A Distributed Approach to Autonomic Long-Term Digital Preservation

CS PhD Seminar - 27.05.2010

Ivan Subotic
Overview

• Motivation
  - Digital Preservation
  - Scenarios

• Few Key Requirements

• Related Work

• **DISTributed ARchival NETwork**
  - Introduction
  - Data Model
  - System Overview
  - DISTARNET (Self-*) Processes

• Next Steps
Overview

• Motivation
  - Digital Preservation
  - Scenarios

• Few Key Requirements

• Related Work

• DISTRIBUTED ARCHIVAL NETWORK
  - Introduction
  - Data Model
  - System Overview
  - DISTARNET (Self-*) Processes

• Next Steps
Digital Preservation

• What is digital preservation?
  - Long-term preservation of information available as digital data

Digital preservation as communication with the future
(modification of [Moore 2008], [Moise 2009])
Digital Preservation (2)

• Physical Representation
  - decay of the physical media in use
    - Optical Disk (CD): 2-50 Years
    - Magnetic (Tape): 2-30 Years
    - Magnetic (Disk): 5-10 Years
Digital Preservation (3)

• Bit-Stream

  - Obsolescence of hardware
    - Kuny: 18 Months, Rothenberg: 5 Years
    - CD-ROM: 1985 - today (>25 yrs.)
    - 3.5” Floppy-Disks: 1987 - today (>23 yrs.)
    - USB: 1996 - today (>14 yrs.)
    - IDE: 1986 - 2003 (17 yrs.)

  - What does obsolete mean?
Digital Preservation (4)

• Logical Representation

  - Obsolescence of software and file formats
    - Software changes over time. Is the old document version still compatible?
    - Software is simply abandoned over time. How can we open a 20 year old WordStar document?
Digital Preservation (5)

- Solution Approaches
  - Migration:
    - media refresh
    - technology migration
    - format migration
  - Emulation:
    - Hardware emulation
  - Computer Museum:
    - hardware and software archiving
A Distributed Approach to Autonomic Long-Term Digital Preservation

Overview

• Motivation
  - Digital Preservation
  - Scenarios

• Few Key Requirements

• Related Work

• **DISTributed ARchival NETwork**
  - Introduction
  - Data Model
  - System Overview
  - DISTARNET (Self-*) Processes

• Next Steps
Scenarios (1)

• Scenario 1: Small Research Institution (A)
  - Characteristics:
    - 15 researchers and 0.25 IT department
    - Data: Images, Audio, Video, Documents
    - single site
  - Goal:
    - geographically distant, redundant (3 copies) and secure storage
    - preservation policy-based fully automated solution
    - through collaboration extend archival network
    - the possibility to share the archived material with other institutions
    - add additional metadata (annotations, links, collections/subcollections)
    - strict access restrictions
    - expendability regarding storage space
A Distributed Approach to Autonomic Long-Term Digital Preservation

Scenarios (2)

- **Scenario 2: Large Corporation (B)**
  - **Characteristics:**
    - IT department but no dedicated archiving specialist
    - Data: Images, Video, Sound, Documents
    - strict preservation policies regarding distance between storage sites and country of storage
    - multiple site
  - **Goal:**
    - geographically distant, redundant (3 copies) and secure storage
    - preservation policy-based fully automated solution
    - through collaboration extend archival network
    - sharing some of the data
    - add additional metadata (annotations, links, collections/subcollections)
    - strict access restrictions
Scenarios (3)

• Conclusion
  - Different scenarios but similar / same functionality
  - A and B could be research partners
  - Goal is a solution that can encompass both scenarios
Overview

• Motivation
  - Digital Preservation
  - Scenarios

• Few Key Requirements

• Related Work

• Distributed ARchival NETwork
  - Introduction
  - Data Model
  - System Overview
  - DISTARNET (Self-*) Processes

• Next Steps
User Functionality (1)

• Ingest
  - Import of digital objects and related metadata into the system (simple and complex objects)

• Access
  - Depending on the access rights, to be able to search and access data over the network
  - Since archiving solution, search does not need to be efficient/fast but effective
User Functionality (2)

• **Links**
  - The ability to make arbitrary virtual connections between different digital objects
  - Provides additional possibilities in finding other digital objects of interest and browsing through the archive

