The Business Case for Database Information Management

Julie Lockner, Data Management Consulting Services, LLC
Gary Zasman, Chair SNIA DMF LTACSI
The material contained in this tutorial is copyrighted by the SNIA.

Member companies and individual members may use this material in presentations and literature under the following conditions:
- Any slide or slides used must be reproduced in their entirety without modification
- The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.

This presentation is a project of the SNIA Education Committee.

Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.

The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.
Abstract

The Business Case for Database Information Management

- Databases consume significant resources in the data center. There are several information management strategies that when applied to these databases, significant savings and improved efficiencies can be realized. Users should walk away with a good understanding of the principles behind some of the more standard Data Information Management practices and how to look at the business case for justification.

- At this point we are only addressing a subset of the numerous Data Information Management practices and principles.
About the SNIA DMF

This tutorial has been developed, reviewed and approved by members of the Data Management Forum (DMF)

❖ The DMF is an industry resource to those responsible for the accessibility and integrity of their organization’s information

❖ The DMF focuses on the technologies and trends related to Data Protection, ILM and Long-term digital information retention

<table>
<thead>
<tr>
<th>DMF Workgroups:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Protection Initiative (DPI)</td>
</tr>
<tr>
<td>Defining best practices for data protection and recovery technologies such as Backup, CDP, Data deduplication and VTL</td>
</tr>
<tr>
<td>Information Lifecycle Management Initiative (ILMI)</td>
</tr>
<tr>
<td>Developing, educating and promoting ILM practices, implementation methods, and benefits</td>
</tr>
<tr>
<td>Long-term Archive and Compliance Storage Initiative (LTACSI)</td>
</tr>
<tr>
<td>Addressing the challenges of retaining, securing, and preserving digital information for the long-term</td>
</tr>
</tbody>
</table>
Agenda

- What is Driving Enterprise Database Information Management
  - Basic Principles

- What is Database Information Management (D.I.M.)
  - How you can save with a D.I.M. Strategy

- The Benefits
  - Improve performance
  - Control costs
  - Mitigate risks

- Building your D.I.M. Return On Investment (ROI)
  - Getting Started
What is driving Enterprise Database Information Management?

Business Owners Need for:
- Compliance
  - Controls for SOX, GLB etc
  - Improve records retention
  - Ensure processes are in compliance
  - Allow records to be discovered for litigation
- Business Information Needs
  - Improve access to reference data
  - Facilitate information mining
  - Use info for organizational improvement
- Reduced Cost of Ownership
  - Optimize infrastructure costs
  - Minimize people costs
  - Improve efficiencies

IT & DBA’s Need for:
- Systems Efficiency
  - Reduce high cost storage needs
  - Reduce load on servers and DB
  - Reduce backup resources
  - Reduce recovery resources
- User Productivity
  - Remove inactive data to improve application performance
  - Reduce backup time
  - Reduce recovery time
  - Improve availability
  - Ease of access to retained data
- Cost Reduction
  - Optimize storage costs
  - Minimize people costs
  - Improve Disaster Recovery
  - Improve and speed up dev/test/QA
D.I.M. Basic Principles

- **Assess & Define Requirements**
  - Assess Application types and storage tiers
  - Define tiered service level and access requirements
- **Classify based on Policy & Service Objectives**
  - Identify business objects to archive
  - Determine retention requirements
  - Identify use cases e.g. audit, eDiscovery, dev/test etc
- **Archive**
  - Develop operational practices
  - Define management architecture
- **Store**
  - Design hardware architecture and targets i.e. tiers etc.
  - Establish security requirements
- **Access**
  - Communicate access policies and capabilities e.g. SLAs etc
- **Delete**
  - Develop deletion policies with all stake holders
Managing **Structured Content**

- Policies & Best Practices
- Classifications & Taxonomies
- Efficient Storage Pools & Service Levels
- Archiving & Deletion
- Test Data Management
- Confidentiality & Privacy
Database Information Classification?

