In this Breaking Analysis we dig into the cloud database market and take a closer look at how Snowflake competes with Amazon Redshift, Google BigQuery and Microsoft Synapse, among others. New Cloud Workloads As we’ve reported, there is a new class of workloads emerging in the cloud. Early cloud was all about infrastructure-as-a-service (IaaS). [...]

Breaking Analysis: How Snowflake Competes with Cloud Native Database Giants  by, David Vellante

June 6th, 2020
In this Breaking Analysis we dig into the cloud database market and take a closer look at how Snowflake competes with Amazon Redshift, Google BigQuery and Microsoft Synapse, among others.

**New Cloud Workloads**

As we’ve reported, there is a new class of workloads emerging in the cloud. Early cloud was all about infrastructure-as-a-service (IaaS). That is, the spinning up of storage, compute and networking resources to support startups, dev/test, SaaS and eventually moving more business workloads into the cloud.

Today’s cloud workloads go beyond infrastructure services and are increasingly diverse. One of the more notable innovations we’re seeing is the practice of leveraging data by infusing AI into applications, simplifying analytics and scaling with the cloud to deliver business insights in near real time. At the center of this mega trend is a new class of data stores and analytic databases. What some call enterprise data warehouses (EDW) – a term that is perhaps outdated for today’s speed of doing business.

In this Breaking Analysis we want to accomplish three things:

1. First we want to cover the basics of the cloud database market space – i.e. what you need to know.
2. Next we will look into the competitive environment and dig into the ETR spending data to see who has the momentum in the market.
3. Finally we’ll close with some thoughts on how the competitive landscape is likely to evolve. We will answer the question – will the cloud giants overwhelm the upstarts and specifically Snowflake? Or will the specialists continue to thrive and if so how?

**Legacy EDW Evolves to Analytic Data Stores**

We are seeing a revolution in the EDW market space, brought on by cloud, data science tooling and modern database technology. EDW has been critical to supporting the reporting and governance requirements for companies, especially supporting the accounting requirements of Sarbanes-Oxley. However historically EDW has failed to deliver on its promises of a 360 degree view of the customer and real time insights. Classic enterprise data warehouses are too cumbersome, complicated and slow to keep pace with the speed of business.
EDW is a $20B market but the analytic database opportunity we think is much larger. Why? Because cloud computing unlocks the ability to rapidly combine multiple data sources, bring data science tooling into the mix, quickly analyze data and deliver near real time insights to the business; or importantly, allow line of business pros to access data in a self-service mode. It’s a new paradigm that uses the notion of DevOps as applied to the data pipeline - think agile data or “DataOps”.

The market for cloud native analytic database is highly competitive. In the early part of last decade we saw Google bring BigQuery to market. But Google was primarily focused on its own ad business and has taken years to make enterprise cloud a priority. Snowflake was founded in 2012 and is a disruptor in the market. Around this time, AWS did a one-time license deal to acquire the intellectual property of the ParAccel MPP database, on which it built Amazon Redshift. In the latter part of the decade, Microsoft threw its hat in the ring with SQL DW which the company evolved into Azure Synapse at its Build conference a few weeks ago. There are other players as well like IBM.

High Stakes Game of Chess

There’s a lot at stake here. The cloud vendors want your data because they understand that is one of the key ingredients of the next decade of innovation. No longer is Moore’s Law the mainspring of growth...rather today it’s data and AI to drive insights at scale with the Cloud.

Here’s the interesting dynamic emerging in the space. Snowflake is the cloud specialist in this field having raised more than $1B in venture capital. And it’s up against the big cloud players who are moving fast and often taking moves from Snowflake and driving customers to their respective platforms. But Snowflake is also a major partner for the cloud suppliers because they help sell infrastructure services.

