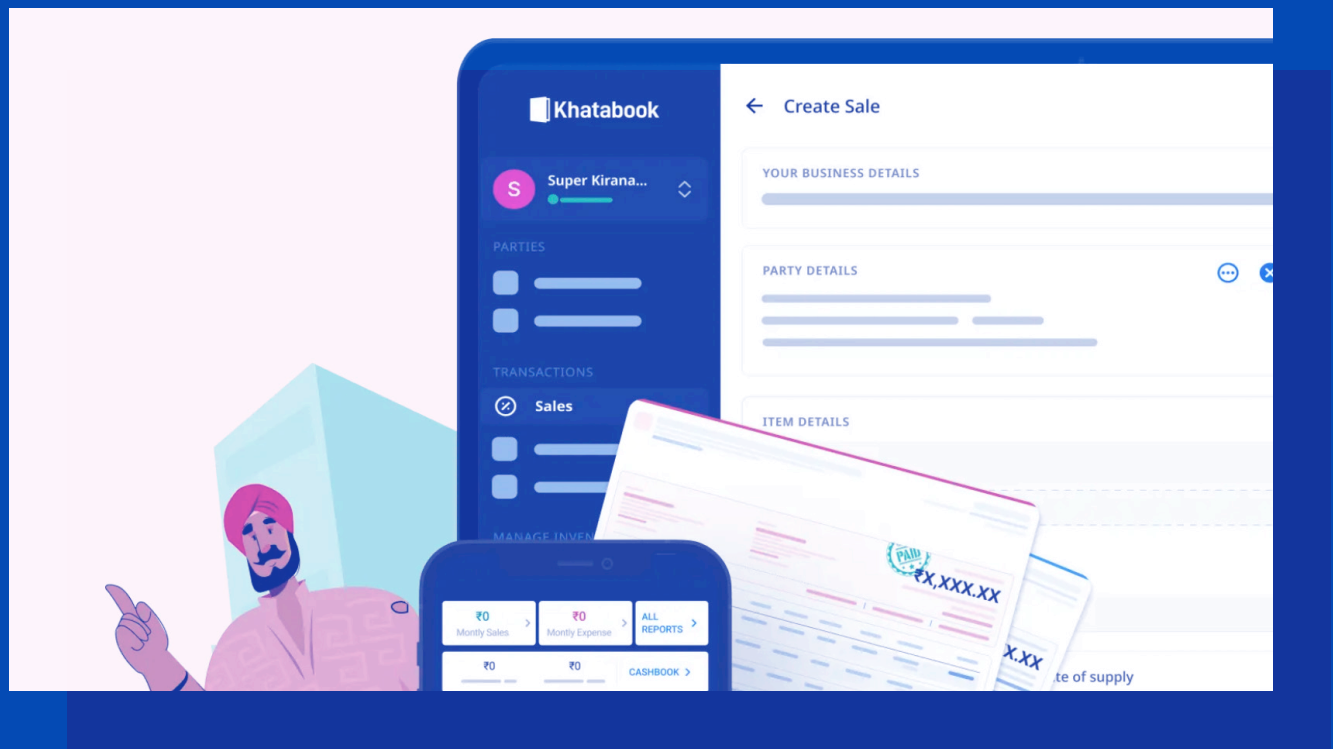rudderstack™

Khatabook Replaces Segment with RudderStack to Allow Cost Effective Scaling to Over 6 Billion Monthly Events

Highlights

- Khatabook moved from Segment to RudderStack to support its explosive growth while managing costs.
- On-prem RudderStack is GDPR-compliant, opening the door for ISO Certification. The hosted enterprise version is also GDRP-compliant because RudderStack doesn't store customer data on its servers.
- RudderStack's warehouse-first architecture enabled Khatabook to collect all their customer data, build a complete view of their customers in their warehouse, and make those profiles available to every team.

Key Stats

- Khatabook processes 4 billion monthly user events on a 64-core 64GB on-prem server and can process double that amount with its current configuration.
- Khatabook engineers took 15 days to validate the company's data sets across all front-end sources and back-end destinations, confirming that RudderStack resolved user identities in the absence of historical data.
- Khatabook reduced CDP spend by 90% over Segment by switching to RudderStack.

