Information Management Along the Lifecycle of Data and Application Systems: Challenges and Solution Approaches

Georg Fischer
SAP AG
ge.fischer@sap.com

Axel Herbst
SAP AG
axel.herbst@sap.com

Abstract: A tried and tested method for improving the quality of information in the business system and control data volume growth, is moving selected, eligible data from the application system to less expensive long-term storage. To ensure that the eligibility of the data is actively taken into account, methods such as archivability checks must be part of this process. Descriptive policies play into the strategy, to allow for the compliant retention of the moved data along its entire life cycle. A further challenge is controlled destruction to complete the life cycle of data according to the law. At the same time, the fact that often the life span of application systems is shorter than that of the data itself, needs to be considered. In this presentation we show how long-standing archiving techniques for business data are challenged anew by the growing complexity of today's legal and business requirements, causing an evolution to information lifecycle management.

Agenda

1. Current Demands on Information Management Given the Lifecycle of Information
2. Data Archiving: Managing Data Volumes and Improving the Quality of Information
3. Retention Management: Dealing with Legal Compliance Aspects in Information Management
4. System Decommissioning: Taking the System Life Cycle Into Account
5. Summary: From Data Archiving to Information Lifecycle Management
Information Management Demands Customers are Facing Today

<table>
<thead>
<tr>
<th>Increasing Complexity</th>
<th>Internal Drivers</th>
<th>External Drivers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>External Drivers</strong></td>
<td>– High costs for hardware and administration</td>
<td>– Legal Retention Requirements</td>
</tr>
<tr>
<td>– Policies and service level agreements</td>
<td>– Product liability</td>
<td>– Product liability</td>
</tr>
<tr>
<td>– Risk of litigation</td>
<td>– Lawsuits (Legal holds, e-Discovery)</td>
<td>– Tax Reporting, Audits</td>
</tr>
<tr>
<td>– Company-specific processes</td>
<td>– New technologies</td>
<td>– New technologies</td>
</tr>
<tr>
<td>– System landscape harmonization/centralization</td>
<td>– Mergers and acquisitions</td>
<td>– Legal Retention Requirements</td>
</tr>
<tr>
<td>– Mergers and acquisitions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Lifecycle of Data in a Business Application System

<table>
<thead>
<tr>
<th>Database</th>
<th>File System</th>
<th>Storage System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non changeable</td>
<td>Business complete</td>
<td>Audit</td>
</tr>
<tr>
<td>Creation</td>
<td></td>
<td>Lawsuit</td>
</tr>
<tr>
<td>Residence time</td>
<td>Lawrence</td>
<td>Information Destruction</td>
</tr>
<tr>
<td>Access Frequency</td>
<td>Time</td>
<td></td>
</tr>
</tbody>
</table>
Data Analysis and Classification to Facilitate Archiving and Retention of Data

Bottom Up Approach
- Based on the complete content (filled tables) of the system
- Relevant for data extraction

Top Down Approach
- Based on the legal and reporting requirements and those tables that are required for these reports
- Relevant for reporting

General strategy: Combination of both approaches

Agenda

1. Current Demands on Information Management Given the Lifecycle of Information
2. Data Archiving: Managing Data Volumes and Improving the Quality of Information
3. Retention Management: Dealing with Legal Compliance Aspects in Information Management
4. System Decommissioning: Taking the System Life Cycle Into Account
5. Summary: From Data Archiving to Information Lifecycle Management
Application-Oriented Database Archiving

Architectural classification by provider of archiving service

- DBMS-based
- DBMS-integrated

Advantages of the DBMS-Based Approach

- DBMS-independence
  - Availability: DBMS-integrated approach only in DB2 RAM (restrictions!)
  - Even “Standard” SQL needs to be unified

- Archive storage integration
  - Needs to be vendor-independent as well: Dedicated Content Mgmt or ILM-aware storage systems, connected via certified interfaces; physical location matters!

- Application “awareness” of archiving
  - DB schema hardly contains application semantics (almost no integrity constraints on DB level, business context for archive access, ...
Data Archiving Process

Agenda

1. Current Demands on Information Management Given the Lifecycle of Information
2. Data Archiving: Managing Data Volumes and Improving the Quality of Information
3. Retention Management: Dealing with Legal Compliance Aspects in Information Management
4. System Decommissioning: Taking the System Life Cycle Into Account
5. Summary: From Data Archiving to Information Lifecycle Management
Legal Compliance - Record Retention Periods

- **Pharmac./Life Sci. (21 CFR Part 11)**
  - 2 years after distribution: Records relating to the manufacturing, processing, packing of drugs and pharmaceuticals
  - 5 years after end of manufacturing or product: Records relating to the manufacturing of biological products

- **Healthcare (HIPAA)**
  - 5 years: All hospitals must retain records in originally or legally produced form
  - 21 years: Medical records for minors from birth to 21
  - 2 years after patient’s death: Medical records

- **Financial Services (SEC 17a-4)**
  - 3 years: Financial statements
  - End-of-life of enterprise: Member registration for broker/dealers

