Coverage Initiation: Decodable enables teams to build stream processing pipelines in SQL

Analysts - Alex Johnston

Publication date: Thursday, December 15 2022

Introduction

Decodable is a startup that offers an eponymous stream processing platform. It enables application and data engineers to define data pipelines in SQL, using the popular programming language for transformations, aggregations and stateful processing, among other applications. The company incorporated in 2021, took in a $20 million series A funding round in February and formed a strategic partnership with DataStax in August.

The Take

The complexity of establishing and maintaining stream processing architectures is widely acknowledged. As the costs of real-time data have become less prohibitive, skillsets are increasingly the bottleneck to leveraging the technology. Decodable is seeking to address this bottleneck directly by letting teams establish capabilities that can filter, route, enrich or transform data streams using SQL, and easily build streaming applications.

Context

Privately held Decodable was founded in January 2021 by Eric Sammer, chief executive officer and director of the company. Sammer previously cofounded Rocana, then known as ScalingData, in 2014, where he served as CTO. Decodable raised $20 million in series A funding announced in February, adding to $5.5 million raised in its seed round. Investors include Bain Capital Venture Partners, VR Adviser LLC and Harpoon Ventures Management.

The company currently employs over 25 people, with near-term objectives to hire additional talent, and suggests it may double its full-time employees next year. Leadership is conscious of the

macroeconomic environment, however, and keen to not overstretch. Although headquartered in California, the company lacks a physical office — the team is highly distributed with headcount in countries including the U.S., Canada, Germany and Morocco.

Stream processing refers to technologies that support the processing or analyzing of event streams. Companies looking to leverage real-time data require an ingest stream into analytical systems or microservices, or a tool that processes data in motion. A number of open-source projects, such as Apache Flink and Apache Kafka, have modernized enterprise streaming architectures to address this need, but working with them directly leads to developers building at low levels of abstraction.

A burgeoning vendor space has emerged to manage much of the underlying infrastructure associated with these technologies, and to simplify the process of maintenance and development. Decodable is a new entrant into this space, with a platform built on Apache Flink foundations.

Technology

Decodable suggests that its primary objective is to increase the accessibility of stream processing. It is designed to abstract away much of the complexity associated with managing real-time data pipelines, giving developers the ability to apply SQL — a popular programming language and one supported by many database systems.

The platform has been designed so that developers with experience working with SQL queries in Snowflake, Teradata or BigQuery, for example, will be able to easily transition into building event-driven applications or establishing data ingestion pipelines. The company employs consumption-based pricing, using a unit of "tasks." This is a set amount of resources that a customer can designate to a workload and it is treated as a maximum, with Decodable auto-scaling up to that requirement as needed.

For complex infrastructure, developers may need low-level access, but Decodable suggests that much of the work undertaken by developers centers on minor changes or simple applications. For those use cases, the Decodable platform abstracts away the complexity. Users can preview the processing output of pipelines, and the service offers a number of protections to ensure the viability of pipeline transformations coded in SQL.

When a developer provides code, error messages provided by the platform can identify whether the submitted code may break pipelines, isn't valid, or is missing needed information. A major focus area for the business is ensuring that for more challenging designs, the platform can integrate with the technologies developers are comfortable using. Decodable provides a library of connectors for source and sink systems to connect to external data sources and messaging systems.

The platform is built on the foundations of Apache Flink, but a number of additional capabilities have been added to support development at a higher level of abstraction. The other open-source technology leveraged by the platform is Debezium, a tool for change data capture.

Decodable is separated into a centralized control plane, where users submit pipelines and observe lineages, and distributed data planes held in different cloud regions, providers or even on-premises. This is designed to support multinational enterprises that are challenged by data residency requirements. An area of focus within Decodable's product roadmap is visualizing performance metrics, understanding what users want to have visibility into, and the best ways of presenting it.

