

LEISTUNGSFÄHIGE DATENVERARBEITUNG

Centers for Medicare & Medicaid Services:
tiefgreifende Analysen von Abrechnungsdaten

Die Centers for Medicare & Medicaid Services (CMS) haben einen Prozess für fortlaufende, umfassende Analysen der riesigen Mengen von Abrechnungsdaten entwickelt, die sie aus jedem US-Bundesstaat erhalten. Damit kann die Behörde Trends, Ausgabenposten, Auswirkungen bestimmter Diagnosen und Medikamente auf die Öffentlichkeit und vieles mehr ermitteln.

Für die Verarbeitung und Analyse der gesamten Abrechnungsdaten von Medicaid entwickelten die CMS das Medicaid Statistical Information System (MSIS). Laut Louis Gamerman, IT-Experte der CMS, bietet dieses Programm eine nach Bundesstaaten aufgeschlüsselte Übersicht aller Daten. Für die Abfrage kommen drei Data Marts zum Einsatz: In „State Summary“ sind die Medicaid-Informationen nach Bundesstaaten zusammengefasst, „Annual Person Summary“ enthält sämtliche Daten nach Patienten geordnet und „Clinical Services“ umfasst Daten zu Medikamenten und Diagnosen. Vor der Befüllung der Data Marts müssen die Rohdaten in einem komplexen Prozess formatiert und zusammengefasst werden.

„Wir haben uns für die Data Marts entschieden, da wir das Data Warehouse nicht direkt ansprechen dürfen“, erklärt Gamerman. „Es handelt sich dabei um eine riesige relationale DB2-Datenbank auf dem Mainframe, und diese Struktur ist nicht für Abfragen ausgelegt und verfügt daher über keinen Index. Doch auch die schiere Datenfülle würde das Warehouse bei Abfragen in die Knie zwingen. Stattdessen extrahieren wir Flatfiles aus den zugrunde liegenden DB2-Tabellen, wobei jeder Flatfile einer Tabelle entspricht, und wandeln sie von EBCDIC in ASCII um. Für die Zusammenfassung für einen Bundesstaat haben wir sechs große Dateien, die wir anhand von Informationen zur Abrechnung und zur Berechtigung der Begünstigten verkleinern und dann mithilfe komplexer Geschäftsregeln zusammenfügen müssen.“ Unterstützung bei der Entwicklung der Anwendung für die notwendige Datenverarbeitung holte sich Gamerman bei den Beratern von ViPS, einem führenden Anbieter von Datenmanagement- und -analyiselösungen für das Gesundheitswesen.

„Wir wandten uns an Syncsort, da wir nicht über genug Leistungskapazität für diese Datenmengen verfügten.“

– Bruce Crawford, Lead Systems Engineer bei ViPS



BRANCHE

- Gesundheitsbehörde
- Hauptsitz in Baltimore, USA

HERAUSFORDERUNGEN

- Analyse von Abrechnungsdaten aus jedem US-Bundesstaat
- Abfragen des relationalen DB2-Data-Warehouse nicht möglich
- Verarbeitung von ca. 100 bis 200 GB aktueller Abrechnungsdaten pro Patient und Quartal
- Beschleunigung der Verarbeitung von Joins und Aggregationen

DIE LÖSUNG

- Einbindung von Syncsort DMExpress für eine problemlose Beschleunigung der Informatica-Umgebung von CMS
- Steigerung der Anwendungsleistung und Beschleunigung der Analyse sehr großer Mengen von Abrechnungsdaten durch erweiterte Aggregationsfunktionen (Sort, Merge und Join) von DMExpress

GESCHÄFTLICHER NUTZEN

- Senkung der Wartungs- und Gesamtbetriebskosten für die Datenintegrationsumgebung von CMS durch eine effizientere Nutzung der CPU-Zeit sowie des Festplatten- und Arbeitsspeichers
- Nutzung der vorhandenen Informatica-Umgebung, um eine höhere Rendite zu erzielen
- Bereitstellung von Business-Intelligence-Berichten über Trendanalysen, Ausgaben, Medikamente und Diagnosen im Medicaid Statistical Information System (MSIS) und Healthcare Information System (HCIS)