Classifying content in a database based on an Information-Centric Taxonomy

- By Legal, Business, Security, Compliance Classifications
  - Classified, Public, Sensitive, Regulated, Aged, Transaction Status
- By Business Object defined by Database Entities
  - Database, Table, Row, Partition, System
  - Business Object Status & Sensitivity
- By usage
  - Production, Test & Dev, Training, Reporting, Backup, D/R
Matching SLA’s to Information Lifecycle

Tiered Service Level Agreements

Active ➔ Less Active ➔ Historical ➔ Offline Archive

- Storage Pool 1
  - FC
  - Data SLA 1

- Storage Pool 2
  - FC/SATA
  - Data SLA 2

- Storage pool 3
  - SATA/WORM
  - Data SLA 3

- Offline Pool
  - Offline Media
  - Data SLA 4

Use ITIL-style ILM service catalogs and processes to develop a dynamic and consolidated storage service & Save $
Benefits of SLAs

▶ Arrest Spiraling Storage Costs
  ◆ Inefficient asset utilization due to point solutions or project based acquisition
  ◆ Keeping pace with compound annual storage growth rates of 50-100%

▶ Respond Faster – Accelerate Business
  ◆ Provision new projects and applications faster than ever
  ◆ Rapid and reliable recovery across all application tiers

▶ Remove Risk
  ◆ Repeatable processes, standardized designs and the ITIL aligned management framework drives down risk, support costs, and time to market.
Database Archiving Principles

- Migrate Business Object to Archive based on Policy
  - I.e. Closed Transaction Older than 2 Years
- Maintain appropriate levels of access to archived data
  - Determines online / offline archive
  - Native access, separate reporting system, SQL
- Database Archiving Tools are available to assist
  - Database features, independent vendors, application tools

Sample Database Archiving Configuration

App Specific View
- Pool 1
  - SLA 1

App Independent Viewer
- Pool 2
  - SLA 2
- Pool 3
  - SLA 3

Deletion
Reduce Cost & Improve Performance

**Before**

- Mirror 1000GB
- PROD 1000GB
- Patch 1000GB
- Test 1000GB
- Training 1000GB
- Dev 1000GB
- PROD Backup 50% 500 GB

**Prod (1000 GB) x 6.5 = 6.5 TB**

**After**

- Mirror 500GB
- PROD 500GB
- Patch* 500GB
- Test* 500GB
- Training 500GB
- Dev* 500GB
- Archive DB ** 500GB
- PROD Backup 50% 250 GB
- Archive Backup *50% 250 GB

**Prod (500 GB) x 6.5 + Archive DB (500) x 1.5 = 4 TB → ~40% reduction**

* Additional storage efficiency solutions also available, i.e. dedup & snapshots

** Archive can be stored on WORM media, lower cost storage**
Test D.I.M. Principles

- Determine Test Data Set based on Usage Classification
  - Full copies
    - Staging environment, performance benchmarking
  - Subset copies
    - Training, Functional testing
  - Replication & Cloning options
    - Full clones / app / data, sync / async, etc.

Sample Testing Configuration

- Production, DR, Staging
  - Full copies
  - Data Subsets on lower SLA Pools

- QA, Training, Patch
  - Full Copies with highest SLA
Reduce Cost & Test Cycles

Before Test Data Management

- Mirror: 1000GB
- PROD: 1000GB
- Patch: 1000GB
- Test: 1000GB
- Training: 1000GB
- Dev: 1000GB

Test Storage Requirements
- 1000GB x 4 = 4TB
- Cloning Time
- 1 to 3 Days x 4 = 1 to 3 Man Weeks

After Test Data Management

Snapshots & Subsetting

- Mirror: 1000 GB
- PROD: 1000 GB
- Patch: 100 GB
- Test: 100 GB
- Training: 50 GB
- Dev: 50 GB

Test Storage Requirements
- 300 GB (90% reduction in storage)

Cloning Time
- 1 to 3 Hours x 4 = ¼ to ½ Man Days
- (83% reduction in time)
The Old Way

- Production
- Mirror
- Dev 1
- Test 1
- QA
- Test 2
- Dev 2

Lengthy cloning, requires extra storage

The New Way

- Production
- Mirror
- Dev 1
- Test 1
- QA
- Test 2
- Dev 2
- Writable Snapshots

Fast cloning, requires less storage
The Business Case for Database Information Management
© 2009 Storage Networking Industry Association. All Rights Reserved.

