



# How the right database helps Caresyntax improve their operations

*A dbInsight case study for Amazon Web Services*

## Executive Summary

### Trigger

The changes to the global economy that are driving digital business are also driving demand for a new generation of scalable, cloud-ready, modern applications. Compared to traditional, internally-facing enterprise transaction applications, modern applications respond to a wide range of scenarios and data types that in turn drive varying requirements for performance, scale, and geographic reach. There is no single, silver bullet data management recipe for responding to use cases ranging from supporting millions of concurrent sessions for online gaming, to identifying patterns of fraud in financial transactions, and delivering consistent high performance for scenarios with unpredictable demand spikes. With a variety of options, ranging from traditional SQL relational databases to NoSQL key-value stores, document database, graph databases, and others, how can organizations make the right choices?

### Our Take

The use case drives the choice of database. While there is no single formula dictating whether an enterprise adopts one or more databases to address their business use case or scenario, they should depart from preset notions and dissect the problem into its constituent parts, assessing the need for requirements ranging from ACID transaction consistency, referential integrity, latency, availability, and other factors. In some instances, a mix of database systems may prove the best solution, for example with web and mobile commerce applications that may require a mix of simple lookups, search capabilities, and in some cases, leaderboards for showing hot selling items. In other cases, a single solution may address the challenge where the problem is narrow and well-defined, such as sorting through the complex interrelationships of entertainment content.

The first step in choosing the data solution always starts with the business goal and then evaluating the tasks that are necessary to address it. In many cases, the tasks and the requirements for data may be varied. For example, in e-commerce, where the operations often involve a mix of stable and volatile data, there is need for simple lookups and search, and the ability to quickly sort through and retrieve frequently accessed data. With numerous options for databases available, enterprises need not make compromises on choosing the best data solution for the job.

## Caresyntax: Finding the right mix of operational and analytic data stores

### Helping healthcare providers become more efficient

Choosing data solutions requires understanding the data *and* how it will be consumed. Caresyntax delivers solutions for healthcare providers for optimizing their operating rooms. For them, the scalability of the cloud not only impacts the capacity to handle demand spikes, but also the ability to deliver far more granular analysis of the factors that impact both the operational efficiency and the outcomes in its business. The cloud's scale offers another key benefit: elimination of process bottlenecks for data systems based on legacy on-premises architectures.

In 2019, Caresyntax completed a major acquisition that filled a key gap in its business: adding operational analytics for tracking the efficiency by which hospitals book their operating rooms. That complemented its existing business that had been focused on tracking and providing analysis of clinical outcomes. For most hospitals, surgery accounts for the lion's share of billing, equipment, and utilization of specialists. Successful hospitals efficiently schedule and utilize their operating theaters and consistently deliver successful patient outcomes. While analysis of operational efficiency and clinical outcomes have traditionally been highly siloed, there are huge potential advantages if hospitals can understand the impacts of one on the other, and from that generate lessons and best practices.

### Finding the right database for the job

Caresyntax's operational analytics solution for surgical facilities uses several AWS database services including Amazon Aurora MySQL, for transaction processing and light analytic queries; Amazon Redshift, for running analytic queries demanding more scale and granularity; open source Redis for caching "hot" frequently queried data; and Amazon DynamoDB for managing the availability of reporting parameters.

Originally, the key to choosing the transaction database was driven by skills and scaling requirements. The team developed the implementation on a self-managed MySQL database, but as the service grew successful, it required a more scalable system to handle transaction and query traffic. Caresyntax's solution stores basic data on operating room utilization, and as such, fits the profile for a typical transaction database use case: high volumes of read/write/update/delete queries on relatively limited columns or fields. While the range of data is limited, the size of the transaction store is significant, numbering in the millions of rows.

With Aurora, AWS offers a service that provided full compatibility with the team's original MySQL implementation; the team could move to Amazon Aurora with MySQL compatibility without changing any lines of code or the underlying schema. They gained a database service

designed for the cloud that leverages a smart storage tier for accelerating write commits, automatic three-way replication to speed up reads, and much higher availability than possible with the basic self-managed open source implementation. That is critical for a database with high transaction volumes and row counts. While Aurora is used for transactions, it is also used for lightweight queries such as capacity utilization of the operating room, which involves lookups of relatively few columns that are readily handled, even with transaction stores of up to millions of rows.

