Maximize Your Return on Big Data

May 15-18, 2012
What’s New From Informatica To Optimize Data Warehouse Performance & Lower Cost

Mukesh Marodia
Senior Product Manager

Claudia Chandra
Senior Director, Product Management
Informatica
Agenda

• Challenges of data growth in data warehouses
• Data Archive for Data Warehouse
• Demo of smart partitioning
• Nearline for SAP BW
• Customer case studies
Safe Harbor

The information being provided today is for informational purposes only. The development, release and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision. Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.
Challenges of Data Growth in Data Warehouses

- Escalating storage, server, and data warehouse cost
- Increasing effort spent on maintenance
- Diminishing data warehouse performance
- Multiplier effect of data growth in non-production copies
- Lack of visibility to how data is used to prioritize resources
Informatica Data Archive for Data Warehouse
Optimize Data Warehouses

Informatica Data Archive for Data Warehouse

Informatica Data Warehouse Advisor

• Improve Performance

• Reduce Costs

• Gain Visibility

• Ensure Compliance

Informatica Data Archive

Inactive data

Active data

Production

Development/Test/Training Copies

Copy 1

Copy 2

Copy 3

TIME

DATABASE SIZE

Performance

Maintenance
Lean Data Warehouses

Focus on optimizing and tuning the most relevant data – based on usage

Streamline Data Loads.  
Eliminate Loading Unnecessary Data.

Identify unused data and applications.  
Eliminate/retire unused applications and reports.  
Compress and archive data.
Informatica Data Warehouse Advisor (DWA) provides visibility into business activity, data usage and performance to manage data more efficiently, help justify costs and prioritize resources based on data utilization.

Integrated with BI Applications and Informatica PowerCenter
Data Warehouse Advisor
Identify Dormant Tables and Records for Archiving

**Capabilities**

- Monitor and report on unused tables and records that are candidates for archiving
- Identify who uses the data

**Benefits**

- Reduce storage, server, and DW license costs by reducing DW size
- Reduce backup, recovery, and upgrade window and effort
- Reduce DW outage due to maintenance activities

---

Informatica proprietary/confidential. Informational purposes only. No representation, warranty or commitment regarding future functionality. Not to be relied upon in making purchasing decision.
ILM Data Discovery
Discover Table Relationships in Data Warehouses

Capabilities
- Discover relationships (joins) between tables based on profiling matched values between table columns
- Alternatively, discover relationships based on join conditions in queries and reports
- Automatically create business entities based on discovered relationships
- Based on integration with Informatica Data Explorer product

Benefits
- Reduce manual effort required to determine relationships between tables and reverse engineering of business entity definition
- Reduce dependence on application skill-set
Lean Data Warehouses
Monitor Data Usage, then Partition and Archive

- Production Data Warehouse
- Partitioned Production Data Warehouse

PARTITION

- Smart partitioning
  - Individual table
  - Entity-based

- Improve performance, maintain in same database

Production Data Warehouse
Seamless Access
Online Archive Database

- Access archived data through production interface
- Move data out of production
- Less maintenance

Production Data Warehouse
Optimized File Archive

- Archive data to optimized file format for storage reduction
  - Compressed (up to 98%)
  - Immutable
  - Accessible
### Informatica Data Archive
#### Unified Information Lifecycle Policy

**Diagram:**
- Production Data Warehouse
- Partitioned Data Warehouse
- Online Archive Database
- Optimized File Archive

**Legend:**
- **Very Frequent**
- **Less Frequent**
- **Infrequent**
- **Rare**

<table>
<thead>
<tr>
<th></th>
<th>Very Frequent</th>
<th>Less Frequent</th>
<th>Infrequent</th>
<th>Rare</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Production</strong></td>
<td>100%</td>
<td>80%</td>
<td>60%</td>
<td>20%</td>
</tr>
<tr>
<td><strong>Partitioning</strong></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>5%</td>
</tr>
<tr>
<td><strong>DB Archive</strong></td>
<td>100%</td>
<td>75%</td>
<td>100%</td>
<td>2%</td>
</tr>
<tr>
<td><strong>File Archive</strong></td>
<td>100%</td>
<td>25%</td>
<td>20%</td>
<td>5%</td>
</tr>
</tbody>
</table>

- **Maintenance Window:**
- **Storage Size:**
- **Cost:**
Example of Entity-based Partitioning

**ORDER TABLE**

Partition by Order Mode

**ORDER_ITEM TABLE**

Partition by Order Mode from Order Table

**TABLE SPACE 1**

**TABLE SPACE 2**

Informatica proprietary/confidential. Informational purposes only. No representation, warranty or commitment regarding future functionality. Not to be relied upon in making purchasing decision.
Unified Lifecycle Policy

- Discover business entities or leverage prebuilt accelerators
- Analyze data distribution by profiling
- Monitor data usage and identify dormant data
- Centrally manage both partitioning and archiving from a single console

Informatica proprietary/confidential. Informational purposes only. No representation, warranty or commitment regarding future functionality. Not to be relied upon in making purchasing decision.
Smart Partitioning Demo…
Federated Access
Combined Access to Production and Archived Data

- Need to report against combined production and archived data?
- Need to discover data across multiple locations for compliance audits?

