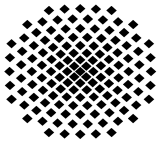

ADSM as a Tool for long-term Archiving?

University of Stuttgart
Regional Computing Center
Gerhard Rentschler
g.rentschler@rus.uni-stuttgart.de

Agenda

- Introduction
- What is long-term Archiving needed for?
- Requirements 1
  RAS
- Requirements 2
  Ownership, Access Control, Quota
- Requirements 3
  Managing changing Technology
- Requirements 3
  Managing changing Technology
- Requirements 4
  User Interface
- Requirements 5
  Metadata
- Does ADSM meet the Requirements?

Introduction

- Books in Libraries are kept for centuries
- What about digital data?
- Long-term archiving: keeping data for more than 5 years
- typically 10 to 30 years
- guarantee for readability of the raw data
- Long-term archiving has special requirements
- IBM doesn't promise ADSM will meet the requirements
- Nevertheless, one might be tempted to use ADSM for this purpose
- This presentation will show the problems

What is long-term Archiving needed for?

- Results of a poll among University of Stuttgart Computer users
  - data used as a base for expert opinions (typically 10 years)
  - Project Documentation, source code, test reports, mails, contracts, minutes of meetings, pictures, photographs, video sequences (typically 10 years)
  - measurement data (e.g. environmental or climatic research, more than 20 years)
- Problem: Long-term archiving cannot guarantee that application-specific data can be used
  - e.g. audio sequences, images etc need application for reproducing
  - need to archive application with source as well along with documentation

  $\Rightarrow$ owner has to care for this
- No direct access to data with open, read, write etc necessary

Requirements 1
RAS

- Availability of System 95 % may be sufficient
- Automatic recycling of media, e.g. copy data every 3 years
- Safety of Data much more important
  - Less than 1 KB per Terabyte in 10 years
  - requires at least 2 or 3 mirrors
  - probably on different platforms
- System may offer different safety classes at different cost

Requirements 2
Ownership, Access Control, Quota

- Unix style ownership not sufficient
- Project data is kept long after the end of the project
  - Definition of inheritance of data is needed, e.g. data ownership goes to department head or to the university
  - Need to register University, department as user
  - Need to define who pays for storing the data after the end of the project
  - No storage of data without expiration date
- Access rights:
  - write and read
  - delete
  - list own objects
  - change expiration date
- Quotas for all types of users and for groups of users

Requirements 3
Managing changing Technology

- Support change to new hardware like new tape technology
  - need long-term vendor commitment to product support
  - probably source code is needed for adapting to new platforms
  - need automatic copy functions
- Support change to a new Software-System
  - need export function for data and metadata for being able to move to a new archiving software
  - Format of exported data needs to be documented

Requirements 4
User Interface

- 20 years ago I used a HP 2648 Terminal
- 10 years ago I used an IBM 3270 Terminal
- Today I use Pentium PC with Win NT
- What will I use 20 years from now?
- Need an architected Interface to the archiving system
- For the time being and several years ahead a WWW interface with Java might be the best choice
- Encryption might be needed for transport over public networks
- Hierarchical Structure needed for storing objects
- Links would be nice

Requirements 5
Metadata

- Archiving system should offer functions for describing the data
  - e.g. format of data, description in free text, categories
- Database is needed with search functions

Does ADSM meet the Requirements?

- ADSM Drawbacks:
  - Ownership primarily based on computer system
  - no cross-platform retrieve between Unix and PC-platforms
  - no metadata, ADSM DB not extendible
- Enhanced ADSM or develop separate System
  - ADSM is very useful for backup
  - probably not wise to overload ADSM with new function
- Short term enhancement
  - make cross-platform restore possible
    - via tar-like front end?
    - Adsmpipe
    - unfortunately no API for PC Platforms