Go-to-market

The personas Decodable targets are application developers and data engineers. The company notes it is not currently focused on the business analyst persona, and that this has framed how its technology has evolved. The need to align with the tools and practices of application developers and
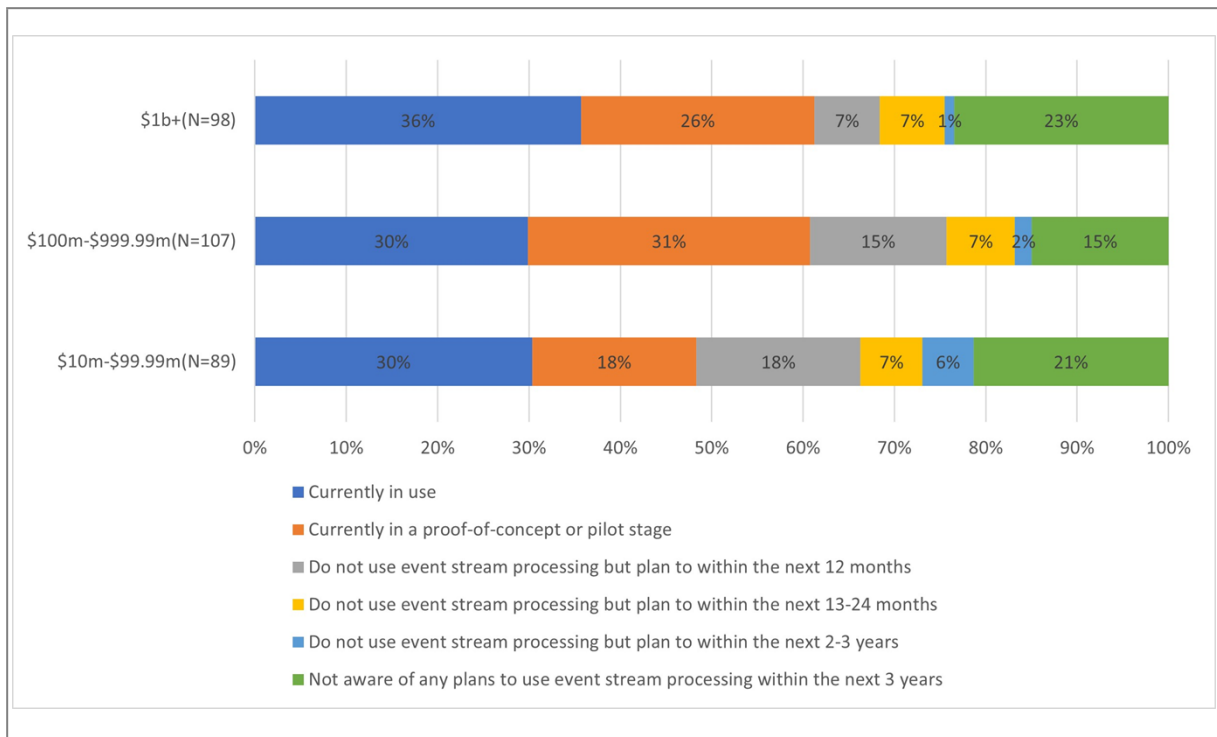
451 Research

S&P Global
Market Intelligence

data engineers has been a guiding force for the business. Decodable suggests that the pain points for these users commonly derive from having to maintain operations for the wide array of data pipelines in use within businesses, often very simple ones.

The company see clear use cases in a diverse range of industries including retail, logistics, gaming, media and advertising technology. It is in the process of putting together public customer references, but noted it is working with at least one company on shipping container movement and another on customer visibility, and is engaging with a prospective customer on gaming cheat protection.

Decodable suggests its platform is less about ingest into analytics products, and more a cloud service that customers are using to build product features or product services. It perceives its strengths to be event-based microservice use cases, such as those surrounding updating email addresses, where a change data capture event is generated within a database, and Decodable will filter and normalize that request before making an API call to ensure a confirmation email is sent.

