



## RingCentral Dials Up a Data Lake for Quality of Service, Fraud Detection and Product Usage Analytics

### Overview

RingCentral, Inc. is an award-winning global provider of cloud-unified communications and collaboration solutions, providing innovative and secure ways to connect employees, boost workforce productivity, share knowledge, and strengthen customer relationships.

It provides an open platform that integrates with today's leading business apps while giving customers the flexibility to customize their own workflows. RingCentral solutions empower today's mobile and distributed workforces to be connected anywhere and on any device through voice, video, team messaging, collaboration, SMS, conferencing, online meetings, contact center, and fax.

RingCentral achieved over \$500 million in revenue last year by handling over 200 million calls a month from over 350 thousand business customers. RingCentral experienced 32% year over year growth and revenue last year.

### Challenges

RingCentral had several strategic goals that involved leveraging data insights, but the volume of data and complexity of data formats, such as CDR (Call Detail Records), made ingesting data, managing data pipelines and applying analytics a real challenge.

RingCentral's Michael Becker stated that *"the challenge for us was to actually take all of that information and weave it together, to give us a 360-degree view of a conference call."*

There were three mission critical areas in building out a 360 view of conference calls which, with the help of Cloudera and the StreamSets DataOps platform, allowed them to overcome many challenges.

### Quality of Service

Handling over 200 million calls per month, RingCentral lives or dies based on whether call quality meets customer expectations. RingCentral had many capabilities for forensically assessing call quality after the fact but aspired to detect and remediate poor quality in real-time. Real-time quality insight would also help them quickly pinpoint root cause of poor call quality, whether in their internal systems or within their providers' or customers' networks.

### Fraud Detection

For communication providers, fraud is an ever-present threat. Malicious users hijack accounts and mount large charges on international carrier networks that adversely impact RingCentral's customer base and the company's brand reputation. Detecting fraud through real-time analysis of network usage and carrier records is key to shutting down bad actors, but with legacy solutions RingCentral could only monitor a limited set of resources.

### Product Usage

As with every SaaS provider, RingCentral, wants to understand how customers are using their technology in order to improve customer experience as well as product stickiness and virality. Marketing wants to understand how their product is getting adopted and which features deliver the most value to customers. Finance wants to understand growth in user counts and activity. Sales wants to have intelligent conversations with their customers based on actual user behavior, and Customer Success needs specific usage data to identify and resolve support tickets.

While a number of departments need usage data, RingCentral lacked a consistent way to stream all the data for a given account, and required many person hours to create reports and run analyses.

## Solution

To address these real-time analytics challenges, RingCentral partnered with StreamSets and Cloudera to build out dataflows into an enterprise data lake to store and analyze its data. RingCentral manages their data pipelines with StreamSets DataOps Platform which helps them ensure relevant and reliable data is immediately available to all areas of the organization.

RingCentral’s data lake architecture and dataops platform combined allows them to run multiple business imperatives on a single set of data across multiple analytics solutions. In addition to their data lake, RingCentral also does widespread ETL and data processing for a variety of data sources using StreamSets. They push that data to a Vertica database for access via Tableau for business intelligence.

By leveraging StreamSets Intelligent Pipeline technology to build their data movement topologies, RingCentral can build dataflows that automatically adjust to data drift, changes in data structure that happen in the normal course of expanding their data sources and modernizing infrastructure. Since network activity is not static, the platform needs to handle traffic bursts, such as when multiple offline edge providers join at the same time. Gracefully handling data drift and traffic bursts ensures that marketing, product, sales, and finance all have timely access to analysis-ready data.