Erhebliche Leistungssteigerung

„Das MSIS verarbeitet ein enormes Datenvolumen und die CMS mussten verschiedene Berechnungen und Distinct Counts für ca. 1 500 Permutationen der Schlüssel durchführen“, berichtet Bruce Crawford, Lead Systems Engineer bei ViPS. „Die CMS legten Wert darauf, Informatica für die Business-Logik der Anwendungen zu verwenden. Wir wandten uns an Syncsort, da wir nicht über genug Leistungskapazität für diese Datenmengen verfügten. Wir wussten, dass wir uns darauf verlassen konnten, dass Syncsort die Aufgabe schnell erledigen würde. DMExpress arbeitet ohne Unterbrechung und wir profitieren zusätzlich von den erweiterten Merge- und Join-Funktionen.“

Die Anwendung verwendet nun eine Kombination aus Syncsort- und Informatica-Funktionen, die um spezielle C-Programme ergänzt wurden. Damit kann sie ca. 100 bis 200 GB aktueller Abrechnungsdaten pro Patient und Quartal verarbeiten. Die Flatfiles für jeden US-Bundesstaat und Puerto Rico werden einzeln verarbeitet. Die Dateien werden mit Syncsort sortiert und gebündelt, Informatica wird zur Anwendung der Geschäftsregeln und für die Aggregation verwendet. „Die Leistung hat sich erheblich verbessert, seit wir vor der Durchführung der komplexen Aggregationen Syncsort verwenden. Wir haben die Funktionen von Syncsort und Informatica eng miteinander verknüpft. Die Anwendung von Geschäftsregeln kann recht komplex sein und häufig müssen wir gleichzeitig eine Suche oder andere Aufgaben durchführen. Außerdem müssen wir die Dimensionstabellen berücksichtigen, daher erstellen wir Ersatzschlüssel. Bei den großen Datenvolumen bietet Syncsort den einzigen effizienten Weg zur Zusammenführung von Dateien“, fügt Gamerman hinzu.

Für das MSIS werden die Daten vierteljährlich auf einem Sun E10000 Server mit Solaris verarbeitet. Die Aggregationen

werden dann in die drei Data Marts geladen und zur Befüllung von IBM Cognos-Cubes für Abfragen verwendet. Diese Daten können anschließend für die Erstellung unterschiedlichster Berichte für jeden Bundesstaat – z. B. über die Ausgaben von Medicaid und die Begünstigten nach Leistung und Altersgruppe oder die Ausgaben für ausgefallene Programme nach Wartungsstatus und Berechtigung – genutzt werden.

Nutzung sämtlicher Funktionen

CMS entschied sich außerdem dafür, Syncsort in das Healthcare Information System (HCIS) einzubinden, da dieses Projekt ähnliche Anforderungen hinsichtlich Verarbeitungsaufkommen und Zusammenfassung von Daten hat. In HCIS werden alle Abrechnungen von Medicare verarbeitet und analysiert. Wie bei MSIS werden die Abrechnungsdaten von HCIS vom Mainframe auf den UNIX-Server verschoben und durchlaufen dort eine Reihe von Verarbeitungsschritten in Syncsort und Informatica. „Wir komprimieren die Daten, wobei wir ca. 200 Permutationen durchführen, und laden sie anschließend in die Oracle-Tabellen“, erklärt Crawford. „Wir führen den gesamten Prozess mit Syncsort und Informatica durch und nutzen sämtliche Aggregationsfunktionen von Syncsort für Join, Merge, Sort usw.“

„Wir haben uns hauptsächlich deswegen für den Einsatz von Syncsort in Verbindung mit MSIS und HCIS entschieden, weil wir ein leistungsstarkes Tool benötigten, das das enorme Datenvolumen bewältigen kann“, fasst Gamerman zusammen. „Syncsort war in der Lage, die Fülle an Daten und Abrechnungen zu verarbeiten, die wir zusammenführen mussten. Nicht zuletzt nutzt es den Festplatten- und Arbeitsspeicher sehr effizient.“

Über Syncsort

Syncsort ist ein weltweit tätiger Softwarehersteller, der Unternehmen eine neue Sicht auf den Business Value ihrer Daten ermöglicht. Syncsort bietet für die Datenintegration und -sicherung hochleistungsfähige Lösungen, die sich dank der einfachen Bedienung innerhalb kürzester Zeit amortisieren. In mehr als 15 000 Implementierungen hat Syncsort eine Basis für neue Entscheidungsmöglichkeiten gelegt und weltweit Tausenden von Kunden zu einem besseren Geschäftsergebnis verholfen. Syncsort Incorporated, mit Hauptsitz in 50 Tice Boulevard, Woodcliff Lake, NJ 07677, USA. www.syncsort.com

syncsort RETHINK THE ECONOMICS OF DATA

Syncsort GmbH, Calor-Emag-Str. 3,
40878 Ratingen, Tel.: 02102 9546-0
www.syncsort.de