• **Annotations**
  - The ability to extend the metadata of digital objects

• **Collections/Subcollections**
  - Grouping of arbitrary digital objects
Functional Requirements

• Access Rights Management
  - access restrictions on the user, group and virtual organization level
  - all user operations will be restricted accordingly

• Fully automated system -> Self-* Properties
  - self-configuration
  - self-healing
  - self-learning
Non-Functional Requirements (1)

- High degree of scalability
  - number of nodes
  - number of users
  - volume of content to be archived
Non-Functional Requirements (2)

• Compliance to the Reference Model for an Open Archival Information System (OAIS)

OAIS Reference Model ([OAIS 2002])
Overview

- Motivation
  - Digital Preservation
  - Scenarios

- Few Key Requirements

- Related Work
  - Distributed ARCHival NETwork
    - Introduction
    - Data Model
    - System Overview
    - DISTARNET (Self-*) Processes

- Next Steps
Related Work (1)

• centralized approach
  - kopal and SHERPA DP
    (+) easy control of archived content
    (-) high infrastructure and maintenance costs
    (-) inflexibility to cope with rapidly increasing loads if they exceed the initial design

• distributed approach
  - LOCKSS, Cheshire 3, and SHAMAN
    (+) lower infrastructure and maintenance costs
    (+) the ability for virtually unlimited growth, and
    (+) allow for a higher degree of availability and reliability.
    (-) through distribution higher complexity regarding the management of archived content
Related Work (2)

• Conclusion
  - distributed approach is the way to go
  - **missing**: full automation \(\rightarrow\) self-* properties
Overview

- Motivation
  - Digital Preservation
  - Scenarios
- Few Key Requirements
- Related Work
  - Distributed ARchival NETwork
    - Introduction
    - Data Model
    - System Overview
    - DISTARNET (Self-*) Processes
- Next Steps
Introduction

• What is DISTARNET 2.0?
  - Distributed ARCHival NETwork
  - Followup on the original DISTARNET idea but completely redesigned
  - Vision: A collaboration between many institutions who take part in the DISTARNET network to deliver a long-term digital archiving solution based on P2P-GRID technology
  - Self-* properties to automate and minimize administration, and to provide an efficient resource allocation and utilization

• Why?
  - There is still no satisfactory fully automated long-term digital archiving solution out there
Overview

• Motivation
  - Digital Preservation
  - Scenarios

• Few Key Requirements

• Related Work

  • DISTributed ARchival NETwork
    - Introduction
    - Data Model
    - System Overview
    - DISTARNET (Self-*) Processes

• Next Steps
Data Model

• Information Object (IO)
  - logical container holding data objects (images, audio/video, text, etc.)
  - metadata: annotations, links, collection/subcollection sets, etc.
  - technical metadata: e.g., file format descriptions, etc.
  - preservation metadata: e.g., checksums, etc.

- Logical Data Model for DISTARNET 2.0:

- Information Object (IO) = Archival Information Package (AIP)
Overview

• Motivation
  - Digital Preservation
  - Scenarios

• Few Key Requirements

• Related Work
  • **DISTributed ARchival NETwork**
    - Introduction
    - Data Model
    - **System Overview**
    - DISTARNET (Self-*) Processes

• Next Steps
System Overview
Layers (1)

- **P2P-GRID Layer**
  - **GRID:**
    - basic-services (used by the Distributed Workflow Subsystem)
    - system-services (authentication/authorization and VO)
    - based on Globus Toolkit 5.0
  
  - **P2P:**
    - building and maintaining the structured P2P network
    - resource discovery,
    - access and management of the global information storage (DHT - Distributed Hash Table)
    - Kademlia based structured network

- **Data Repositories at this Level:**
  - Local: Node Information Repository (NIR) - location, country, free/used space, etc.
  - Global (DHT): Global Replica Repository (GRR)
Layers (2)

- **DISTARNET 2.0 Network**
  - P2P-VOs:
    - multiple VOs with own structured P2P network
    - VOs need additional external data and will be created manually
    - nodes can be part of multiple VOs
    - resources of a node will be reserved per VO
Layers (3)

