



KIXEYE

Building a Cloud Data Warehouse at KIXEYE



CUSTOMER KIXEYE

PARTNER Snowflake



CASE STUDY

Founded in 2007, KIXEYE is a mobile and online gaming company that is redefining the intersection of fidelity and accessibility to create innovative experiences for competitive gamers. KIXEYE currently has over 300 employees based in its San Francisco headquarters, along with international offices in Brisbane, Australia, and Victoria, Canada. KIXEYE is responsible for popular titles like VEGA Conflict, Battle Pirates, War Commander, and Backyard Monsters.

KIXEYE develops a smaller number of high-quality games, many of which are free to play. This approach makes it critically important to ensure that games are not only engaging, but also create revenue from virtual goods and other avenues. Data analytics plays a central role in helping KIXEYE meet those goals, supporting constant experimentation with new features, functionalities, and platforms.

“Data is extremely important to us,” said Josh McDonald, Director of Analytics Engineering at KIXEYE. “We look at millions of events every few minutes to make decisions and iterate quickly on many aspects of our games in order to make them fun, interesting, and profitable.”

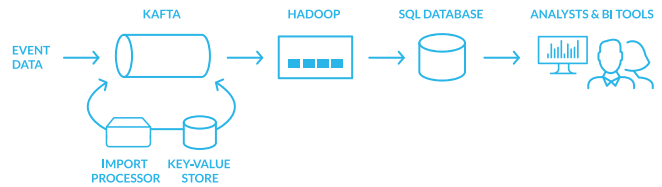
“Snowflake has made it easy to get data to analysts faster. Its unique architecture makes it possible to have all of our teams working with data in one system, without performance impact and without needing to be database administrators.”

— Josh McDonald, Director of Analytics Engineering

KIXEYE'S SCENARIO

KIXEYE's data processing pipeline captures events from games, collects and processes them, and then makes the data available via SQL and business intelligence (BI) tools for exploration, reporting, and predictive analytics. Their event data is primarily in JSON format to make it easy to read, to produce and to manipulate the structure as needed. However, making that data accessible with good performance to SQL analysts and tools like Tableau required transforming the data into aggregated, relationally-structured data in multiple steps.

KIXEYE's previous architecture funneled event data into a Kafka pipeline, where a set of jobs handled initial processing, and then wrote the data into a Hadoop cluster. That data was then imported into Hive tables. It quickly became clear that this architecture would not meet their long-term needs.



“Our architecture had significant complexity,” said McDonald. “It also introduced a lot of latency, in part because it required doing a significant amount of ETL processing in Hadoop, which ultimately wasn’t scalable enough or extensible enough to meet our needs.” In addition, it was painfully expensive for KIXEYE to host and maintain all of the infrastructure needed to summarize, aggregate, and ultimately load data into a SQL database. That database was needed to support querying, reporting, and visualization by SQL analysts and BI tools such as Tableau.

FINDING A BETTER SOLUTION

In exploring alternatives, KIXEYE wanted a solution that was scalable, extensible, timely, accurate, cost-effective, and user friendly. In McDonald's words, “Our desired system needed to be able to handle the constant change in the structure of our logging event data and do so without significant engineering overhead or operational cost.” KIXEYE heard about Snowflake and decided to take a closer look. “We brought in Snowflake, sat them down and grilled them on how their system was really going to do all that they said it was going to do,” recalls McDonald.

One reason that KIXEYE became interested in Snowflake was its ability to support their data transformation and processing needs in a far simpler

way. Snowflake's flexibility and its native support for semi-structured data would make it possible for them to build a significant share of their data pipeline inside Snowflake.

KIXEYE also valued Snowflake's ability to support diverse concurrent workloads. “We had been taking raw event data and flowing it into separate data marts for each game and team because teams run all kinds of ETL and analytic jobs at all times of the day and we needed to avoid contention,” says McDonald.

“Snowflake's separation of data from processing allows all of our teams to work with data in one system, without worrying that one group or job will impact another.”

Another important selling point for Snowflake was that it did not require a lot of management and tuning. “I’m not interested in having database administrators on staff that spend all their time administering databases,” said McDonald. “With our previous system, we had to have

people constantly making sure that the key-value store was working appropriately, that Hadoop was doing the things that we needed it to do, that data was being stored appropriately, and so on. With Snowflake, I don’t have to be a database administrator. I really like that.”

SEEING RESULTS

Based on their evaluation, KIXEYE made the decision to deploy Snowflake. In their new architecture, KIXEYE outputs the JSON event data directly from Kafka to Amazon Web Service’s S3 storage service, from which they load into Snowflake to make it accessible for processing and analysis. All of KIXEYE’s data processing now happens inside Snowflake, with each group able to operate independent virtual warehouses for transforming, slicing, aggregating, and analyzing data.

KIXEYE can now make data available for analysis an order of magnitude faster than before. “Previously, it took more than two hours from the time that data arrived to the time that it was fully queryable. That just wasn’t timely,” says McDonald. “With Snowflake we can process upwards of 1 million batches of events every couple of minutes.”

KIXEYE is also taking advantage of Snowflake’s elasticity to give each group the resources it needs without breaking the bank. “The ability to have warehouses sized appropriately for processing our loading of data and internal ETL, each happening in parallel so that our marketing, finance, and gaming teams can work at the same time, without contention, is huge,” notes McDonald. “Our internal analytics engineering processes run 24x7, but our game teams aren’t actually running their analytics queries 24x7. With Snowflake we only pay for what we use. That saves a huge amount of money compared to our previous hosted environment—in a hosted or leased environment you are overpaying by definition because you are forced to purchase capacity in advance.”

Snowflake’s native support for semi-structured data is also getting heavy use. “I can’t say enough about how fantastic the native JSON support is,” said McDonald. “I’ve never actually seen anything that worked until now. Using SQL, Snowflake lets us load our JSON data as-is into our event tables and then parse that into views that tools like Tableau can understand and it works fantastic. My analysts are really happy about this.”

KIXEYE DATA FLOW WITH SNOWFLAKE