For example. Snowflake’s largest cloud partner is AWS. Snowflake drives lots of Amazon EC2 sales. Yet AWS has Redshift, which directly competes with Snowflake. Redshift often announces features that Snowflake has popularized. Here’s an example that we reported on at last year’s AWS reinvent. The article below by Tony Baer from ZDNet talks about how AWS RA3 separates compute from storage. Of course, this was a founding architectural principle for Snowflake.
And here’s another example from The Information reporting that Microsoft, another Snowflake cloud partner, is turning up the heat on Snowflake. And you see the highlighted text below where the author talks about Microsoft trying to divert customers to its database.

So you have this weird dynamic. Snowflake doesn’t run on prem. It only runs in the cloud. It runs on AWS, Azure and GCP. The cloud players all want your data to go into their database and they push hard on customers to use captive services. At the same time they need ISVs like Snowflake to run in their clouds.
because it sells infrastructure services, expands customer optionality and evolves the ecosystem.

Should Snowflake perhaps pivot to run on-prem as a way to differentiate from the cloud giants? We asked Frank Slootman, Snowflakes CEO, about the on-prem opportunity earlier this year and his comments below are crystal clear:

Snowflake is 100% cloud native – period... https://t.co/Y5jMyqQpqr @SnowflakeDB #cloud #aws #azure #gcp #analytics #database pic.twitter.com/vrWr0eH0xC

— Dave Vellante (@dvellante) June 6, 2020

These are unequivocal statements by Slootman. The question we want to pose next then is can Snowflake compete given the conventional wisdom that we saw in the media articles that the cloud players are going to hurt Snowflake in this market? And if so how will Snowflake compete?

Customer Spending Data Shows Snowflake Poised to Win

The chart below shows two of our favorite metrics from the ETR data set. Net Score, which is on the Y-Axis – that’s a measure of spending momentum – and Market Share on the X-Axis. Market share is a measure of pervasiveness in the data set, not conventional share. It’s a calculation of mentions of a company divided by total mentions within a sector. Below we show some of the key players in the EDW and cloud native analytic database market.

The following points are noteworthy:

- We show survey data from the April ETR survey, which was taken at the height of the COVID lockdown. The survey captured responses from more than 1,200 CIOs and IT buyers asking about their spending intentions for analytic databases for the companies shown on the chart.
- The higher on the vertical axis, the stronger the spending momentum. You can see Snowflake with a 77%
Net Score leads all players with AWS Redshift very high as well.

- In the box on the lower right of the chart you can see the exact Net Scores for all the vendors and the Shared N. Shared N is the number of citations for that vendor within the survey N of 1269. So you can see the overall sample is large and the vendor mentions are large enough to feel comfortable with some of the conclusions we will make.
- Microsoft has a huge footprint and somewhat skews the data with its very high market share due to its volume. You can see where Google sits with good momentum but not as much presence in the market.
- We’ve added Teradata and Oracle for context – two companies that primarily compete with on-prem offerings.

The bottom line is twofold: 1) The cloud native analytic database market is capturing share of wallet; and 2) Snowflake, as it has for the past several surveys, continues to lead all players with the highest spending velocity.

**Snowflake and Redshift both Strong on AWS**

Let’s look at how Snowflake performs inside of the “Big 3” clouds. We’ll start with AWS.

The chart below shows the customer spending momentum inside of AWS accounts. We cut the total sample to isolate only on those ETR survey respondents running AWS – an N of 672. The bars show the Net Score granularity for Snowflake and Amazon Redshift.

![Snowflake Winning in AWS Accounts](image)

We show that there are 96 shared N responses for Snowflake and 213 for Redshift within the N of 672 AWS accounts. The colors show 2020 spending intentions relative to 2019. Reading left to right: Replacements (bright red), spending less by 6% or more (pinkish), flat spend (gray), increasing spending by more than 6% (forest green) and adding the platform new (lime green).

