Data Engineering at Rezdy

5 December 2018

A bit about Rezdy

Rezdy started our journey about 7 years ago. We are an online reservation platform **specialising** in tours and activities, through which we currently process up to **1,600** bookings an hour.

We make our suppliers lives easier by providing the tools necessary to scale their operations and save money in streamlining their booking and customer experience.

We also connect their service offering through several channels, providing the online solution to grow their business and reach new customers with our extensive supply and demand network.

Insights team values

rezdy

We build tools

Give a man a fish and you feed him for a day. **Teach a man to fish** and you feed him for a lifetime

rezdy

We are the paved path

Let people use the systems they want to use. However always strive to be the easiest way for people to get what they are after.

We have a single source of truth

Have one place you **know** is accurate.

We build for scale

Short term gains are the **killer** of future progress

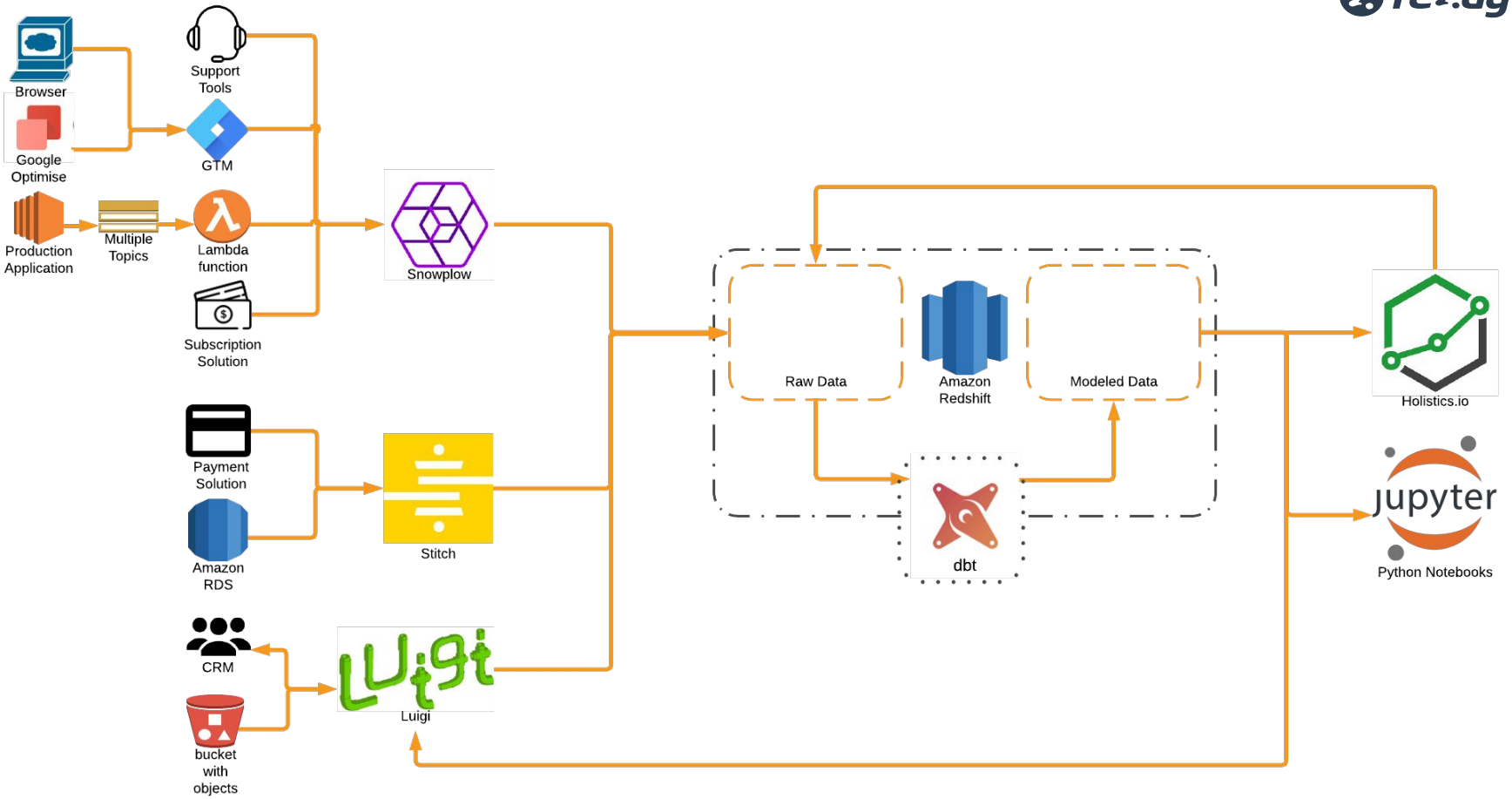
We do the simple things

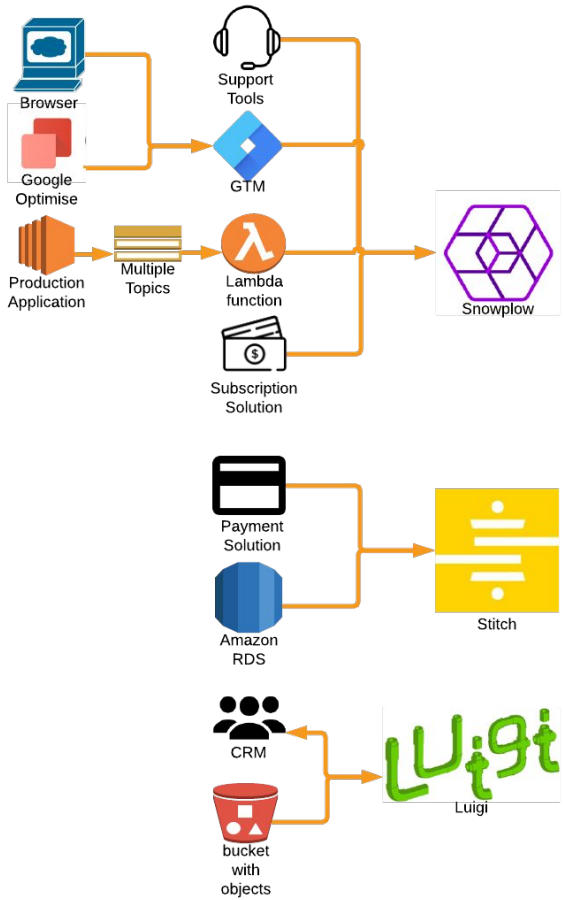
If something requires **20% of the work** to get you 80% of the way then you need to access **how important that extra 20% of the way is.**