- **Content Management Layer**
  - Responsible for Management of:
    - information objects, i.e. the archived digital objects (e.g., image, audio/video, text document, etc.)
    - corresponding metadata (e.g., annotations)
    - relations (e.g., links, sub-/collections) between other information objects
    - properties (e.g., access rights, availability requirements)
    - technical metadata
    - preservation metadata
  
- Allows:
  - r/w access to information objects via OID (subject to access rights)
  - partitioning of the storage area for multiple VOs
Layers (4)

- Access Management Layer
  - Comm. to the “outside world”:
    - Ingest Services (used by the content providers)
    - Access Services (used by the users)
A Distributed Approach to Autonomic Long-Term Digital Preservation

DWS

- Distributed Workflow Subsystem
  - Provides
    - decentralized global service invocation
    - reliable and fault-tolerant distributed execution and management
    - operations implemented as basic-services (P2P-Grid layer)
  - Contains
    - DISTARNET processes implemented as workflows
  - Implementation
    - based on OSIRIS
Overview

• Motivation
  - Digital Preservation
  - Scenarios

• Few Key Requirements

• Related Work
  • DISTributed ARchival NETwork
    - Introduction
    - Data Model
    - System Overview
    - DISTARNET (Self-*) Processes

• Next Steps
Self-* Processes (1)

- Self-Configuration
  - ability of the system to automatically detect and react to changes in the network (nodes joining/leaving)
  - underlying structured P2P network is self organizing

- Automated Dynamic Replication (ADR):
  - finds suitable storage nodes by using information from the Node Information Repository
  - policy-based restrictions are taken in to account
  - estimating the optimal number of replicas
  - evaluation of existing availability of the IO in the network
  - initiating the creation of replicas (if required)
Self-* Processes (2)

- Self-Configuration example: Ingest Process
  - The Access Management Layer receives an Ingest Request and triggers the Ingest Process
  - Content Management Layer stores it to the filesystem, generates an unique ObjectID (OID), and initiates checksum generation.
  - Metadata Manager stores corresponding representation information (structure and semantic information), the preservation metadata (e.g., checksum), and other properties (e.g., access restrictions, availability requirements, etc.) to the database records linked to the OID.
  - Indexing Service indexes the newly created information object
  - ADR evaluation:
    - optimal (PP)
    - current replica numbers (GRR)
  - current < optimal:
    - find suitable storage nodes (NIR s.t. PP)
    - initiate creation of additional replicas
Self-* Processes (3)

- Self-Healing

<table>
<thead>
<tr>
<th>Event</th>
<th>DIST 2.0 Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node-Lost</td>
<td>ADR via info from the GRR</td>
</tr>
<tr>
<td>Corrupted Data Object</td>
<td>ADR via info from the GRR</td>
</tr>
<tr>
<td>Unreliable Connection</td>
<td>ADR via info from the NIR</td>
</tr>
<tr>
<td>Low Storage Capacity</td>
<td>ADR via info from NIR</td>
</tr>
</tbody>
</table>

- Node-Lost Event:
  - react and initiate countermeasures by reevaluating IOs
  - information in the Global Replica Repository is continuously refreshed. When a node is lost, then this information

- Corrupted Information Object:
  - periodic integrity checks ensure detection
  - the ADR process restores IO and redundancy
Self-* Processes (4)

- Self-Learning
  - Adaptive Parameters:
    - periodicity of triggers will be dynamically adapted to the state of the network
  - Information Dissemination:
    - dissemination of the information needed for node operation through attaching small bits of information to existing communication messages
    - e.g., the Node Information Repository stores data distributed this way
Overview

- Motivation
  - Digital Preservation
  - Scenarios
- Few Key Requirements
- Related Work
  - DISTRIBUTED ARchival NETwork
    - Introduction
    - Data Model
    - System Overview
    - DISTARNET (Self-*) Processes
- Next Steps
Next Steps

• Short Term
  - Implementation (3 Layers and DWS)
  - Improve the DISTARNET Self-* Processes
  - Optimize the Information Dissemination Process

• Medium-Term
  - Evaluate prototype in the lab

• Long-Term
  - Evaluate the system in the “real world” with partners
A Distributed Approach to Autonomic Long-Term Digital Preservation

Literature

- [Kopal], http://kopal.langzeitarchivierung.de
- [SHAMAN], http://shaman-ip.eu/shaman/
Literature (2)