# RAFAY



# Distributed Data Ingestion and Preprocessing

## Challenge

Smart watches, car sensors, wireless access points - everywhere we look there are billions of devices and endpoints generating terabytes of data daily. With advances in artificial intelligence and machine learning, we can process all this data and come up with meaningful, actionable insights.

But this data is being generated on endpoints, which is physically far away from the scalable cloud infrastructure that needs to process it. Hence, it must be moved from the end device to a data center or public cloud, and then kept in a highly available storage service. Not all data being generated on endpoints may be relevant, so an ETL (Extract, Transform, Load) service may be used to determine which data is useful. Consider that 90 percent of the data could potentially be deemed not useful, which means 90 percent of the resources applied to the problem are wasted, not including the time it takes to carry out this complex processing.

Instead, what if the data being generated by devices and endpoints could be processed closer to the end point?
- Data could be economically processed in small chunks and aggregated over time.
- Application owners could write logic to prune, pre-process, summarize or transform data without needing to bring it all back to the cloud or data center.
- With only relevant, actionable data being sent, data analysis or machine learning clusters can efficiently process data and generate actionable insights faster.
- Network, storage and processing requirements could be significantly reduced.

Many enterprises would prefer to adopt such a strategy to deal with device-generated data, but lack access to distributed infrastructure and tools to internally develop such data preprocessing workflows.

## Solution

Rafay enables enterprises to deploy distributed data ingestion and preprocessing across a network of distributed edge clusters - without any DevOps effort. Called the Programmable Edge™, Rafay's platform delivers powerful yet easy-to-use APIs for deploying microservices worldwide in minutes. Enterprises can leverage Rafay's multi-tenant network to gain instant access to global locations close to end users or deploy Programmable Edge technology in their own private edge application network.

The platform applies a variety of patent-pending algorithms to optimize microservice placement decisions based on configured policies. The platform also provides a suite of application lifecycle management tools to help developers deploy and run their applications across a number of edges, so you don't have to.

By leveraging Rafay's Programmable Edge platform, developers can deploy specific microservices responsible for distributed data ingestion and preprocessing closer to endpoints. As a result, enterprises can target aggressive time-to-value goals while deploying increasingly complex pre-processing workflows while reducing overall ownership costs.

## Benefits

- **On-Demand Expansion of Your Application's Global Footprint:** All traffic from endpoints is handled by your containerized microservices, which are dynamically placed close to endpoints.

- **Highly Improved End User Experience – Globally:** Distributed API routing, global load balancing, end point authentication and policing ensure that end users experience a consistently high performing application.

- **Developer-Friendly Application Lifecycle Management Tools:** A suite of application life cycle features empower developers to deploy microservices to the large number of edge clusters without worrying about internally developing code and artifact deployment, log aggregation, and a variety of other tools.

# RAFAY

530 Lakeside Drive
Suite 210
Sunnyvale, CA 94085
info@rafay.co
(669) 247-2551