SISENSE

ANSWERED:
YOUR MOST BURNING QUESTIONS ABOUT ANALYTICS WITH LIVE DATA FROM REDSHIFT DATABASES

INTRODUCTION:
IS YOUR DATA ALWAYS ONE STEP BEHIND?

MOVING DATA FROM ONE PLACE TO ANOTHER PUTS A LOT OF PRESSURE ON YOUR BI SYSTEM. IN ORDER TO ANALYZE YOUR DATA, TYPICALLY YOU NEED TO EXTRACT IT FROM THE ORIGINAL DATABASE OR SYSTEM AND PUT IT IN A DIFFERENT ONE, WHERE YOU CAN MASH IT UP, MANIPULATE IT, AND GENERALLY WORK WITH IT.

This batch or micro-batch approach to BI is a convenient way of doing things, and it's totally fine for the data analysis needs of many organizations, especially if you tend to run the same queries at regular intervals, or know in advance which data sources you'll be using. In some cases, though, it's less than ideal. If you need immediate or near-immediate results, the added loading time that comes with importing the latest data all over again can be frustrating.

Thanks to recent developments in the data landscape, companies that fall into this category no longer need to put up with a lag. Fast-performing databases like Redshift are used to improve connections to live data and strip out delays from BI processes. Meanwhile, BI platforms are getting more and more flexible, creating new ways to capitalize on live databases and allowing you to analyze data as fast as it can be updated.

The best of both worlds would be to use a hybrid solution that feeds live data and updates into your dashboards directly from external databases while combining dashboards and analysis from historical data in the system.

Technological evolutions like these mean you can embrace lightning-speed, live connection access to data in a Redshift environment, making ad-hoc and complex queries easier and more accurate than ever.

SISENSE

WHY YOU NEED A
LIVE CONNECTION

BEFORE WE GET INTO THAT, LET'S REVIEW WHY YOU NEED A LIVE CONNECTION TO YOUR DATABASE OR DATA WAREHOUSE IN THE FIRST PLACE.

In sectors like sales and marketing, trends change quickly. Depending on your area of business, we could be talking about weeks, days or even hours. If one product is flying off the shelves, you need to know so that you can place orders or shift stock into stores. If a campaign is tanking, the sooner you know, the faster you can switch tactics. Any delay means lost opportunities, increased risk, lagging behind the competition and ultimately, taking a hit to your bottom line.

In fields like fraud detection and cybersecurity, you need to be alerted instantly if something looks suspicious. It's not much good if your system is telling you that, yes, there **_was_** a string of unusual transactions on a card, or hackers managed to find a weak spot.

BY THAT STAGE, YOU'RE ALREADY CLEANING UP THE MESS.

QUICK AND EASY
DECISION-MAKING

Of course, you also need a way to run queries on all that data quickly and efficiently. Increasing your query performance and your access to the full scope of data is fundamental to helping you make better, data-backed decisions.

By giving business users a way to access a high-performance database directly and immediately, they get reliable insights faster. They can analyze that data in moments, reducing uncertainty and improving swift, accurate decision-making.

SISENSE

HOW REDSHIFT LIVE
CAN HELP

THIS NEED FOR RAPID ACCESS HAS MADE AMAZON REDSHIFT VERY POPULAR, AND WITH GOOD REASON.

As a cloud-based database, Redshift gives you scalability and the safety net of automatic backups while keeping maintenance costs low. It's also very fast, outperforming traditional databases and at times even leading data lake technologies like Hadoop.