Accelerating and Improving Testing

Time

The Old Way

Cloning

TEST

Cloning

TEST

The New Way

TEST

TEST

TEST

TEST

TEST

Create Writable Snapshot
Copies

- Improved Speed
  - DB copies made in minutes

- Improved Quality
  - More copies to more people
  - Time to do more testing
Maintain Confidentiality

- Assign Masking Policies to Business Objects based on Usage (i.e. Production) and Sensitivity
  - Understand testing requirements as well
  - Encryption alone is not a comprehensive masking strategy

Sample Data Masking Policy

<table>
<thead>
<tr>
<th>Data</th>
<th>Type</th>
<th>Policy</th>
<th>Masking Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer Credit Card</td>
<td>(string)</td>
<td>Sensitive</td>
<td>PCI standards - Encryption or Masking</td>
</tr>
<tr>
<td>Employee SSN</td>
<td>(string)</td>
<td>Sensitive</td>
<td>Null</td>
</tr>
<tr>
<td>Corporate Sales Forecast</td>
<td>(number)</td>
<td>Sensitive</td>
<td>Substitution</td>
</tr>
</tbody>
</table>

The Business Case for Database Information Management
© 2009 Storage Networking Industry Association. All Rights Reserved.
How Can you Save? Hard Savings

- **Tiering & Archiving**
  - By removing aged data from production, all copies will be smaller

- **Test Data Management**
  - By creating subsets or snapshots, additional copies will be smaller & test cycles will be faster

- **Security & Confidentiality**
  - By masking sensitive data, risk of unauthorized access is more easily controlled

- **Optimize Infrastructure to Meet Business Requirements**
  - Storage - both disk and tape backup
  - Server requirements – CPU usage and utilization
  - Software licenses – CPU based licenses
How Can you Save? Soft Savings

▶ Tiering & Archiving
  ◆ Reduced cloning times, smaller backup / recovery windows
  ◆ Reduced upgrade and patching downtime
  ◆ Reduced performance related maintenance costs

▶ Test Data Management
  ◆ Reduced cloning times, smaller replication windows
  ◆ Reduced test cycles
  ◆ Faster time to market with new features

▶ Security
  ◆ Cost avoidance – risk mitigation
Benefits

❖ Improve Performance
  ◆ Improved availability
  ◆ Speed backup and recover and dev/test
  ◆ Improve application performance

❖ Control Costs
  ◆ Reduce expensive production storage
  ◆ Reduce license fees and hardware costs
  ◆ Reduce labor overhead
  ◆ Reduce app decommissioning and dev/test costs

❖ Mitigate Risks
  ◆ Store in an immutable format that cannot be altered
  ◆ Index for easy retrieval for e-Discovery
Getting Started – Assess what you have

- Inventory all databases
  - For each database, determine total annual cost
  - Include License, Server, Storage, & People costs
- Factor in Database Growth Rates
  - Future annual costs need to consider growth of production and how it impacts all copies

Define classifications & Business Objects

- Quantify how much the Business Objects are consuming
  - As a percentage of the total cost
  - Typically does not include Master / Setup objects
  - Only those tables that contain data that could be archived or subset
Determine D.I.M. Policies

- Define what is needed –
  - For Each Business Object
- Production Systems
  - Define Retention & Archive Policies w/ Business Owners
  - Quantify how much could be archived
- Non-Production Systems
  - Define Test Data Policies w/ Developers
  - Quantify how much could be used for testing
  - Quantify how much sensitive data resides in copies
Prioritize Projects Based on Business

