

REPORT REPRINT

This Impact Report was published by 451 Research as part of our syndicated Market Insight subscription service and subsequently licensed for commercial use by AtScale.

AtScale expands purview, recasts as a universal semantic layer for BI tools

KRISHNA ROY

07 JUN 2017

The startup has spread its wings beyond an initial focus on providing native analytics for Hadoop and is now casting itself as a purveyor of one semantic layer for any BI tool.

THIS REPORT, LICENSED EXCLUSIVELY TO ATSCALE, DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, SHALL BE OWNED IN ITS ENTIRETY BY 451 RESEARCH, LLC. THIS REPORT IS SOLELY INTENDED FOR USE BY THE RECIPIENT AND MAY NOT BE REPRODUCED OR RE-POSTED, IN WHOLE OR IN PART, BY THE RECIPIENT, WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.



©2017 451 Research, LLC | WWW.451RESEARCH.COM

AtScale initially came out of the gate with an 'OLAP-on-Hadoop' play designed to enable the creation of hierarchies, measures, metadata and calculations into a virtual cube, which a business intelligence tool can use to more effectively query data live in Hadoop. While the vendor continues to serve this use case in order to enable enterprises to use their existing BI tools with Hadoop – an environment for which many weren't designed – it has also expanded its remit. AtScale's virtual cubes are now operating as a universal semantic layer, so-called because they enable business logic to exist directly on top of a big-data cluster and provide interfaces (SQL, MDX or API) to hook into BI tools.

THE 451 TAKE

AtScale's new go-to-market strategy – along with the enhancements it has made in the past six months – place the startup in a stronger position than it was this time in 2016 when we were concerned it was boxing itself into a niche by focusing on Hadoop only. As the popularity of SAP BusinessObjects Universes – from which AtScale's universal semantic layer drew inspiration – attests, enterprises need semantic model management. However, we'd like to see the vendor continue to build out support for additional back ends, as it would serve to further reinforce its value as the Switzerland of semantic layers.

CONTEXT

AtScale is no longer positioning for OLAP-on-Hadoop, as was the case when the startup first came to market, promising to bring business intelligence to the masses, in mid-2015. While it can certainly be said that the democratization of analytics in the enterprise remains a core endgame, the company is tackling this mission in a wider manner, which has led to its reframing as a provider of a universal semantic layer for BI tools.

Management notes that AtScale always intended to provide an offering that was broader than OLAP-on-Hadoop, but it chose this arena initially in order to address a particular pain point. That pain point involved enabling organizations to perform fast and interactive analytics on data in Hadoop using their existing BI tools, without moving data, creating extracts or data marts.

AtScale has now enveloped other data stores in the cloud and on-premises. And, together with fresh enterprise-oriented enhancements, it is looking to provide organizations with one semantic layer, which could be used for a variety of use cases including migration to a public cloud, or legacy database migration, without causing disruption to BI users.

The startup signaled its intentions to get its BI beyond Hadoop strategy underway when we last checked in with management in December 2016. At that time, AtScale expected to have 30 customers by year-end. The vendor cites 40 customers in total now, and a headcount of roughly 85 staff.

PRODUCT/TECHNOLOGY

AtScale 5.0 set the stage for the firm's new market focus as a provider of a universal semantic layer. The release, which became generally available in March, saw AtScale embrace Teradata, Google Dataproc and BigQuery environments and in so doing kicked off a strategy to enable enterprises to connect to any back end and get a consistent, trustworthy view (or semantic layer), while preserving investments in existing BI tooling. AtScale is essentially virtualizing how the data is stored and where it is stored to create the semantic layer.

In addition, AtScale 5.0 included fruits from an approved patent to enable its scalable dimensional calculation engine (otherwise known as its OLAP engine), to be used with any third-party BI tool. AtScale is designed to connect to the standard BI drivers an organization already uses – API, JDBC/ODBC, OLE DB or XMLA – and promises sub-second performance when used with them. AtScale also provides a schema on read approach for query optimization and is designed to capture data in formats that aren't compatible with many existing analytic offerings, such as log files in JSON documents.

