Is Big Data Dead? MotherDuck Raises $47M to Prove It

November 16, 2022

Brian Heater

Big Data is dead. Or so says MotherDuck, the leader of a serverless analytics platform based on DuckDB. The company’s founders say they learned from years-old users that a vast majority of workloads do not require the high overheard costs of big data distributed company thanks to recent advances.

"The fact is, Big Data is dead," the simplicity and the ease of making sense of your data is now a lot more important than its size," said Jordan Tigani, CEO and co-founder of MotherDuck, in a release.

DuckDB, an open-source, in-process database similar to SQLite for analytics workloads. According to MotherDuck, the SQL OLAP database management system has garnered widespread adoption because of its ability to run everywhere (proven by, for instance, query data from anywhere without preloading it, and even execute quick analytical queries based on up-to-date academic research. OLAP workloads are complex with long-running queries that process significant portions of a stored dataset, and changes to the data are made with several lines of business or large portions of tables being changed or added at the same time, according to DuckDB.

"To efficiently support this workload, it is critical to reduce the amount of CPU cycles that are expected per individual query. The trade-off of the set in data management to achieve this are either vectorized or join-in-line query execution engines. DuckDB employs a columnar- vectorized query execution engine, where queries are still interpreted, but a large batch of values (a 'vector') is processed in one operation," says DuckDB creator. (This greatly reduces overhead present in traditional systems such as PostgreSQL, MySQL or SQL which process such workloads in a pipelined fashion.)

"In a company like DuckDB, DuckDB Labs commented on the vision of the partnership with MotherDuck, "When the first ideas that eventually led to DuckDB were first released, we were in a phase of the wisdom in both industry and research. The only massive scale and distributed data processing would be the way forward. From our interactions with data professionals, we became convinced that while massive datasets exist, they are mostly found in organizations that already have the technological expertise to handle them. Nowadays, we see an efficient and effective engine for analytics, and we are very happy that the MotherDuck team shares this vision, especially given the team’s background.”

DuckDB Labs was founded by Michael Mühleisen and Jaap Raasveldt to build a serverless analytics platform based on DuckDB. The company says the funding will be used to further this collaboration, as well as build out its engineering and GTM teams.

"Laptops today are faster than a data warehouse. With sub-second in-database computation is no longer necessary for most workloads," said Tigani. "Cloud data warehouses are focused on the performance of 100TB queries, which is not necessarily ideal for the vast majority of users, but also offers low vendor's ability to deliver a great user experience. We are taking the power of DuckDB and combining it with serverless analytics to help scale up and scale down with ease."

"Every data lake currently stores Petabytes of data. We can now turn any of this data into insights in less than a second! This is for all workloads, from OLAP queries to ad-hoc analysis, and we are very happy that the MotherDuck team shares this vision, especially given the team’s background.”

DuckDB is an open-source, in-process database similar to SQLite for analytics workloads. According to MotherDuck, the SQL OLAP database management system has garnered widespread adoption because of its ability to run everywhere (proven by, for instance, query data from anywhere without preloading it, and even execute quick analytical queries based on up-to-date academic research. OLAP workloads are complex with long-running queries that process significant portions of a stored dataset, and changes to the data are made with several lines of business or large portions of tables being changed or added at the same time, according to DuckDB.

DuckDB is an open-source, in-process database similar to SQLite for analytics workloads. According to MotherDuck, the SQL OLAP database management system has garnered widespread adoption because of its ability to run everywhere (proven by, for instance, query data from anywhere without preloading it, and even execute quick analytical queries based on up-to-date academic research. OLAP workloads are complex with long-running queries that process significant portions of a stored dataset, and changes to the data are made with several lines of business or large portions of tables being changed or added at the same time, according to DuckDB.

"To efficiently support this workload, it is critical to reduce the amount of CPU cycles that are expected per individual query. The trade-off of the set in data management to achieve this are either vectorized or join-in-line query execution engines. DuckDB employs a columnar-vectorized query execution engine, where queries are still interpreted, but a large batch of values (a 'vector') is processed in one operation," says DuckDB creator. (This greatly reduces overhead present in traditional systems such as PostgreSQL, MySQL or SQL which process such workloads in a pipelined fashion.)

"In a company like DuckDB, DuckDB Labs commented on the vision of the partnership with MotherDuck, "When the first ideas that eventually led to DuckDB were first released, we were in a phase of the wisdom in both industry and research. The only massive scale and distributed data processing would be the way forward. From our interactions with data professionals, we became convinced that while massive datasets exist, they are mostly found in organizations that already have the technological expertise to handle them. Nowadays, we see an efficient and effective engine for analytics, and we are very happy that the MotherDuck team shares this vision, especially given the team’s background.”

DuckDB Labs was founded by Michael Mühleisen and Jaap Raasveldt to build a serverless analytics platform based on DuckDB. The company says the funding will be used to further this collaboration, as well as build out its engineering and GTM teams.