The beauty of Redshift is that it allows any user to set up a database with millions or even billions of rows, at a far lower price point than a traditional database - and with no hardware (for the data storage at least, running queries is another matter - more on that in a moment). This makes Redshift an extremely useful source of data for live Business Intelligence and Business Analytics.
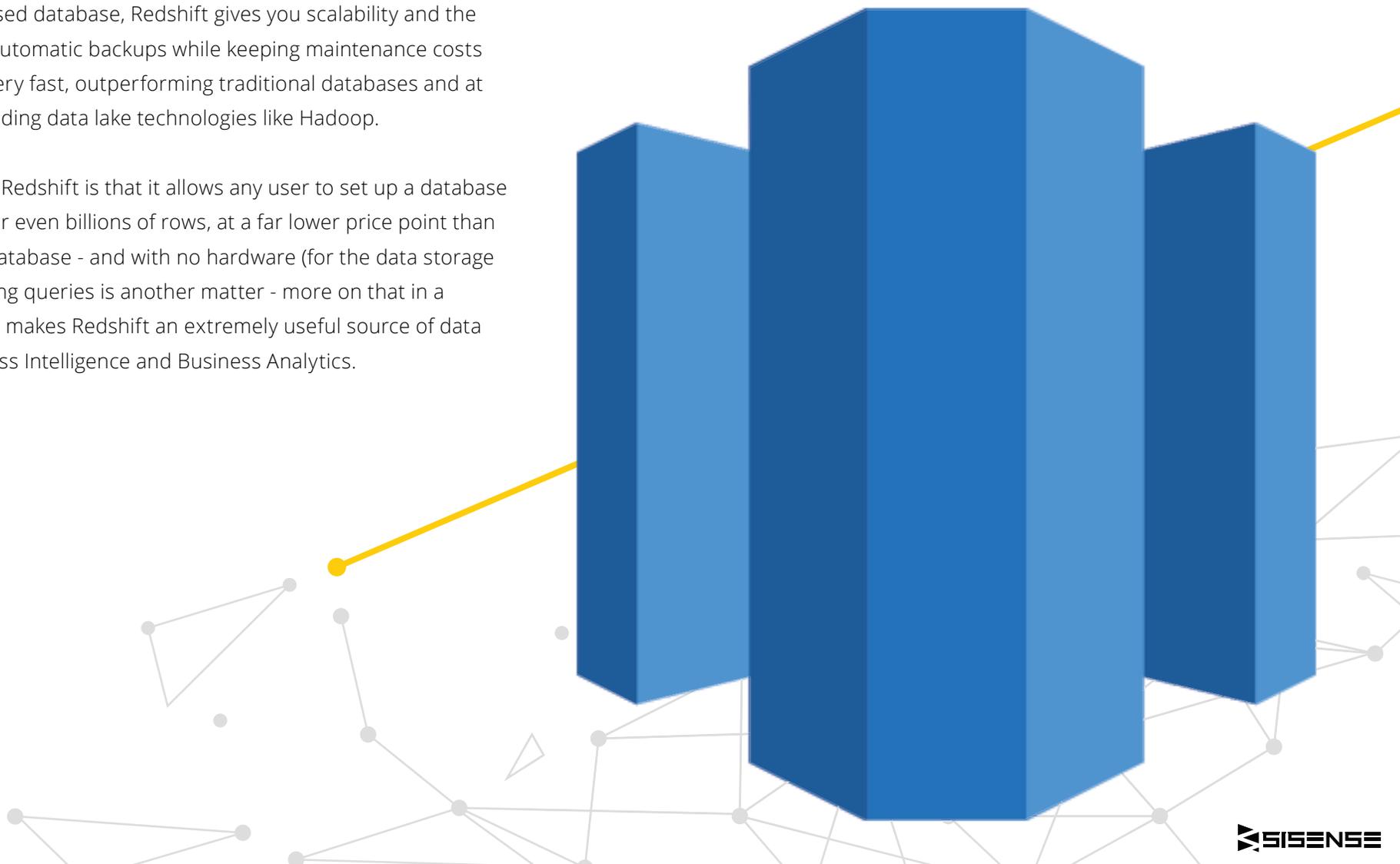
COMBINING LIVE AND HISTORICAL DATA
TO ENRICH ANALYTICS

With a live connection to data in a Redshift database, live data runs through the system and is cached, not stored. It's also replaced immediately every time there is new live data. This means you can run 100% up-to-date analysis in real time, in order to make well-supported on-the-spot decisions - even from constantly changing data.

This builds on the capacity of our live data models in a few important ways. With Sisense, you can query any data that's already in your data warehouse or other directly-connected data sources, but you may have to wait a little longer if you want to run ad-hoc or complex analysis, if the data hasn't yet been loaded into the data warehouse, or if your data warehouse infrastructure can't quite keep up with how fast you need to create a complex analysis model or integrate data.

NOW LET'S IMAGINE YOU WANT TO IDENTIFY BROADER TRENDS IN YOUR DATA OR DEVELOP PREDICTIVE MODELS.

For this, you need to work with static and live data together... but you also need to sound out which data from which sources will tell you what you want to know before you go to the trouble and expense of moving the data from place to place, or storing it again. Hybrid analytics gives you a great way to model this, playing with data from different streams to work out how you'll use it before you commit. You can play with static and live data to see how different systems, departments, and groups interact with Redshift data.

SISENSE

WHERE DOES
SISENSE COME INTO IT?

THIS NEED FOR HYBRID DATA SOURCE FUNCTIONALITY LED THE TEAM AT SISENSE TO DEVELOP LIVE CONNECTION CAPABILITIES. HERE'S HOW THAT WORKS.