The data also drove the requirement for a separate analytics solution. The factors impacting operating room utilization venture beyond scheduling. It also requires tracking inventory along with the events involved with conducting surgery. That extends from the point where the patient is rolled into the administration on anesthesia, the conduct of the operation, and the steps required to transfer the patient to recovery. With so many data elements, analytic queries are more optimally handled in a columnar database that also has the capability to perform federated query to analyze occasionally-used “cooler” data. Federated query provides a more cost- and time-efficient alternative to constantly performing ETL and persisting cooler data in the data warehouse.

Caresyntax uses Amazon Redshift, a data warehouse service, for its ability to store massive amounts of data that includes far more entities, or columns compared to the transaction system. Besides operating room utilization, Redshift also is used for storing data encompassing the details of the equipment, implanted devices, and surgical steps or events. Consequently, while Aurora can be used for analyzing the level of overall utilization of the operating room, Redshift can drill down with far more granularity. For instance, it can help compare performance by individual surgeons, such as why the same (or different) surgeon took differing amounts of time to perform the same procedure, and in turn, analyze the sequence of steps to help formulate best practices. This level of detail would not have been possible with a traditional on-premises data warehouse that could not efficiently autoscale to manage the data and provide analytic access to it.

For now, Caresyntax uses Redshift strictly for analyzing data in place, but Redshift also provides future-proofing for a company that has data stored in Amazon S3 storage for its clinical care analytic solutions; in the future, Redshift Spectrum might be used to query metadata from video stored in S3 for Caresyntax’s clinical analytics application, Qvident, that records video from surgery so that individual surgical steps can be analyzed.

Complementing Aurora and Redshift, Caresyntax also runs Redis in Amazon EC2 as a caching layer for “hot” data that is frequently queried. Additionally, Amazon DynamoDB is used for configuring reporting, an important capability that enables Caresyntax to support the varying reporting requirements across different healthcare systems – requirements that are continually evolving thanks to changing regulatory regimes. By using a NoSQL data store, Caresyntax can store requirements and not be burdened with the overhead of having to change the schema each time a report requirement changes.

## Takeaways

Depending on the use case, there may be widely diverging requirements for managing and processing data. The data in a product lookup table that is stable will have different requirements from the data representing public health exposure from interpersonal contacts that are constantly changing. But the decision is not purely a matter of architecture. The same body of data that might have a relational structure may require different modes of access depending on whether the data is stable, such as a fact table, or volatile, such as tracking the latest product, gaming, or media content choices by the consumer. Can the requirements for reading and writing data be broken up into separate steps, such as member lookup and next-best action or recommendation, and therefore be addressed by database solutions that are designed for the purpose? Furthermore, does the nature of the use case translate to a need for extreme autoscaling and/or the need to read and/or write data with processes that are localized or highly distributed? And what is more important to the business case: returning data that is absolutely consistent, or ensuring that the service be highly available?

When choosing the database(s), look at the business use case first. With a multitude of database options out there, the choice does not have to be whittled down to an either-or.

## Author

Tony Baer, Principal, dbInsight

[tony@dbinsight.io](mailto:tony@dbinsight.io)

Twitter @TonyBaer

## About dbInsight

dbInsight LLC® provides an independent view on the database and analytics technology ecosystem. dbInsight publishes independent research, and from our research, distills insights to help data and analytics technology providers understand their competitive positioning and sharpen their message.

Tony Baer, the founder and principal of dbInsight, is a recognized industry expert on data-driven transformation. *Analytics* named him as one of its Top 100 influencers for [data](#) and [cloud](#) in 2019 and 2020. *Analytics Insight* named him one of the [2019 Top 100 Artificial Intelligence and Big Data Influencers](#). His combined expertise in both legacy database technologies and emerging cloud and analytics technologies shapes how technology providers go to market in an industry undergoing significant transformation. His regular ZDnet “*Big on Data*” posts are read 25,000 – 30,000 times monthly.

dbInsight® is a registered trademark of dbInsight LLC.