

COVERING THE WORLD OF DATA: BI, INTEGRATION & ANALYTICS

VOLUME 14, NUMBER 7, JULY 2004

DMTM REVIEW

www.dmreview.com • www.dataWarehouse.com



syncsort

www.syncsort.com

SyncSort for Windows/NT

SyncSort Slashes Patient Data Processing Time by 60% for Presbyterian Health Plan

REVIEWER: Bruce Alexander, data warehouse manager for Presbyterian Health Plan.

BACKGROUND: Presbyterian Health Plan (PHP) is operated as a division of Presbyterian Healthcare Services, New Mexico's largest locally owned healthcare system. We offer a statewide healthcare delivery system and 17 years of experience in managed care. PHP provides more than 280,000 members with a comprehensive provider network, a quality medical management program, commitment to customer service and cost-effective, consumer-driven managed healthcare services.

PLATFORMS: We use Intel running Windows NT Server 4 with four CPUs and 1.5GB of memory. The database engine is Oracle8i.

PROBLEM SOLVED: We need to extract approximately 37 million rows of medical claims data for our disease and population management program. The rows must be sorted in patient and service date sequence prior to input into another product, Symmetry's Episodic Treatment Group, which examines the data and identifies medical episodes. We were using SQL and PL SQL to prepare the data. We would unload seven gigabytes of raw patient claims data from an Oracle database in delimited format. Next, we would sort the data on the patient ID and service date. Additional processing was then performed in order to group the data by clinically relevant classifications. The data was then loaded into Oracle for subsequent analysis. The process was taking more than five hours, and we started hitting some of the limitations of what we could do with the SQL tools – particularly where sorting was concerned. Because I previously worked with SyncSort on a mainframe, I decided to test the NT version to see if it would improve

MANUFACTURER



SyncSort for Windows/NT

Syncsort Incorporated
50 Tice Boulevard
Woodcliff Lake, NJ 07677
201-930-8200
www.syncsort.com

the performance of the application. With SyncSort, I was able to achieve the performance results I needed. Now the application begins with a simple text file extract from Oracle. The output of this script is read by SyncSort, which sorts the records in the proper patient and service date sequence prior to grouping and reloading back into the Oracle database.

PRODUCT FUNCTIONALITY: Using SyncSort, we created an application to achieve the performance results we needed. The application helped us reduce the sort processing time by more than 60 percent. Additional tuning of the SyncSort process has resulted in further gains.

STRENGTHS: We selected SyncSort because of its ability to quickly process a large amount of

data. Once we started using it, we found other capabilities such as the ability to merge or split files.

WEAKNESSES: It would be helpful to have the ability to assign a random number or introduce truly random sorts as opposed to having to add bolt-on extensions to it. Likewise, the ability to perform basic mathematical functions would be another helpful feature. However, as new features are being added to SyncSort, we are hopeful that these features will be added to the product.

SELECTION CRITERIA: When I started looking for a better solution to accelerate this application, I turned to SyncSort because of its reputation. I received an evaluation copy, and we made sure that it would process our data within an acceptable time frame.

DELIVERABLES: Using SyncSort, we were able to cut the procedure down to well under two hours, slashing processing time by more than 60 percent. Once the data is processed, our Medical Staff Affairs Group is able to perform the subsequent disease management, population management and utilization management analysis.

VENDOR SUPPORT: I talked to the Syncsort technical support staff about helping us create a random sampling extension to help solve another problem. They gave us an approach to assign a random number to individual records. We used that as a sort key in order to put the data in a truly random fashion for sampling. In all instances, we have found the technical support staff to be very helpful.

DOCUMENTATION: The documentation was very good and met all our needs.