototype are recommended to be implemented in the NLS Geoserver system.

To achieve a high level of data security in the accessible database it is recommended to have it inside the firewall in the NLS Geoserver. A test database should be inside the firewall, only for testing of data, applications and other internal use.

The Webserver is within reach of external users as a common Internet application.

Appendix 9-1: Computer System in Sweden

Appendix 9-2: Principals for Hierarchical Storage Management, HSM.

Appendix 9-3: Calculated mean search time for each media type.

Appendix 9-4: Draft to catalogue structure for DHM data in 3 different HSM

2.1 Formats

Original files are stored as TIFF and compressed as MrSid format.

2.2 Storage

Files for the prototype system are stored on harddisk in a Webserver (NLS).
Sweden.

Original files and compressed copies are stored on tape (DLT7000) in a HSM system.

Denmark:

Original files and compressed copies are stored on CD's and stored of Manage unit in KMS.

Germany:

Original files and compressed copies are stored on CD's.

Report

Status:

Version:

Date

WP 2

Final

1.0

2000-09-30

Requirements and structure for Storage

Page 3

WP2-App9

3 Checklist and Questions

To use as an checklist and basis for requirement specification in relation with configuring and adjustment to IT-Systems

Appendix:

Appendix 9-1: Computer System in Sweden

Appendix 9-2: Principals for Hierarchical Storage Management, HSM.

Appendix 9-3: Calculated mean search time for each media type.

Appendix 9-4: Draft to catalogue structure for DHM data in 3 different HSM

3.1 Questions for each product

3.1.1 Storage hierarchical in Computer system

Responsible for maintenance: _____

Group responsible: _____

Product / data set: _____

1. Are all roles of responsible and contact persons known?
2. Shall some specifications change for the product?
3. Naming of catalogues and files?
Based on the actual status, investigate needs for changes. Is it possible to automate the changes, for example when changes are made in the storage system?
4. How to design the catalogue structure?
Co-ordination with other applications using the same storage system!
5. How to handle earlier versions of the same items? For example from different epochs/ revisions.
6. Demands for volumes:
 - Size of files?
 - How many exists today?
 - Volume today?
 - Increase next year and forwards (five years)?
7. Demands for reading/writing data to the storage?
See estimated mean times for different media types.
8. Is there any pattern in access of information?
Does this imply new levels for parts of the data set when setting up the HSM system?
9. Are there any plans to use or develop a search/retrieve system for access of the underlying file structure? If so, when?

Report	Status:	Version:	Date
WP 2	Final	1.0	2000-09-30
Requirements and structure for Storage			Page 4
WP2-App9			

Will the Oracle DBMS be used for linking search concepts, overview images and the basic data storage?

10. Are there any other software necessary to run the application needed to be installed on the system?
11. Demands for access? Hours a day, weekends, holidays?
12. Are there any further security demands to clear out?
13. Need for user groups with separated access rights?
Maintenance personnel, etc.
Contact person for administration of users and access to single catalogues/files (ACL)?
14. Time schedule for conversion work?
 - Start time and rough estimation of work.
 - Demands and priorities towards other applications.
 - Allocation of resources for the work.
15. Documentation of the system functions for end users.
 - Are ther any descriptions today?
 - When to make completions and changes. Routines for information.

3.1.2 Storage in Oracle databases

Responsible for maintenance: _____

Group responsible: _____

Product / data set: _____

1. Are all responsibility roles and contact persons known?
For example Oracle-DBA.
2. Are any specifications for the products to be changed?
3. Demands on storage volume:
 - Status today?
 - Increase of volume next year and in the future (five years)?
4. Version of the Oracle DBMS? Demands based on this?
5. Are there any other software affecting the DBMS, for example SDE and FME?
6. Availability during the day, in weekends and holidays?
7. Security matters to be cleared out?
8. Need for user groups with separated access rights?
Maintenance personnel, etc.
Contact person for administration of users and access to single catalogues/files (ACL)?
9. Time schedule for change to the new storage environment?
Start time and rough estimation of work.
Demands and priorities towards other applications
Allocation of resources for the practical work.

Report

Status:

Version:

Date

WP 2

Final

1.0

2000-09-30

Requirements and structure for Storage

Page 5

WP2-App9

Report

Status:

Version:

Date

WP 2

Final

1.0

2000-09-30

Requirements and structure for Storage

Page 6

WP2-App9

3.2 Test server/domain

Use of specialised test domain before implementation in production environment.