Management notes that large companies use its wares to satisfy queries from customer-facing apps, and for internal usage that is a core part of daily business operations. That means it is vital for AtScale's Query Service to be deployed in a clustered mode, as this mode guards against downtime in the event of an individual server failure, or failure of a cloud service.

The latest version, AtScale 5.5 – unleashed in June – is designed to better support production-ready high availability. AtScale has, essentially, introduced a new capability known as 'hot, hot,' which is designed for when the offering is configured in high-availability mode. It enables queries to be configured (via a load balancer) to any node at any time. When configured in high-availability mode, AtScale is also able to consistently maintain query graphs to ensure the consistency of query results. It will also synchronize all metadata services across all nodes (using PostgreSQL), and ensure that SQL engines (Spark and Hive) are up and running using so-called Watcher Services.

In addition, AtScale 5.5 houses a makeover of the offering's Adaptive Cache, which has been renamed Machine Learning Performance Optimizer. The vendor has, in essence, developed a series of algorithms that are designed to anticipate certain query patterns and create aggregates to support them before an end-user query is executed against an AtScale virtual cube. They are called prediction-defined aggregates.

The query patterns AtScale anticipates are based on observations the startup has noted from production deployments with customers. AtScale developed a statistics system built on these insights, which is designed to constantly evaluate row counts and attribute cardinality and join quality of the underlying data sets. These statistics are fed into a series of algorithms, which are designed to be able to predict the potential value of creating aggregate tables in order to satisfy anticipated query patterns.

The fruits of AtScale 5.0 and 5.5 mean that the vendor has laid the groundwork for semantic model management and its pitch to provide an easy-to-use interface to model virtual cubes – while providing other features to support its positioning as a universal semantic layer for BI tools. These additional capabilities include supporting system metadata in order to maintain a graph of data lineage back to raw data and providing additional features beyond merely bolstering query performance for BI tools. Multi-fact metrics for per member, per month analysis, semi-additive metrics for counting unique customers, and multi-level metrics for calculating store vs. region performance are examples of the latter.

COMPETITION

AtScale has been grouped with fellow startups Arcadia Data and Kyvos Insights because all three got off the blocks with an OLAP-on-Hadoop play. Kyvos Insights remains focused on OLAP-on-Hadoop, having bolstered security and BI tool compatibility in the past year. Arcadia Data – like AtScale – has expanded beyond an initial focus on native Hadoop analytics. But Arcadia Data provides data visualization, discovery and BI via its own client, which is a different strategy from AtScale, which has always looked to play nicely with existing analytic offerings in order to make them shine.

AtScale's new capabilities and positioning for semantic management differentiates it from Arcadia, Kyvos and others with a native Hadoop analytics focus such as Datameer and JethroData. Now known simply as Jethro, the startup is targeting interactive BI on Hadoop using a SQL-based approach, which employs both indices and cubes to accommodate varying types of queries. Jethro could, therefore, be used as an alternative to AtScale for certain scenarios, as could Datameer, but only in the Hadoop realm.

Last, it is worth pointing out that many enterprises continue to use MicroStrategy, Tableau, Qlik, Microsoft Power BI, Excel, IBM Cognos and SAP BusinessObjects – all tools AtScale supports – without its middleware layer, despite the boons it can bring.

SWOT ANALYSIS

STRENGTHS

AtScale is morphing into a highly differentiated offering capable of addressing a wide variety of use cases and workloads. Moreover, the startup's customer base seems to be building nicely, demonstrating there is a real-world demand for its offering.

WEAKNESSES

Some enterprises won't 'get' AtScale in its new guise. Some evangelism is therefore likely to be required to elucidate its benefits. The startup will also need to communicate its differences from other semantic layers, notably a BusinessObjects Universe, which remains one of the most well known.

OPPORTUNITIES

Migration to cloud infrastructures and movement off more traditional MPP-based environments seem good new use cases to court. Although the startup has always taken a judicious approach to funding, some additional venture capital could help elevate its business to the next level.

THREATS

AtScale may have an edge in its approach but better tools don't always win out. The vendor's approach to semantic management also makes it more appealing as an acquisition target, particularly for a BI vendor looking to modernize its analytics architecture.