Net Score is derived by subtracting the reds from the greens. And you can see that Snowflake has more spending momentum in the AWS cloud than Amazon Redshift by a small margin.
Adding the green bars shows that 80% of AWS accounts plan to spend more on Snowflake in 2020 relative to 2019. With 35% of that number coming from customers adding Snowflake as new. For Redshift, 76% of AWS customers plan to spend more in 2020 relative to 2019 with 12% adding new. So both companies show very strong spending velocity with minimal red.

It will be critical to see in the June ETR survey – which is now in the field – if Snowflake is able to hold on to these new accounts.

**How is Snowflake doing inside of Azure?**

Let’s take a look at that data from the ETR survey to answer this question.

So we’re showing above the same view of the data here except we isolate on 677 Azure accounts within the survey. We show Snowflake and Microsoft cuts for analytic databases with 83 and 393 shared N responses respectively. So enough we feel to draw some conclusions.

Note the Net Scores. Snowflake again winning with 78% versus 51% for Microsoft. Once again you see massive new adds at 41% for Snowflake, whereas Microsoft’s Net Score is being powered by growth from existing customers. And again – very little red for both companies.

**How is Snowflake doing in GCP accounts?**

Let’s dig in to that data from the ETR surveys.
Here’s the same view of the data in the chart above. The difference is now we isolate on 298 GCP accounts running Snowflake and Google analytic databases. The Snowflake shared N at 49 is smaller than on the other clouds because the company just announced support for GCP about a year ago. But still large enough to draw conclusions from the data. And you can see Google’s shared N at 147.

Once again, Snowflake is winning by a meaningful margin as measured by Net Score or spending momentum with 77.6% versus Google at 54%. Adding the two green bars, again we see that 80% of Snowflake customers running GCP expect to increase spending with with Snowflake in 2020.

Both Google and Snowflake show very little red- a positive sign.

The bottom line is our data shows that Snowflake has greater spending momentum than the captive cloud provider in all three of the big U.S.-based clouds.

**Can Snowflake Hold Serve and Continue to Grow?**

As we said, this is a very competitive market. We have reported how Snowflake is taking share from some of the legacy on premises data warehouse players like Teradata and IBM. And from what our data suggests, Oracle too. We have reported how IBM is stretched thin on its research and development budget. Oracle is more targeted toward database and can direct more of its free cash to database than IBM, but Amazon, Microsoft and Google don’t have free cash flow problems.

This is a challenge for Snowflake in our view. The big cloud players will invest and continue to try and keep pace with Snowflake. Below is an example. It’s a partial list of recent innovations in this space by Snowflake and AWS. We show here a set of features that Snowflake has launched in 2020 and AWS since re:Invent last year.
Many of these features will resonate with database prose (e.g. materialized views, etc.) and have been around for a long time. Cloud native data stores must continue to add critical features that mature on-prem stacks have had for years - especially important are governance and security features. But the point is the new leaders are adding these features in cloud native form.

And we know that AWS is no slouch at adding features. Amazon spends 2X more on research and development than Snowflake is worth as a company.

So why do we like Snowflakes chances?

There are several reasons we think Snowflake can continue to lead. First every dime Snowflake spends on engineering, go to market and its ecosystem, goes into making its database better for customers.

We asked Frank Slootman in the middle of the lockdown how he was allocating precious capital during the pandemic. Below is his response which underscores this point:

Slootman hires in engineering with no reservations because it's the future
https://t.co/BQPGScnnK2 @SnowflakeDB #redshift #Azure #GCP #googlecloud #AWS #database #analytics #research #development #Focus pic.twitter.com/4m7Cc1wPWa

— Dave Vellante (@dvellante) June 6, 2020

But this is only part of the story.

**Building a Data Fabric Across Clouds**

As many of you know we’ve been skeptical of multi-cloud up until recently. We’ve said multi-cloud is a symptom of multi-vendor and largely vendor marketing to date.

That’s beginning to change. We see multi-cloud as increasingly viable and important to organizations,
especially as it relates to data, data locality and global scale.