We do the important things

Just because something is **simple doesn't mean that it is important.** The most important thing is the most important regardless of effort.

Where did all of that get us


Browser
Google Optimise
Support Tools
GTM
Production Application
Multiple Topics
Lambda function
Subscription Solution
Snowplow
Payment Solution
Amazon RDS
Stitch
CRM
bucket with objects
Luigi
Raw Data
Amazon Redshift
Modeled Data
dbt
Holistics.io
jupyter
Python Notebooks

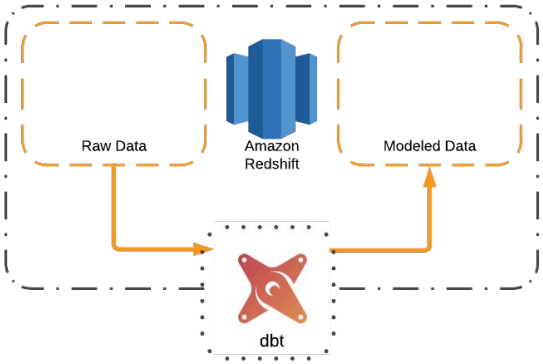


Ingest

3 main Ingest services:

- Snowplow
  - Webhook or event based
- Stitch
  - Bin log replication
  - Plug and play integrations
- Luigi
  - Cron/Batch based workloads

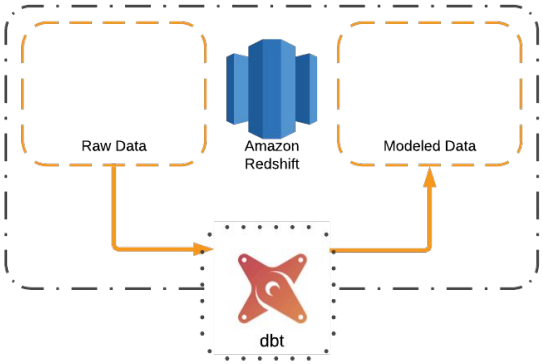
Storage



For long term analytical workloads you really need a columnar database (Commonly referred to as a data warehouse).

We chose Redshift.

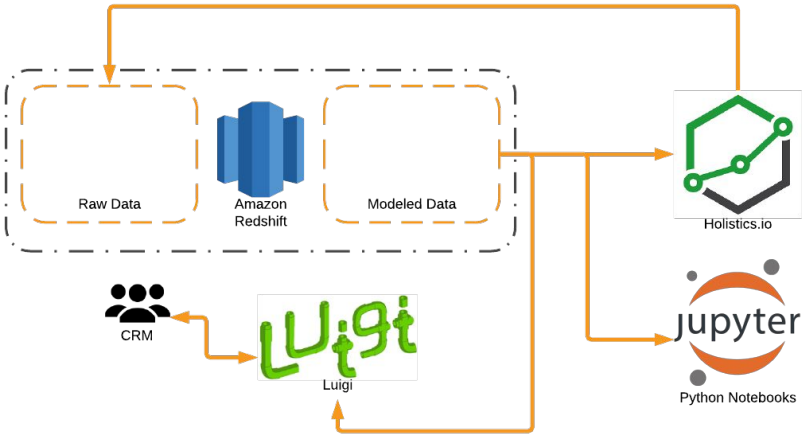
Transform



Any DAG based structure that enforces good engineering practices is a good fit here.

Avoid monolith queries at all costs.

Consumption



3 supported ways we let people access our data:

- Through visualisation in Holistic.io
- Programatically (I primarily use Python)
- Through our CRM which we push data to through Luigi

Thanks!

Any questions?