## Quality of Service

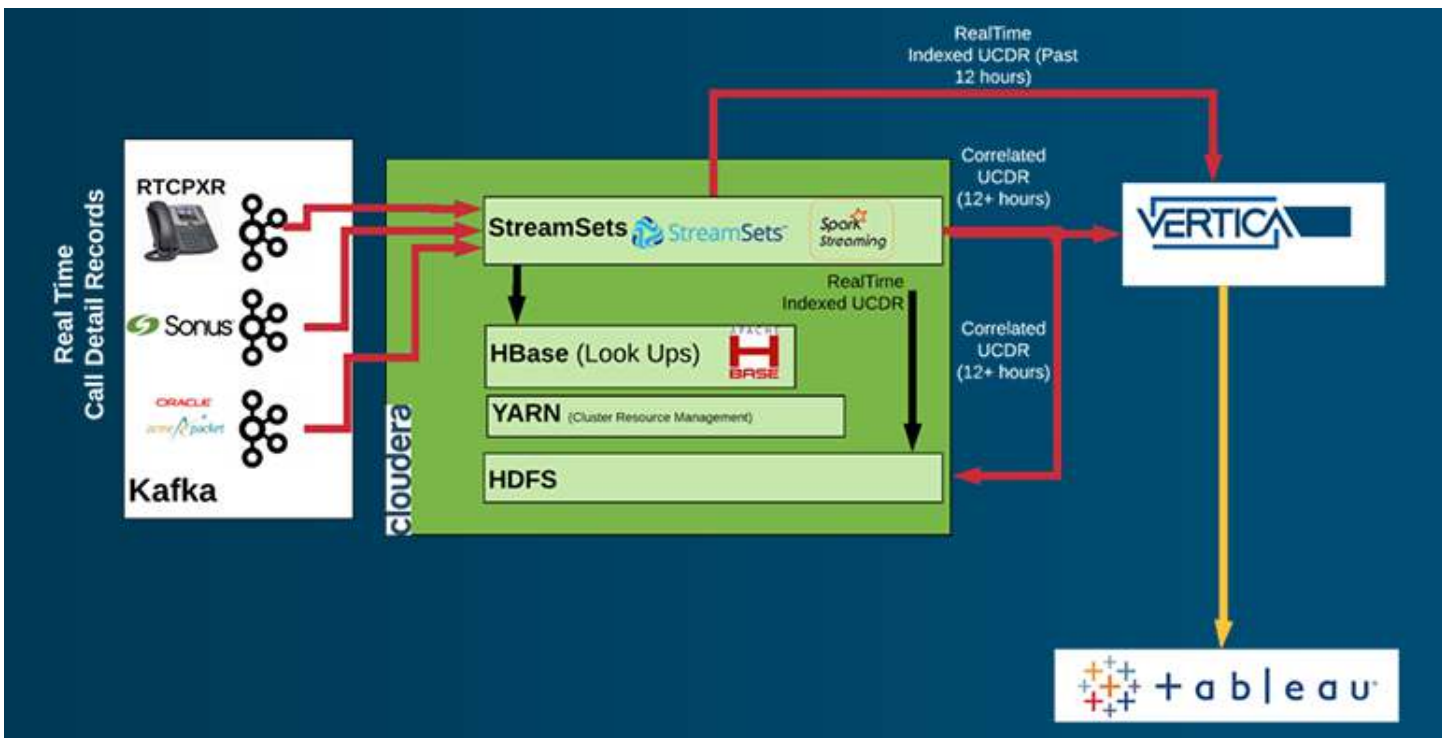
RingCentral leveraged StreamSets Data Collector and scaled resources so that call detail records (CDR's) are evaluated in real-time through the use of Apache Spark. The distributed processing and elastic scaling capabilities of Data Collector and Apache Spark allowed RingCentral to burst processing capacity to more frequent quality sampling so that analysis happens faster and with higher fidelity.

## Fraud Detection

The product usage data lake also alerts when usage and carrier rates exceed thresholds based on historical norms, which usually indicates a fraud incident. Leveraging this alerting they can now identify and take action against bad actors faster.

## Product Usage

RingCentral created an enterprise data lake, which records all product usage and ties that information to a specific user, tracking daily activity. This gives them full visibility on product trends and customer feature utilization. These insights are shared between departments to optimize a variety of processes and feed real-time reports and dashboards.



The StreamSets DataOps platform allows RingCentral to gain real-time call quality insight

## Results

By leveraging StreamSets and Cloudera, RingCentral has addressed its 3 targeted initiatives, while also creating new capabilities and opportunities for product and solution innovation.

### Quality of Service

RingCentral now addresses quality of call service in real-time and can make immediate adjustments to their network and communicate issues to their carrier partners and customers. By detecting sub-optimal call quality, they can directly reach out to customers and get ahead of potential satisfaction issues. Given the importance of low churn rates to a services business this ability to attend to quality in real time can substantially impact revenue.

### Fraud Detection

RingCentral now monitors usage and carrier charges in near real-time allowing them to identify bad actors before they incur large charges for their customers. This helps them to more efficiently prevent fraud and protect their brand reputation.

### Product Usage

RingCentral now has a single place to retrieve product usage information, with roll-up information based on each use case. Product usage is now analyzed in real-time and they continue to enhance this product usage database. Sales, marketing, engineering, operations and support all have self service access to the data for use in their specific analytics activities, complete with access controls.

## The Future

RingCentral is delivering unified communications with the power and intelligence of real-time data. By building a unified data lake and managing dataflows holistically they were able to address existing initiatives.

But, each application creates a new opportunity for future capabilities. RingCentral is looking to implement more democratized data science work on their data. They are exploring a standard set of tools that allow for data scientists and data engineers to work more efficiently together and self-serve on the growing data volume. By opening up more users to predictive analytics, RingCentral can begin predicting potential issues like customer churn, network connectivity, identifying fraud, marketing spend analysis, and discovering associated products and services.

For more on RingCentral and StreamSets including architecture and performance [watch the video](#).

## ABOUT STREAMSETS

StreamSets transforms how enterprises flow big and fast data from myriad sources into data centers and cloud analytics platforms. Its DataOps platform helps companies build and operate continuous dataflow topologies, combining award-winning open source data movement software with a cloud-native Control Hub. Enterprises use StreamSets to enable cloud analytics, data lakes, Apache Kafka, IoT and cybersecurity.

Founded by Girish Pancha, former chief product officer of Informatica, and Arvind Prabhakar, a former engineering leader at Cloudera. StreamSets is backed by top-tier Silicon Valley venture capital firms, including Battery Ventures, New Enterprise Associates (NEA), and Accel Partners.

For more information, visit [streamsets.com](http://streamsets.com)

## LEARN MORE

To view more StreamSets customer stories, go to [www.streamsets.com/case-studies](http://www.streamsets.com/case-studies)

For additional information about StreamSets or StreamSets Services, please contact your StreamSets Account Representative or visit us online at [www.streamsets.com](http://www.streamsets.com)