"Khatabook's digital ledger app revolutionized how India's 63 million MSMEs record and track everyday cash and credit transactions. Our user-friendly mobile interface empowers rural businesses and neighborhood shops to move from paper to digital bookkeeping. With 50 million downloads, and 10 million monthly active users generating 4 billion events every month, our Segment CDP struggled to route data to downstream destinations and cost too much. Switching to on-prem RudderStack reduced our CDP spend by 90% and sped up routing to more destinations while ensuring GDPR compliance."

Overview

Founded in 2019 in Bangalore, Khatabook is a fintech start-up that launched a free mobile digital ledger app for India's 63 million MSME (micro, small, and medium enterprise) users. The app allows businesses to record and track cash and credit transactions and send payment reminders to customers. Users include rural businesses and neighborhood shops that are working with digital accounting software for the first time.

To ease the transition from paper-based ledgers to digital accounting, Khatabook modeled its app on popular communications tools like WhatsApp, making it instantly familiar to neophyte users. Available in 13 languages and used in 3,000 Indian cities, Khatabook has over 50 million downloads, 10 million monthly active users, and two million daily users.

Khatabook acquired Biz Analyst for $10 million in March 2021. The mobile app syncs to Tally ERP 9/Prime, India's leading enterprise management platform. It has 150,000 monthly paid users, and Khatabook hopes to double that number to 300,000 by December 2023. One of India's fastest-growing companies, the start-up secured $100,000 million in funding led by Tribe Capital and Moore Capital Ventures in August 2021.

Khatabook is now moving into the digital lending space and piloting its lending products in partnership with four non-banking finance companies. The start-up is also exploring supply chain financing as it looks to expand its revenue-generating products and offer new digital-first financial solutions to businesses in the world's sixth-largest economy.

Challenge: Segment Was too Slow and Costly to Scale as Khatabook Grew

As a start-up on the path to profitability, Khatabook was massively scaling. To fuel growth, the company needed to increase visibility into customer events while reining

in costs. To strike the right balance, Khatabook needed to re-evaluate the price-per-formance ratio of its technology stack. A deep dive into its tools revealed a significant pain point.

Within a few weeks of launching Khatabook in 2019, app users were generating one billion events a month. "By September 2022, we had grown to 125 million events daily and nearly four billion monthly events," says Pathak. "I expect us to reach 250 million daily events and six billion monthly events in the next year. To continue scaling, we needed a cost-effective way to see what our users were doing on our platforms, and we found our CDP lacking."

Khatabook was using Segment as its CDP on a pay-per-use plan, and the costs were skyrocketing as the company's active user base grew. At the same time, Khatabook began working toward ISO certification, which required the company to employ GPDR-compliant infrastructure.

> "As Khatabook grew, we had to limit what events we sent to Segment. It was far from ideal," adds Pathak's colleague, Sakshi Barnwal (also a Software Development Engineer III). "To control costs, we had to prioritize the metrics we wanted to analyze, which limited our ability to run experiments and try new things. It was already holding back our ability to scale. And another problem lay ahead of us. Segment stored user event data on its servers instead of ours. Even if we could do everything we wanted with Segment, we would have to give it up at some point because it would block our route to ISO certification."

Cost was not the only issue. Khatabook struggled to route user event data from its front end to its back end, and there were also speed issues. "Our sources include Android and iOS apps, Flutter, JavaScript, and Node.js," adds Barnwal. Destinations include Mixpanel, CleverTap, and Amazon S3. We wanted to push data directly to Snowflake instead of going through S3, but there was no way to do this directly from Segment, and flowing data into CleverTap was painfully slow. Faced with these limitations and rising costs, we decided to look for another CDP and see how it

compared to Segment."

Pathak and Barnwal chose RudderStack as the potential successor and spent the next two weeks running head-to-head tests to see how well the two platforms performed. RudderStack emerged as the winner, taking Khatabook in a new direction.