If you're not familiar with the Sisense ElastiCubes, this is Sisense's unique, high-performance analytics database engine with super-fast data stores that are specifically designed to withstand extensive querying typically required by business intelligence applications.
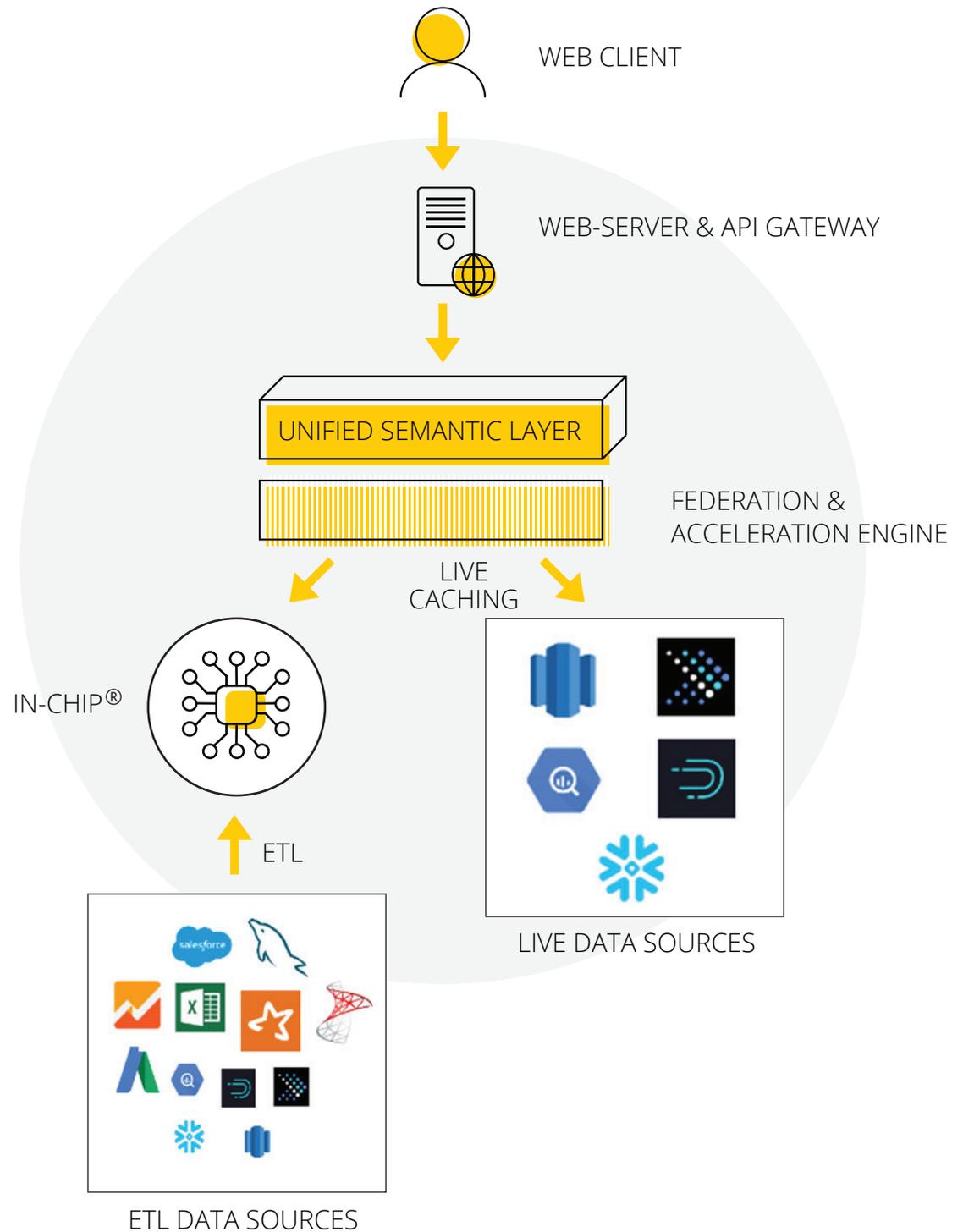
ElastiCubes allow you to bring in data from multiple sources, and then merge, manipulate and query the data as if it was one consolidated dataset. These unique cubes are what makes it possible to do Business Intelligence and Business Analytics even without a data warehouse a sort of a built-in data warehouse for analytics and BI. For Live Data Models, you go a step further and bypass the ElastiCubes and link directly to Redshift (as well as other types of databases, SQL Servers, PostgreSQL and so on). With ElastiCubes, you work with close to real-time data. Now it really is immediate.

What's more, live connections are managed through a graphical UI that requires very little technical understanding and is easy to navigate for business users. This reduces business users' reliance on IT and allows them to manage live connections themselves. They can decide which types of data they want to work with and what kind of queries they want to run without adding to IT's workload and can get the answers they need in the blink of an eye.