- Risk of Not Applying D.I.M. Principles
- Cost / Benefit Analysis
- Readiness of organization to support & enforce policies

Distribution reports show data volumes by business process

Growth charts illustrate forecast requirements for storage
Cost Savings Estimates

**Business As Usual – vs. – with Database Information Management**

- **Storage Savings**
  - $$/GB savings
  - Cost Avoidance
- **Server Savings**
  - Server consolidation
  - Avoid server upgrade costs
- **Software license Savings**
  - CPU based license costs
- **People Savings**
  - DBA time spent performance tuning or waiting around for a clone
- **D.I.M Investments**
  - Determine what tools / costs need to be deployed to implement D.I.M.
  - Need to factor in these costs for a true ROI calculation
## Sample Assessment

### Archive Potential by Business Object

<table>
<thead>
<tr>
<th>Application</th>
<th>Retention Policy (months)</th>
<th>Data Volume Before Archiving (GB)</th>
<th>Data Volume Eligible for Archive (GB)</th>
<th>Data Volume After Archiving (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Ledger</td>
<td>24</td>
<td>54</td>
<td>15.84</td>
<td>38.16</td>
</tr>
<tr>
<td>Accounts Payables</td>
<td>12</td>
<td>47.93</td>
<td>13.14</td>
<td>34.79</td>
</tr>
<tr>
<td>Workflow</td>
<td>1</td>
<td>28.42</td>
<td>4.38</td>
<td>24.04</td>
</tr>
<tr>
<td>Content Management</td>
<td>12</td>
<td>1.77</td>
<td>0.77</td>
<td>1</td>
</tr>
<tr>
<td>Accounts Receivables</td>
<td>12</td>
<td>32.54</td>
<td>11.9</td>
<td>20.64</td>
</tr>
<tr>
<td>Inventory</td>
<td>12</td>
<td>0.11</td>
<td>0</td>
<td>0.11</td>
</tr>
<tr>
<td>Human Resources</td>
<td>12</td>
<td>25</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Fixed Assets</td>
<td>12</td>
<td>4.12</td>
<td>0.4</td>
<td>3.72</td>
</tr>
<tr>
<td>Project Accounting</td>
<td>12</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Other ERP Apps</td>
<td>95</td>
<td>6.07</td>
<td>0</td>
<td>6.07</td>
</tr>
<tr>
<td>Custom Tables</td>
<td>95</td>
<td>0.97</td>
<td>0</td>
<td>0.97</td>
</tr>
<tr>
<td>ERP Non-Archive Tables</td>
<td>95</td>
<td>12.75</td>
<td>0</td>
<td>12.75</td>
</tr>
<tr>
<td>System Tables</td>
<td>95</td>
<td>22</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>235.78</td>
<td>56.43</td>
<td>179.35</td>
</tr>
</tbody>
</table>

**Before** 236 GB  
**After** 179 GB  
**Savings** 56 GB
Make It happen

- Typical benefits
  - 100 to 500% ROI
  - Payback period < 12 months
- Calculate your ROI
  - Look for tools from vendors & software features that can help

![Graph showing accumulative yearly savings and investment over time](image)
Interested in Getting Involved?

- Join the Database Information Management Special Interest Group

  http://www.snia.org/forums/dmf/programs/ltacsi/dim_sig/

- Participation will provide many opportunities for you and your company:
  - Ground-floor leadership role in developing database information management best practices and standards
  - Collaboration on educational, marketing, and outreach activities
  - Learn from academic and industry research, and to keep up-to-date on this rapidly-evolving area of our industry
  - An opportunity to be informed, to help frame research and standards development, and to help influence projects that SNIA will undertake
Please send any comments on this tutorial to SNIA at: trackdatamgmt@snia.org

The DMF would like to thank the following individuals for their contributions to the development of this tutorial:

Julie Lockner
Gary Zasman
Michael Peterson
Jan Rosenberg

It’s easy to get involved with the DMF!

- Find a passion
- Join a committee
- Gain knowledge & influence
- Make a difference

www.snia.org/dmf