Solution: Taking Control of Khatabook's CDP Costs and Technology Stack with RudderStack

With cost as a critical factor, Khatabook chose RudderStack and tested it against Segment. Pathak and his team ran a POC to evaluate how well the platform routed user event data from the company's front end to its back end.

"We looked at how well RudderStack flowed events to back-end destinations, and the results were fantastic," says Pathak. It was much faster than Segment and started flowing events to CleverTap immediately. We also routed data directly from RudderStack to Snowflake without going through S3, and RudderStack's warehouse-first architecture sped up that process too."

Khatabook also benefitted from the support of the RudderStack community. Before installing the new CDP, Pathak and Barnwal read the online documentation and joined the community support Slack channel. They asked a lot of questions about the feature set and were pleased with the rapid response.

> "Within a couple of hours of registering for the Slack channel, the community, including developers from RudderStack, started answering our questions," says an enthusiastic Pathak. "It was much faster than Segment's enterprise help desk, which took 12 to 14 hours to respond to a ticket. The community helped us roll out RudderStack to our internal cloud and supported us throughout the process. We did a blue-green deployment to resolve any issues, coded some custom logic that allowed us to start scaling right away, and kept Segment running until we could phase it out completely."

Bringing RudderStack in-house gave Khatabook complete control over the cost of its new CDP. "RudderStack is extremely efficient," adds Barnwal. "Our instance runs on a 64-core server with 64 GB of RAM. We process 125 million daily user events and can easily scale to handle twice that capacity with our current configuration. We were struggling with Segment and fretting about the cost per user. Now we're free to experiment and route as many events as needed from our front-end sources to our back-end destinations. Unlimited event tracking has given us room to grow."

While RudderStack's event stream product is sufficient for Khatabook's present-day needs, the company has an eye on the future.

"RudderStack gives us the capacity and basic functionalities to scale our Khatabook and grow our paid Biz Analyst subscription base," explains Pathak. "As we move into lending and supply chain financing, I expect we'll add more destinations and start using ETL and reverse ETL."

Result: A CDP that Can Scale into an Enterprise Solution

Replacing Segment with RudderStack reduced Khatabook's CDP spend by 90%. "During the testing phase, our Engineers determined that the RudderStack SDK makes 33% fewer API calls than Segment. That's a tremendous benefit," Pathak shares.

Hosting RudderStack on the company's internal cloud complies with the GDPR and opens the door to ISO certification. The company will remain compliant if Kathabook upgrades to the enterprise version of RudderStack, sparing the company's engineers and data analysts from learning a new CDP.

> "RudderStack's warehouse-first approach puts us firmly in control of our customer's data, explains Barnwal. "Unlike Segment, RudderStack will never store our data on its servers. RudderStack collects data from our front end, transforms it, and routes it to our back-end destinations, including our Amazon S3 data lake and our Snowflake data warehouse. This workflow is 100% GDPR-compliant."

Khatabook is using RudderStack Transformations to optimize its data stream by filtering the events it sends to downstream destinations. The data engineering also used RuddersStack to validate the data being routed downstream.

"We took 15 days to analyze our data sets across all sources and destinations, and everything checked out," continues Barnwal. "We expected issues with identity resolution because it requires historical data, and fortunately, they never materialized. The transition to RudderStack was seamless in that respect."

"Decision-making is vastly improved," adds Pathak. "We can build a funnel in runtime and start running campaigns on top of that. Khatabook can target the right customers in a timely manner and scale faster, which is a big win."

Thanks to RudderStack, Khatabook can continue to scale as it grows its paid subscriber base, moves into digital lending and supply chain management, and works toward profitability.

"The Khatabook app was an instant success," concludes Pathak. "Our 10 million monthly active users generate four billion events every month. That number will increase 25% to 50% over the next year as we grow our Biz Analyst paid subscriber base and venture further into the fintech space. RudderStack allowed us to optimize our technology stack at substantial savings. It also offers future savings should we move to the paid Enterprise tier to use the platform's ETL and reserve ETL function-alities as we grow our product portfolio, increase our paid user base, and scale our business."

Destinations	Mixpanel, CleverTap, S3, Snowflake
Sources	Android, iOS, Flutter, Javascript, Node.js

Warehouse | Snowflake