"Laptops today are faster than a data warehouse. With sub-second in-database computation is no longer necessary for most workloads," said Tigani. "Cloud data warehouses are focused on the performance of 100TB queries, which is not necessarily ideal for the vast majority of users, but also offers low vendor's ability to deliver a great user experience. We are taking the power of DuckDB and combining it with serverless analytics to help scale up and scale down with ease."

"Every data lake currently stores Petabytes of data. We can now turn any of this data into insights in less than a second! This is for all workloads, from OLAP queries to ad-hoc analysis, and we are very happy that the MotherDuck team shares this vision, especially given the team’s background.”

DuckDB Labs was founded by Michael Mühleisen and Jaap Raasveldt to build a serverless analytics platform based on DuckDB. The company says the funding will be used to further this collaboration, as well as build out its engineering and GTM teams.

"Laptops today are faster than a data warehouse. With sub-second in-database computation is no longer necessary for most workloads," said Tigani. "Cloud data warehouses are focused on the performance of 100TB queries, which is not necessarily ideal for the vast majority of users, but also offers low vendor's ability to deliver a great user experience. We are taking the power of DuckDB and combining it with serverless analytics to help scale up and scale down with ease."

"Every data lake currently stores Petabytes of data. We can now turn any of this data into insights in less than a second! This is for all workloads, from OLAP queries to ad-hoc analysis, and we are very happy that the MotherDuck team shares this vision, especially given the team’s background.”

DuckDB Labs was founded by Michael Mühleisen and Jaap Raasveldt to build a serverless analytics platform based on DuckDB. The company says the funding will be used to further this collaboration, as well as build out its engineering and GTM teams.

"Laptops today are faster than a data warehouse. With sub-second in-database computation is no longer necessary for most workloads," said Tigani. "Cloud data warehouses are focused on the performance of 100TB queries, which is not necessarily ideal for the vast majority of users, but also offers low vendor's ability to deliver a great user experience. We are taking the power of DuckDB and combining it with serverless analytics to help scale up and scale down with ease."

"Every data lake currently stores Petabytes of data. We can now turn any of this data into insights in less than a second! This is for all workloads, from OLAP queries to ad-hoc analysis, and we are very happy that the MotherDuck team shares this vision, especially given the team’s background.”

DuckDB Labs was founded by Michael Mühleisen and Jaap Raasveldt to build a serverless analytics platform based on DuckDB. The company says the funding will be used to further this collaboration, as well as build out its engineering and GTM teams.

"Laptops today are faster than a data warehouse. With sub-second in-database computation is no longer necessary for most workloads," said Tigani. "Cloud data warehouses are focused on the performance of 100TB queries, which is not necessarily ideal for the vast majority of users, but also offers low vendor's ability to deliver a great user experience. We are taking the power of DuckDB and combining it with serverless analytics to help scale up and scale down with ease."

"Every data lake currently stores Petabytes of data. We can now turn any of this data into insights in less than a second! This is for all workloads, from OLAP queries to ad-hoc analysis, and we are very happy that the MotherDuck team shares this vision, especially given the team’s background.”

DuckDB Labs was founded by Michael Mühleisen and Jaap Raasveldt to build a serverless analytics platform based on DuckDB. The company says the funding will be used to further this collaboration, as well as build out its engineering and GTM teams.

"Laptops today are faster than a data warehouse. With sub-second in-database computation is no longer necessary for most workloads," said Tigani. "Cloud data warehouses are focused on the performance of 100TB queries, which is not necessarily ideal for the vast majority of users, but also offers low vendor's ability to deliver a great user experience. We are taking the power of DuckDB and combining it with serverless analytics to help scale up and scale down with ease."

"Every data lake currently stores Petabytes of data. We can now turn any of this data into insights in less than a second! This is for all workloads, from OLAP queries to ad-hoc analysis, and we are very happy that the MotherDuck team shares this vision, especially given the team’s background.”

DuckDB Labs was founded by Michael Mühleisen and Jaap Raasveldt to build a serverless analytics platform based on DuckDB. The company says the funding will be used to further this collaboration, as well as build out its engineering and GTM teams.

"Laptops today are faster than a data warehouse. With sub-second in-database computation is no longer necessary for most workloads," said Tigani. "Cloud data warehouses are focused on the performance of 100TB queries, which is not necessarily ideal for the vast majority of users, but also offers low vendor's ability to deliver a great user experience. We are taking the power of DuckDB and combining it with serverless analytics to help scale up and scale down with ease."

"Every data lake currently stores Petabytes of data. We can now turn any of this data into insights in less than a second! This is for all workloads, from OLAP queries to ad-hoc analysis, and we are very happy that the MotherDuck team shares this vision, especially given the team’s background.”