- **OSHA**
  - 30 years after completion of audit: Employee and medical records of individuals exposed to toxic substances
  - End of account +6 years: Trading account records

- **Sarbanes Oxley**
  - 4 years after completion: Original correspondences from financial audits or publicly-traded corporations
  - 21 years+: Medical records for minors from birth to 21

Source: Enterprise Storage Group, May 2003

Evolution of Archiving Functions Towards ILM

- **Archiving**
  - Archive business-complete objects after an archivability check
  - Policy engine determines how long and where the data will be stored

- **Snapshot**
  - No archivability check
  - Also writes non business-complete objects to archive files
  - No data deletion possible on the DB (always redundant data)

- **Data Destruction**
  - No archivability check
  - Only data that can be destroyed immediately is written
  - Allows the removal of non business-complete objects (old non closed processes)
  - File is not stored, but deleted
Define Retention Parameters: Policies and Rules

Example of Policies and Rules for Financial Documents in a Global Enterprise

<table>
<thead>
<tr>
<th>Subsidiary</th>
<th>SAP Company Code</th>
<th>Location of Data Retention</th>
<th>SAP-Residence Time in the Database (Days)</th>
<th>Legal Retention times (Years)</th>
<th>Retention Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>1000</td>
<td>D</td>
<td>1</td>
<td>10</td>
<td>CRep A1</td>
</tr>
<tr>
<td>Great Britain</td>
<td>2000</td>
<td>GB</td>
<td>1</td>
<td>7</td>
<td>CRep A2</td>
</tr>
<tr>
<td>USA</td>
<td>3000</td>
<td>USA</td>
<td>1</td>
<td>7</td>
<td>CRep A3</td>
</tr>
<tr>
<td>Canada</td>
<td>4000</td>
<td>CAN</td>
<td>1</td>
<td>30</td>
<td>Directory …/canada</td>
</tr>
</tbody>
</table>

Archive Hierarchy Based on Retention Rules

Archive Files & Snapshots

Attachments

Certified ILM WebDAV Interface

Certified ILM WebDAV Interface
Agenda

1. Current Demands on Information Management Given the Lifecycle of Information
2. Data Archiving: Managing Data Volumes and Improving the Quality of Information
3. Retention Management: Dealing with Legal Compliance Aspects in Information Management
4. **System Decommissioning**: Taking the System Life Cycle Into Account
5. Summary: From Data Archiving to Information Lifecycle Management
System Decommissioning Process

Multiple source systems are decommissioned into...

...one central ILM system connected to ILM-certified storage...

...and data can be presented for audits using data warehouse queries

Archiving Functions in System Decommissioning Scenario

Retention Policy Engine

<table>
<thead>
<tr>
<th>Existing Archive Files</th>
<th>ILM File Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online Database</td>
<td>Classic archiving and snapshots</td>
</tr>
<tr>
<td>Transactional Data</td>
<td>Context Data Extraction</td>
</tr>
<tr>
<td>Master Data</td>
<td>Context Data Extraction</td>
</tr>
<tr>
<td>Customizing</td>
<td></td>
</tr>
<tr>
<td>Data Dictionary</td>
<td></td>
</tr>
</tbody>
</table>
Agenda

1. Current Demands on Information Management Given the Lifecycle of Information
2. Data Archiving: Managing Data Volumes and Improving the Quality of Information
3. Retention Management: Dealing with Legal Compliance Aspects in Information Management
4. System Decommissioning: Taking the System Life Cycle Into Account
5. Summary: From Data Archiving to Information Lifecycle Management

From Data Archiving to Information Lifecycle Management

**Information Lifecycle Management**

**Data Archiving**
- Analyze and categorize data
- Improve information quality in online system by weeding out data from closed business processes
- Securely move data from the database to long-term storage
- Access archived data on demand

**Challenges:**
- Data access requirements vs. volume management
- Self-contained objects for long-term interpretability
- System-independence

**Retention Management**
- Manage all retention policies across the enterprise
- Manage responsible destruction of data based on policies
- Enforce retention policies
- Use secure ILM-certified storage
- Perform e-Discovery and set legal holds

**Challenges:**
- Media management (technological advances)
- End2End retention enforcement
- Thorough destruction

**System Decommissioning**
- Decommission legacy systems
- Enforce retention policies on data from shut-down system
- Run queries on data from decommissioned system
- Benefit from independently understandable archive

**Challenges:**
- SW-support/automation for old releases
- Heterogeneous legacy systems
- Adequate access methods

**Manage Database Volume**

**End-Of-Life Data**

**End-Of-Life System**
Benefits of Information Lifecycle Management

1 Data Archiving
   Improves information quality for online system. Higher system availability and better performance with shorter response times.

2 Retention Management
   Improves quality of retained data, based on legal and business requirements.
   Supports the complete lifecycle: creation – preservation – destruction.

3 System Decommissioning
   Takes system lifetime aspect of information management into account, by facilitating the long-term use of data, even though the original systems no longer exist. Helps avoid high costs of unnecessarily maintaining legacy systems.

References