First we want to reiterate that new workloads are emerging in the cloud. Real-time AI, insight extraction and AI inferencing is going to be a competitive differentiator. The new innovation cocktail stems from machine intelligence applied to data with data science tooling, simplified interfaces that enable scaling with the cloud.

As such we see cross cloud exploitation as a differentiator for Snowflake and others that build high quality, cloud native capabilities for multiple clouds.

What does this mean for Snowflake?

**Differentiation with a X-Cloud Data Fabric**

- New workloads - Data + AI + Cloud - unlock data insights
- RT AI inference
- Multi-cloud will emerge as a differentiator
  - Require lots of innovation
- Cloud native will be a critical capability across the stack
- Prediction - winner will not have an IaaS agenda

---

**Building capabilities natively for the cloud - versus putting a wrapper around your stack and making it run in the cloud - is a critical differentiator.**

**Why Cloud Native?**

Because cloud native means taking advantage of the primitive capabilities, features and APIs within the respective clouds to create the highest performance, lowest latency, most efficient services possible. And it delivers the most secure experience for customers. The best experience will be enabled by natively building in the cloud and its why Slootman is dogmatic on this issue.

Multi-cloud is a differentiator for Snowflake. Data lives everywhere and you want to keep data where it lives. On AWS, Azure or whatever cloud is holding that data. If the answer to your query requires tapping data that lives in multiple clouds across a data network and the app needs fast answers then you must have low latency access to that data.

Snowflake’s game in our opinion, is to automate its portion of the data flow by abstracting complexity related to data location/latencies, metadata, bandwidth concerns, time to query, time to answer, etc.; and optimizing its portion of the stack to get insights irrespective of data location.
A differentiating formula is to not only be the best analytic database but be **cloud agnostic**. AWS, for example, has a cloud agenda. As do Azure and GCP. Their best answer to multi-cloud is put everything on our cloud. Sure they’ll have offerings across cloud but Snowflake will make it a top priority and must be the best at it. Cloud providers will pursue multi-cloud only after they’ve explored captive options. It’s a nuanced dynamic but one that we’ve seen in the market for decades.

Companies without a cloud platform agenda will have a strong argument and currently we think, in this market, Snowflake has the most compelling position in the market.

**Let’s Wrap**

The ETR spending data shown here confirms the anecdotal information we get from customers, theCUBE network in Silicon Valley and the general sentiment about Snowflake in the market. Having data back up (or refute) the conventional wisdom gives us greater confidence to make conclusions and we feel in this case that if Snowflake can continue to execute it will steadily march toward IPO and thrive as a public company.

Remember these episodes are all available as **podcasts wherever you listen**. - just search **Breaking Analysis Podcast** and please subscribe to the series. Check out ETR’s Web site. We also publish a full report every week here and on SiliconANGLE.

Ways to get in touch: Email: david.vellante@siliconangle.com | DM @dvellante on Twitter | Comment on our LinkedIn posts.

Also, check out this **ETR Tutorial we created**, which explains the spending methodology in more detail.

Thanks for all the great feedback we get on these segments – we appreciate you being part of our community.

Watch this week’s full video analysis:
David Vellante is co-CEO of SiliconANGLE Media, as well as co-founder and Chief Analyst at The Wikibon Project, the world’s leading open source technology research community. Dave is a long-time tech industry analyst, entrepreneur, writer and speaker. As co-host of theCUBE – “The ESPN of Tech,” Vellante has interviewed over 5,000 experts since 2010. He is also a co-founder of CrowdChat, an angel funded startup based in Palo Alto using big data techniques to extract business value from social data. Prior to these exploits, Dave founded a CIO consultancy and spent a decade growing and managing IDC’s largest business unit. He lives in Massachusetts with his wife and four children where he is active in town activities including serving as the president of his town’s local “Kiddie Sports” association. Dave holds a B.S. in Applied Mathematics from Union College.