"But how does it work?" you may be wondering. Well, it's wonderfully simple, really. We host and manage our live connection to Redshift in the Cloud. From the viewer side, you establish a direct connection with your Redshift data warehouse and select the kind of data you want to look at. Since you're working with live data, this is automatically refreshed the moment it is updated, and everyone in your team continues to use the same data source. This is your single version of the truth.

Here's where it gets really interesting, though. When you also want to use static data, you can incorporate this using ElastiCubes. These link to your data warehouse(s), or wherever else you store data, allowing you to bring this in alongside your live data from Redshift. As a result, you can work with current and historical data together without taking a hit on performance.

HYBRID DATA SOURCE FUNCTIONALITY WITH **LIVE CONNECTOR TO REDSHIFT**

WEB CLIENT

WEB-SERVER & API GATEWAY

UNIFIED SEMANTIC LAYER

FEDERATION & ACCELERATION ENGINE

LIVE CACHING

IN-CHIP®

ETL

LIVE DATA SOURCES

ETL DATA SOURCES

6

OPTIMIZING REDSHIFT
ON SISENSE DATA HUB

COMBINING SISENSE AND REDSHIFT CREATES A FAR MORE FLEXIBLE APPROACH.

Using both live and static data sources, you can customize your logic and create custom tables and columns not tied to the database, opening up a whole new world of options for analysis and insights.

It helps you get the most out of Redshift, which isn't a high-concurrency database and so struggles with multiple users/sessions at the same time. Feeding the data into Sisense means you optimize how Redshift data is accessed and distributed, making it much more efficient. You can also use this to help you develop your agile analytics architecture by upgrading Redshift, developing and testing it on Live, then downgrading it and moving it to ElastiCube for analysis (or vice versa).

Plus, it saves you money. Today, many businesses run their live BI on Redshift, but with Redshift charging per query, this can get very expensive, very quickly. Bringing the data into Sisense means you can downgrade the Redshift data and run much of the intensive work through the In-Chip ® memory, columnar database system, dramatically reducing the pressure on your machines and in turn, your hardware costs.

For example, let's say you currently use a DC3 node to power your queries. By using a live connection, you can load data into the ElastiCube Data Model for analysis instead, spreading your computing workload and allowing you to rely on a much cheaper DC2 node.

And since Sisense does not rely on writing SQL statements against Redshift, any employee can use the visual, menu-based user interface to create Data Models to Redshift or ElastiCube Data Models without the need for any third-party tools.

SISENSE

PRIMARY USE CASES

SO: WHAT IS A LIVE CONNECTION GOOD FOR? BROADLY SPEAKING, IT HELPS YOU DO THREE THINGS REALLY WELL.

Faster Results

First, as we've explored in detail, it's a great way to generate faster results - on-time or even real-time, for incredibly accurate and relevant insights.

Cost-Efficiency in Maintaining Data Models

Second, it gives you a time-saving, cost-efficient way to maintain existing data models based on data in your database or data warehouse. Instead of refreshing the whole database every time you add something new, you stick to using this for historical data, and your live connection to Redshift for live data.

Single BI Platform for Existing Data Models & Live Data Models

Third, it streamlines everything. Rather than switching between multiple platforms to run piecemeal analyses that you'll later have to stitch together yourself, you get to run everything through one single BI platform. Even more importantly, you get that operational simplicity without having to sacrifice on detail or scope - the system gives you both depth and breadth of analysis.

WHAT CAN'T IT DO?

IS A LIVE CONNECTION THE MAGIC BULLET?

Well, not in every situation. There are still some things that the system can't pull off perfectly... at least not yet.

Firstly, it's hard to use a live connection when you're running an ad-hoc analysis as this requires too much complex caching to do on the fly. It's much better at handling preset queries.

Next, while you can get right down to day-by-day data, the system doesn't support hour and minute granularity, and you can't create pulse alerts. Bear with us though, as we're working on these.

FINAL THOUGHTS:
GETTING UP TO SPEED

Adopting hybrid analytics with Redshift and Sisense's live connection is an ideal way to optimize your existing data and BI investments while making the whole process easier for IT and business users.

It becomes cheaper and less time-consuming to maintain your data architecture, you can capitalize on existing data infrastructure rather than replacing or upgrading it all, and it's far easier to use, meaning that business users don't need to seek technical assistance every time they want to query live data connections.

Most importantly, it takes the accuracy and timeliness of your analysis to new levels. Query performance is improved, time to insight reduced - and your team provides insights that they can put into practice fast, with greater confidence, spurring on business growth and efficiency all along the chain.

Intrigued? Try it out on your own data! Take a test drive and see how Sisense can connect directly to your Redshift data by scheduling a Custom POC.