Io-Tahoe®

Case study

Smart data discovery helps financial services firm streamline data relationships for greater value

August 2017

About the company

This leading global financial services firm manages securities exchanges, equities services and capital markets processes, and offers a range of data services to improve analytics, mitigate risk and enhance efficiency. Having been in operation for over 40 years, the company has accrued a significant repository of internal data assets, stored across servers that live on the Sybase platform and an array of data centers. The sprawling nature of this data real estate prompted the firm to seek a more streamlined approach to data management.

The situation: a server migration project predicated on better understanding data

The firm was exploring options to support a new migration project, aimed at consolidating their data centers and shifting platforms from Sybase to Oracle. This would require moving 80 servers.

The main objective was to achieve a better understanding of their data. Given the volume of their data assets, understanding how data flows through internal systems and how one element relates to another was key to yielding value from this data in the first place. Understanding how the applications were connected with one another would also help in establishing the order of database migration.

By better understanding its own data relationships, the firm hoped to manage and leverage that insight to better position their company for making business transformation decisions backed by accurate, reliable data.

The initial approach: mobilizing a small army of experts

Approximately 20 subject-matter experts (SMEs) were recruited to pore through the data and work toward three distinct goals:

- Migrate and re-platform 80 servers from Sybase to Oracle
- Understand data relationships within and across these databases
- Simplify and remove redundant data, enabling a smoother transition between platforms

The end result would, ideally, have been 80 replatformed servers, now living on Oracle, with quality data relationships mapped out and redundancies both identified and eliminated – all in all, a streamlined, consolidated and more efficient system, that made managing and leveraging data assets much easier.

However, this approach quickly presented two serious problems.

The challenge: manual data discovery only scratches the surface

For one, manual migration of all 80 servers would have taken the SMEs four years to accomplish.

In fact, just analyzing the sheer volume of data on those servers to reduce redundancies, eliminate identical data sets, centralize data flow management and streamline a complex maze of data relationships would have taken between 18 and 24 months. Suffice to say, that was much too slow and inefficient. The company wanted a fast and efficient way to consolidate and process the value of their data relationships.

Compounding the problem was the fact that this effort would still only scratch the surface of the data relationships. Despite a dedicated team of experts to analyze and review the outcomes, it would still be impossible to manually analyze every data element. Even a best-effort attempt would still result in a selective data discovery process, only covering the most critical elements. The scope of this analysis would be limited in the amount of data relationships discovered and still proceed too slowly, covering only about 10 to 15 elements per week.

While the team was comprised of specialized SMEs, human error is inevitable. Manual discovery is prone to mistakes and limited by the knowledge of its users, even ones who are subject-matter experts in their field. Given the length of time that manual data discovery takes, it is not farfetched to think that those elements would change several times in between the time discovery started and when it finally finished. The results would be rendered obsolete as soon as the data elements changed.

Moreover, data discovery is not a one-time process. Rather, it has to be executed multiple times to ensure that data relationships are accurate, reliable and free of redundant data. A years-long, intensive manual process cannot meet this outcome.

The solution: automated smart data discovery with Io-Tahoe

After deciding the manual approach was not realistic, the Head of Data Office, who reports to the CIO, turned to Io-Tahoe's smart data discovery solution to help uncover 'hidden' data relationships and redundant data. The executive and the Io-Tahoe team agreed to an initial trial run. As part of this arrangement, Io-Tahoe's automated smart data discovery solution would run on four of the 80 servers. Once the onboarding process was complete and server access provided, Io-Tahoe was able to run in the environment, processing 785 tables, 8,800 columns and 1.6 billion rows of data elements – in other words, seven times more data than the SMEs could process manually.

All within 14 days.

This wasn't just limited to data discovery, either; it also included conducting a real-time impact analysis and removing redundant data found throughout the system. As a result, the SME team only needed to review the output produced by Io-Tahoe's final analysis, effectively reducing their commitment by fivefold. In addition to an expedited migration, reductions in test time and error rates also produced tangible financial benefits, including decreased testing costs.

The benefits weren't solely focused on speed and convenience. With automated smart data discovery, 'hidden' or undocumented data elements were uncovered, all while increasing accuracy and defect resolution and reducing human error. The automated nature of Io-Tahoe means it can run multiple times across the same databases to incorporate updated information. All of this achieved the firm's ultimate goal: to make understanding data relationships one of their core competencies.

The Head of Data Office decided to use Io-Tahoe's automated discovery tool across the other 76 servers. When the company first set out on this migration project, they were looking at an 18-24-month effort; by automating data discovery and analysis, Io-Tahoe effectively cut that time in half.