A small commercial team meant it had to carefully target its go-to-market strategy. North American clients are its initial focus — the company feels best placed to service them with its existing infrastructure. Paying customers are currently single digits, but Decodable is targeting middle- and upper-middle-revenue enterprises in an attempt to scale that number quickly, seeking to avoid the lengthy sales cycles associated with the largest enterprises. As shown below, this market has a keen interest in event stream processing technologies, and appears to be a key growth market.

Event Stream Processing Technology Growing Quickly in Middle Market

Source: 451 Research's Voice of the Enterprise: Data & Analytics, Data Platforms, 2022
Q. Which of the following best describes your organization's event stream processing adoption plans?
Base: All respondents

Decodable suggests around half of prospective customers that are in discussions already have some form of vendor relationship to address relevant use cases, but many companies have either not

451 Research

S&P Global
Market Intelligence

invested in event stream processing technologies at all, or have a self-managed approach. For those with existing stream architectures, Decodable can be used to strip out the existing stream processing components. But commonly when the technology is being applied, it is used to integrate different messaging and streaming technologies, the company said.

Partnerships

Decodable has a prominent partnership with DataStax, which is best known for Astra DB, a cloud database as a service based on Apache Cassandra. In June 2021, it announced the beta release of Astra Streaming, adopting Apache Pulsar as another core project.

The partnership brings together a streaming stack built on open-source foundations, in offering an AstraDB database service built for scalability and availability, event streaming architecture through Astra Streaming, and Decodable's Stream Processing technology. This alliance is a logical one, and the company suggests that database service providers, as well as data streaming companies without a processing component, are obvious partners.

Competition

Direct competitors include Ververica, an Apache Flink stream processing and analytics vendor, and Aiven, which has an Apache Flink service within its managed open-source offering. AWS is also active in this space with Amazon Kinesis Data Analytics, another fully managed service. Companies such as Databricks and Cloudera with Apache Spark streaming engine offerings, an in-memory batch and streaming processing project, can also be seen as competitors.

Apache Spark Structured Streaming is a micro-batch system, and has some limitations around correctness and recovery, but is often applied to the same use cases as Apache Flink. Confluent, a cloud-native data-streaming platform based on Apache Kafka, is another company Decodable comes up against in commercial discussions.

Decodable's competitive positioning stresses simplicity and development speed as differentiators. It is confident about the strength of Apache Flink in addressing real-time enterprise use cases, but perceives the complexity of adopting the technology as its greatest barrier to broad adoption. Unlike many other real-time pipeline tools that employ drag-and-drop interfaces to simplify development, Decodable can be interfaced with through a RestAPI, a user interface providing an IDE-like experience, or directly using its command line tool.

It sees its differentiated SQL support as aligning more closely with the needs of its core personas. Although Decodable acknowledges that some of its competitors have SQL interfaces, it suggests these are not adequately intuitive, and that the developer experience is commonly poor.

SWOT Analysis

Strengths	Weaknesses
A focus on simplicity and aligning with the experiences of application developers and data engineers is a well-thought-out value proposition. The depth of SQL expertise already within many businesses will widen the accessibility of the technology.	Decodable, being powered by Apache Flink, will not be able to differentiate itself from better established Flink competitors through latency. Although growing, it currently lacks the marketing and sales infrastructure to contest all prospective markets, and will need to pick its battles.
Opportunities	Threats
The event stream processing market is growing at pace, and Decodable's decision to predominantly target the upper middle	As with all vendors building from open-source projects, low barriers to market entry may pose a threat. The

451 Research

S&P Global
Market Intelligence

| market appears to be a sensible one. The cost of "real time" architectures historically restricted performant streaming capabilities to the largest enterprises, but accessibility has, and is, improving markedly. | growing popularity of stream architectures could encourage a wider array of data platform providers to establish or integrate competing offerings. |

Source: 451 Research