Lower Environment Data Exposure

Meridian Technologies
5210 Belfort Rd, Suite 400
Jacksonville, FL 32257
meridiantechnologies.net

Overview

It is currently common practice in the IT industry to use production data in lower environments for use in software development, testing and training. The reasons for this are generally due to the complexity of production data and the difficulty in generating representative test data that meets all of the naturally occurring combinations of data conditions found in raw production data. Unfortunately, this practice dramatically increases the exposure an organization has to the possibility of private customer data being leaked outside of the organization either accidentally or deliberately. This increased risk of a data breach exposes the organization to legal, financial and reputational risk, including damage to the company's brand, fines, and even imprisonment.

Risks

Production environments have many safeguards in place to prevent loss or misappropriation of data. The production hardware is logically and/or physically separated from the rest of the organization's network. Access to these environments is tightly controlled, and user activity is closely monitored with logging and audit trails. These controls ensure accountability for employees who access production systems, which lower the risk of misuse, or misappropriation of sensitive data. Furthermore, software and databases developed for production use undergo scrutiny across many disciplines in an organization, such as Subject Matter Experts (SMEs), Data/Application Architects and Security professionals. These individuals are trained to assess the exposure risk of an organization's data and have accountability for any data breach.

More than 80% of an organization's data footprint exists in lower environments. This data is exposed to individuals that are not as well trained or adequately motivated to protect the organization's data. For this reason, **60% of a company's risk of a data breach is in their lower environments**. Also, the security controls and monitoring that exist in production environments do not exist in the lower environments. A user could conceivably write a query to copy all of an organization's customer data to a file on the user's local PC or laptop without triggering any alarms. From there, it would not be a difficult task to encrypt the data in such a way as to defeat any content monitoring applications designed to prevent accidental or intentional data exposure. **More than 50% of data breaches are caused by internal employees, accidental or otherwise, and most organizations use 3rd party sub-contractors develop and test their systems.**

Lower Environment Data Exposure

Contingent and Full Time resources, including junior/entry level resources have access to Non-Production environments

- System Administrators
- System Operators
- Database Administrators
- Programmers
- Analysts
- Testers
- End Users
- Users new to the organization for training purposes

Remediation

In order to mitigate the risks discussed above, a test data management plan should be put in place. This plan, at the very least, should severely limit the amount of sensitive data available in lower environments and ideally, completely eliminate instances of sensitive data in non-production environments. If sensitive data does need to exist in a lower environment, the same audits and controls should be put in place as if it were in production. A test data management plan should keep the following in mind:

Access Level	Roles
No access to row level sensitive data	System Administrators System Operators
Access only to data they are accountable for	Database Administrators End Users
Access only to de-identified subsets of data	Programmers Analysts Testers Users new to the organization for training purposes

In order to mitigate these data security risks and to assist our customers with executing an effective test data management plan, Meridian developed AcceleTest. The AcceleTest solution enables our customers to:

- Create de-identified test data from live production data in an efficient and secure fashion

- Create test data sets that provide a full representation of data combinations found in production
- Quickly isolate differences between test data sets (e.g., before and after test execution views) in support of test data analysis tasks

AcceleTest was designed to help businesses streamline and automate the processes associated with the creation and management of the data needed for the testing, development and user training for business applications. The capabilities available with the AcceleTest solution are focused on helping an organization improve how work related to test, development and training data is accomplished – Faster, Better, and Cheaper. We accomplish this by ensuring our solution can be adapted to your processes rather than force you to adapt to the needs of the tool. In developing AcceleTest, we kept your business goals in mind to ensure a solution designed to:

- **Reduce development cost:** Subsetting features allow you to right size your test data relative to testing needs thus reducing overall data storage costs. Automated processing and a streamlined workflow further reduce cost by eliminating expensive manual data recreation and refresh tasks. The result is better data faster and at lower overall cost.
- **Reduce project risk:** High quality data is needed for effective testing, and your product data sources offer the best data available. By sanitizing production data through use of data transformation functionality, you can fully leverage this quality data without the risk associated with unnecessary access to sensitive customer information.
- **Improve efficiency:** Automated discovery of data relationships and a requirements oriented subsetting interface eliminate the need to create complex SQL and so minimize the amount of data knowledge needed to define a data set. Once a data set is specified, data can be recreated at the push of a button. Specifications can also be cloned and modified as needed for re-use on future projects. Automated data comparison ensures differences in data sets can be quickly identified and properly analyzed.
- **Improve quality:** Better data translates into better testing. Data can be transformed to ensure compliance to data privacy policies as well as to support specific test requirements such as date/time windows (e.g. month end processing). Specialized data requirements can be met by appending hand crafted data sets to production samplings.