Flight Planning Using Data Lakes and AI/ML

Use published schedules, flight information, airport weather observations/predictions, historical estimated and actual holding times, and on-route/surface aircraft data to improve hold time prediction, and estimation of departure clearance times.

1. Discover & Catalog
2. Conversion between file formats (CSV, JSON, PARQUET)
3. Propagating changes in incoming data to S3 processed for
4. Building reportable data sets
5. Scheduling and workflow management

Use published schedules of all airlines from OAG to create a full picture of the flights arriving at an airport by the time of day.

Augment the schedule with actual flight data through real-time updated flight information. (This data is limited to the airline currently implementing this solution.)

Use flight plans, surface movement, and on-route information of all the flights from all the airlines (FAA SWIM in US) to create an accurate up-to-date picture of airspace and airport taxiways.

Use weather observations reported through Meteorological Terminal Air Report (METAR) format and weather forecasts reported through Terminal Aerodrome Forecast (TAF) to correlate hold time to adverse weather conditions.

Use historical estimated and actual hold times to build predictive models by airport, time of day, and weather forecasts. Also, use historical data to remove outliers like landing time restrictions.

Tiered data lake architecture using Amazon S3 allows for ingestion and processing of data from a variety of batch and real-time data feeds. In addition, this architecture allows for adding new data feeds and propagating data changes easier.

Use AWS Glue and Amazon EMR to discover, catalog, process inputs, and create the processed data in Amazon S3. These services process batch data to create PARQUET format for all the data.

Use Amazon SageMaker to access, label, build, train, tune, deploy, and manage predictive models to predict hold times and estimated departure clearance times. Use Amazon SageMaker Debugger for facilitating training and tuning of model. Use Amazon SageMaker Model Monitor to detect and remediate concept drift.

Build an operational data store to ingest real-time data and use that data to run a prediction model to adjust predictions based on the latest data. Provide services and events for integration into operational systems like flight planning.