



SyncSort

Success Story: Centers for Medicare & Medicaid Services

When It Comes to Processing Sensitive Data, The Centers for Medicare & Medicaid Services Relies on SyncSort

Organizational Profile

- Serves Medicare and Medicaid beneficiaries as well as healthcare providers

Business Need

- Speed up join and aggregate processing for a business intelligence application

Environment

- Sun E10,000
- Solaris

Benefits

- Improved the performance of the application
- Utilize the advanced merge and join functionality
- Effectively handles large volumes of data
- Works as a complement to Informatica

"We then turned to SyncSort because there's so much data that we just needed the performance."

Bruce Crawford
Lead Systems Engineer
VIPs

The Challenge:

The Centers for Medicare & Medicaid Services (CMS) is developing a process that can provide ongoing, in-depth analysis on the tremendous amount of claims data received from each state. This will enable CMS to determine trends, how money is spent, the public impact of certain diagnoses and drugs, and much more. In order to handle the processing and analysis of all Medicaid claims data, the Medicaid Statistical Information System (MSIS) project has been developed. Louis Gamerman, IT Specialist at CMS, described the project, "MSIS provides a summary of data at the state level. It consists of three data marts that are used for querying: the State Summary mart contains Medicaid summarized information by states, the Annual Person Summary mart includes all of the Medicaid data summarized by individual, and the Clinical Services mart, which we're currently working on now, includes drug and diagnosis related information." Prior to populating the data marts, the data is formatted and aggregated using a complex process.

"The reason we set up the data marts is because we're not allowed to query the data warehouse," Gamerman explained. "It's a giant relational DB2 structure on the mainframe and that structure is not indexed for querying. Also, because of the sheer volume of data involved, querying would bring the warehouse to its knees. Instead, we extract flat files from the underlying DB2 tables, where each flat file is equivalent to one table, and then we convert them from EBCDIC to ASCII. For a state summary we have six large files that we bring down with a combination of claims and beneficiary eligibility information, and we need to join them together and perform aggregates using complex business rules." In order to complete the necessary processing, Gamerman turned to the consultants at VIPs to help develop the application.

The Solution:

According to Bruce Crawford, Lead Systems Engineer at VIPs, "The MSIS takes in a huge amount of data and the CMS group had to calculate various sums and distinct counts across almost 1500 permutations of the keys. It was a requirement of CMS to use Informatica to perform all of the business logic in the application. We then turned to SyncSort because there's so much data that we just needed the performance. We knew that we could count on SyncSort to complete the work and to return in a respectable amount of time. The product has been performing like a real workhorse and we've also been taking advantage of its advanced merge and join functionality."

The application now uses a combination of SyncSort and Informatica steps, along with custom C programs, to process approximately 100 to 200 gigabytes of year-to-date claims data each quarter. The flat files for each state and Puerto Rico are processed separately. The files are ordered and joined

Success Story: Centers for Medicare & Medicaid Services



together using SyncSort, while Informatica is used to apply the business rules and perform the aggregation. "The performance is drastically better using SyncSort prior to performing the complex aggregations. We also do a lot of passes and a lot of steps between SyncSort and Informatica. Applying the business rules can get complex, and then we've got to do other tasks such as lookups, plus there are dimension tables involved, so we're doing surrogate key generation. And because of the volumes of data, SyncSort provided us with the only sensible way to join the files," Gamerman added.

For the MSIS, the data is processed on a quarterly basis using a Sun E10,000 running Solaris. The aggregates are then loaded into the three data marts and used to populate COGNOS cubes for querying. The information can be used to create reports that show a state's Medicaid expenditures by type of service and age group, Medicaid beneficiaries by type of service and age group, Medicaid expenditures for programs broken down by maintenance assistance status and basis of eligibility, and much more.

CMS has also decided to integrate SyncSort into its Healthcare Information System (HCIS) project because of its similar requirements for high-volume number crunching and summarization. Currently in development, the HCIS involves the processing and analysis of all Medicare claims. As with the MSIS, the claims level data in the HCIS is migrated from the mainframe to the UNIX server, then a series of SyncSort and Informatica steps are completed. Crawford explained, "We compress the data, performing approximately 200 permutations of the data, and then load it into Oracle tables. We complete the entire process with SyncSort and Informatica. We also use all of SyncSort's features to do our joins, merges, sorts and more."

The Benefits:

Gamerman added, "The primary reason why we chose SyncSort for both the MSIS and HCIS projects was because we needed a really powerful tool that could manage the volumes of data that we were talking about. SyncSort was the solution that could handle the sheer volumes of data and the sheer volume of claims we had to join together. It has also made efficient use of disk space and memory."

About SyncSort

SyncSort is a high-performance application accelerator that improves the performance of multiple applications and reduces elapsed time for a broad range of applications. It speeds ETL transactions by up to 90% and facilitates data mining and click-stream processing. SyncSort merges, aggregates, cleanses, and converts data. Other features include filtering, pattern matching, and partitioning. SyncSort will save you time in operations like data warehousing, data mining, data marts, CRM, ERP, DSS, BI, Oracle Financials, and legacy migration. Using Visual SyncSort, SyncSort applications can be created through a Windows-like GUI. For more information or to arrange for a free trial, call Syncsort at (201)930-8200 or visit www.syncsort.com.