Use ILM Federation Option
- Access combined production and archived data virtually, w/o creating a physical data warehouse
- Access combined data on demand
Data Warehouse Archiving
Federated View

Production Data Warehouse
(less than 2 years old)

Archive Data Warehouse
(2 – 7 years old)

Optimized File Archive
(over 7 years old)
Data Warehouse Archiving
Lean Data Warehouse Health Check Program

• Data warehouse is just like any other databases
• Volume grows faster due to historical data
• Lean Data Warehouse Health Check Program
• Free assessment to run Data Warehouse Advisor to identify dormant data
• Data Archive to relocate dormant data to less expensive infrastructure
• Promotional bundle – lower than buying the 2 products separately
• Optimize your data warehouse infrastructure and reduce costs
Lean Data Warehouse Health Check

No Charge 14 day Engagement

Deploy and use Informatica DWA

• Track and measure user and application activity and their patterns and trends

• Identify unused or dormant data that are unnecessarily loaded, stored, and maintained

• Identify heavy user and application activity associated with poor query performance to guide tuning and optimization
# Lean Data Warehouse Health Check

## Assessments

<table>
<thead>
<tr>
<th>Category of Analysis</th>
<th>Reports</th>
</tr>
</thead>
</table>
| **Data Usage**       | Activity trends and patterns of peak periods  
                       | Schema and Table Utilization  
                       | Unused Data (Schemas/Tables/Columns/Records) |
| **User Activity**    | Top Applications and BI reports impacting data warehouse  
                       | Top users (and associated Applications, Reports)  
                       | Frequently used data impacting performance and their users |
| **Performance**      | Expensive query activity and users  
                       | Candidates for optimal indexing  
                       | Repetitive and problematic operations |
Informatica ILM Nearline for SAP BW
Manage Data Growth in SAP BW

- Moves infrequently used data from SAP BW to nearline storage
- Maintain access from SAP application
- Market leader for SAP BW nearlining, first certified with SAP nearline
Case Study: Large Financial Services Firm

Expects to Save $30M from Application Retirement and Live Archiving

<table>
<thead>
<tr>
<th>THE CHALLENGE</th>
<th>INFORMATICA ADVANTAGE</th>
<th>RESULTS/BENEFITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>• 90 applications to be retired from acquisition</td>
<td>• Seamless access to archived data</td>
<td>• Retirement to be completed within first year</td>
</tr>
<tr>
<td>• 160 TB is 2 Oracle ODS, growing at 75% / year</td>
<td>• Efficient, high performance archiving</td>
<td>• Expect to archive 10TB of data from production apps within first year</td>
</tr>
<tr>
<td>• Storage, backup costs too high</td>
<td>• Low impact (&lt; 5%) to production performance</td>
<td>• Additional 20% will be archived in subsequent years</td>
</tr>
<tr>
<td>• Batch process overruns</td>
<td>• Ability to restore 2TB of data in less than a day</td>
<td>• Anticipated $30M cost savings over 5 years</td>
</tr>
<tr>
<td>• Undersized processing power, slow performance</td>
<td></td>
<td>• Faster restore, reduced risk of data recovery</td>
</tr>
<tr>
<td>• Long outages due to database size</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Business Imperative:** Reduce costs from another bank acquisition, operational efficiency

**IT Initiative:** Application retirement & Database Archiving

Informatica proprietary/confidential. Informational purposes only. No representation, warranty or commitment regarding future functionality. Not to be relied upon in making purchasing decision.
400+ Informatica ILM Customers
Improve Efficiency and Reduce Costs

- **Performance Improvement**
- **Improve SAP BW Application Performance**
- **Manage SAP BW Data Growth**
- **Storage Reduction**
- **Archiving & Application Retirement**
- **Manage Oracle data growth**

---

**Archived 16 TB’s of data to improve application performance by 30% overall**

**Reduced TCO of SAP BW (30+ TB) implementation, achieving compression rate of 91% with low impact performance in BW reports using nearlined data**

**Nearlined largest InfoProviders in size, achieving compression rates of 98% and increasing the responsiveness of online system up to 10 times.**

**Saved $1.5M by reducing data volumes by 70%, improved performance 4x, eliminated storage and database growth**

**Retired 18 applications from M&A in 6 months, eliminated hardware & software costs, achieved payback within 9 months**

**Large financial services company retiring 90 applications from another bank acquisition, archive 10 TB of 160TB in 2 Oracle DB within first year, 20% in subsequent year, providing seamless access, $6M annual cost savings**
For Additional Information

- Hands-on lab
- Lean Data Warehouse white paper
Thank You