Remote Sensing for Forest Fires

Cloud Data Storage Design

UBC Cloud Innovation Centre

Capstone Team CG-23

Mark Lee, Austin Li, Ethan McKeen, Gurjot Sandhu, Christy Zhang

2024-04-30

Table of Contents

Executive Summary

This document outlines the data storage specifications for handling satellite data, sensor data, notification subscription data, and machine learning models using AWS services. We utilize Amazon DynamoDB for its low-latency data retrieval capabilities, ensuring that our diverse datasets are managed efficiently and securely. This infrastructure supports our project's need for scalability, reliability, and security, facilitating improved efficiency across the application.

1. Data Storage and Structure

Our system utilizes Amazon DynamoDB to manage a variety of data crucial for operational effectiveness. The data is categorized into distinct DynamoDB tables: satelliteData, sensorData, sensors, and userSubLocations. The primary key schema for each table comprises a partition key and a sort key. Each table is tailored to store specific data types such as satellite information, sensor readings, sensor metadata, and user notification information, respectively. For a detailed description of each table's schema, including attribute definitions and data types, please refer to the "Database Schema" document. The structure outlined there provides the foundation for our indexing strategies and data access patterns.

2. Scalability, Security, and Reliability

DynamoDB is inherently scalable, with automatic scaling features that adjust capacity to maintain performance as demand fluctuates. This ensures that the throughput adapts in response to application activity. By utilizing partition keys that are designed for an even distribution of data, DynamoDB can spread the load evenly across partitions, thus avoiding hotspots and maintaining fast, predictable performance.

Amazon Cognito authorization user pool is used to control access to the Dynamodb data tables. Some private information, such as private sensor locations, private sensor readings, and notification subscriptions are only visible to the owner users. Similarly, only these authorized users can make changes to these data entries. Amazon Cognito manages all the read and write permissions to our database and ensures security.

DynamoDB's reliability is underscored by its data replication feature, which automatically replicates data across multiple AWS Availability Zones. This multi-AZ data distribution ensures high availability and fault tolerance, protecting against service disruptions and data loss.