

NC STATE UNIVERSITY

Problem Overview

Every month, Thomson Reuters Healthcare performs a rebuild of their medical claims database. This rebuild is both computationally complex and time consuming.

Because of this long rebuild time, Thomson Reuters has issues meeting certain requirements with a few of their Service License Agreements they hold with other companies.

Our job is to fix the rebuild so that it takes much less time to compute.

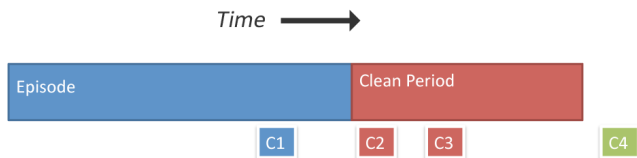
Overview Concepts

Terms

Claim A single insurance claim that has been filed inside the Thomson Reuters system.

Episode A group of claims for one person that, as a group, define a single medical incident.

Clean Period The length of time after an episode such that after that point, no new claims may be added to the episode.



Episodes can be comprised of 1 or more claims, and can also be considered either acute (such as breaking a leg) or chronic (diabetes)

Tools

MEGEE An application that takes as its input a set of claims, and from those claims, produces a set of episodes that it constructs using internal rules.

MultiLoad An application that writes large files containing SQL statements to a database.

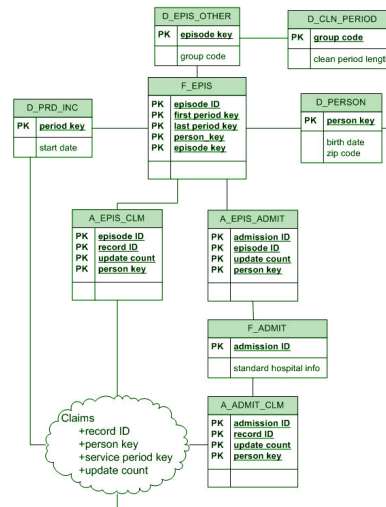
Old Method

As it is currently implemented, the rebuild process simply takes all claims from the database and constructs the set of episodes. Because of this, rebuilds take longer and longer to compute with each iteration.

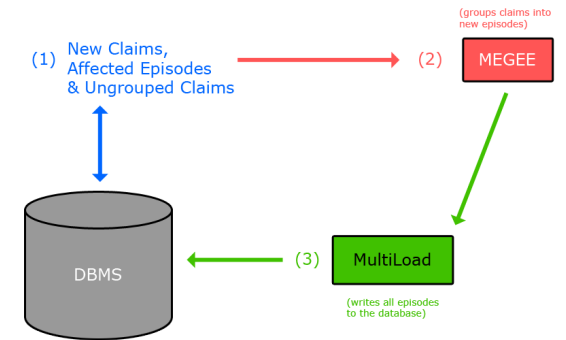
This rebuild requires the following:

- 1) Two powerful Sun v890 build servers with 8 dual-core CPUs and 64 GB of RAM each.
- 2) 54 hours of uninterrupted computing time each server.

Database Schema

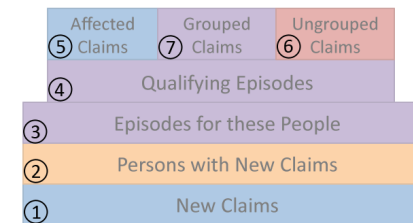


New Method

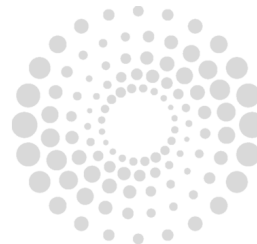


Pull only selected claims from the database, these claims are then passed to MEGEE to be grouped into episodes. The new set of episodes are then written back to the database.

SQL Design



This diagram shows how we have layered our SQL queries on top of one another. We begin by pulling the new claims from the database, and using a date-related heuristic, pull only the claims that could possibly be affected.



THOMSON REUTERS