



Data Archival Testing



Abstract

Today, there is an exponential rise in the amount of data being generated by organizations. This explosion of data increases IT infrastructure needs and has an immense impact on some important business decisions that are dependent on proficient data analytics. These challenges have made data archival extremely important from a data management perspective. Data archival testing is becoming increasingly important for businesses as it helps address these challenges, validate the accuracy and quality of archived data and improve the performance of related applications.

The paper is aimed at helping readers better understand the space of data archival testing, its implementation and the associated benefits.

Introduction

One of the most important aspects of managing a business today is managing its data growth. On a daily basis, the cost of data management outpaces the data storage costs for most organizations. Operational analytics and business intelligence reporting usually require active operational data. Data that does not have any current requirement or usage, known as inactive data, can be archived to a safe and secure storage. Data archiving becomes important for companies who want to manage their data growth, without compromising on the quality of data that resides in their production systems.

Many CIOs and CTOs are reworking on their data retention policies and their data archival and data retrieval strategies because of an increased demand for data storage, reduced application performance and the need to be compliant with the ever changing legislations and regulations.¹

Data Archival Testing – Test Planning

Data Archival is the process of moving data that is not required for operational, analytical or reporting purposes to offline storage. A data retrieval mechanism is developed to restore data from the offline storage.

The common challenges faced during data archival are:

- Inaccurate or irrelevant data in data archives
- Difficulty in the data retrieval process from the data archives

Data archival testing helps address these challenges. While devising the data archival test plan, the following factors need to be taken into consideration:

Data Dependencies

There are many intricate data dependencies in an enterprise's architecture. The data which is archived should include the complete business objects along with metadata, that helps retain the referential integrity of data across related tables and applications. Data archival testing needs to validate that all related data is archived together for easy interpretation, during storage and retrieval.

Data Encoding

The encoding of data in the archival database depends on the underlying

hardware for certain types of data. To illustrate, data archival testing needs to ensure that the encoding of numerical fields such as integers also archives the related hardware information, for easier future data retrieval and display of data with a different set of hardware.

Data Retrieval

Data needs to be retrieved from archives for regulatory, legal and business needs. The validation of the data retrieval process ensures that the archived data is easily accessed, retrieved and displayed in a format which can be clearly interpreted without any time consuming manual intervention.

Data Archival Testing – Implementation

The data archival testing process includes validating processes which encompass data archival, data deletion and data retrieval.

Figure 1 below describes the different stages of a data archival testing process, the business drivers, the different types of data that can be archived and the various offline storage modes.



Figure 1: The Data Archival Testing Process



1

Test the Data Archival process

- Testing the Data Archival process ensures that business entities that are archived includes master data, transaction data, meta data and reference data
- Validates the storage mechanism and that the archived data is stored in the correct format. The data also has to be tested for hardware independence



2

Test the Data Deletion process

- Inactive data needs to be archived and moved to a secure storage for retrieval at a later point and then deleted from all active applications using it. This validation process would verify that that the test data deletion process has not caused any error to any existing applications and dashboards
- When archived data is deleted from systems, verify that the applications and reports are conforming to their performance requirements



3

Test the Data Retrieval process

- Data that has been archived needs to be easily identified and accessible in case of any legal or business needs
- For scenarios that involve urgent data retrievals, processes for the same need to be validated within a defined time period

Benefits of Data Archival Testing

The benefits of data archival testing are often interrelated and have a significant impact on the IT infrastructure costs for a business. Some of the benefits are:

Reduced storage costs

Only the data that is relevant gets archived and for a defined time period which reduces hardware costs and its maintenance costs significantly.

Minimize business outages

Archived data that is deleted from production systems does not have an impact on the related applications' performance and functionality, leading to smooth business operations.

Improved application performance

Data is retrieved faster; the network performs better as only relevant data is present in the production environment. All these factors enhance application performance.

Data Compliance

Easy retrieval and availability of archived data ensures higher data compliance with the legal and regulatory requirements.

Accomplishing all these benefits determine the success of a data archival test strategy.

Conclusion

Due to the critical business needs for data retention, regulatory and compliance requirements and a cost effective way to access archived data, many businesses have started realizing and adopting data archival testing. Therefore an organization's comprehensive test strategy needs to include a data archival test strategy which facilitates smooth business operations, ensures fulfillment of all data requirements, maintains data quality and reduces infrastructure costs.

About the Author

Naju D. Mohan

Naju is a Group Project Manager with Infosys with about 15 years of IT experience. She is currently managing specialized testing services like SOA testing, Data Warehouse testing and Test Data Management for many leading clients in the retail sector.

REFERENCES

1. 'Data overload puts UK retail sector under pressure', Continuity Central, February 2009
 2. 'Data Archiving, Purging and Retrieval Methods for Enterprises', Database Journal, January 2011
-

For more information, contact askus@infosys.com



© 2017 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.