

REINVENTING HOTEL BOOKINGS

WITH IRONWORKER POWERED ETL

HotelTonight has reinvented the task of finding and booking discounted hotel rooms at travel destinations. Designed for last-minute travel planners and optimized for the mobile era, HotelTonight connects adventure-seeking, impulse travelers with just-in-time available hotel rooms wherever they land.

This model has the market-enhancing effect of reducing excess inventory of unused hotel rooms, while delivering a seamless user experience and deep discount for budget travelers who enjoy impulse travel and adventure.

What most travelers may not realize, is that behind the scenes at HotelTonight lies a massive business intelligence system that uses a sophisticated cloud-based ETL platform that collects, converts, and stores data from multiple external services.

WHAT IS EVENT-DRIVEN ETL?

Extract, Transform, Load (ETL) has been around in IT circles for a long time, dating back to tape storage and the mainframe era. Event-driven ETL makes use of cloud-based services along with loosely-coupled and flexible processing to move data between systems in real-time.

ETL & ASYNC TASK PROCESSING

The key component powering HotelTonight's ETL process, which also allows it to be scalable and event-driven is IronWorker, an asynchronous task-processing service provided by Iron.io. IronWorker provides concurrent workload processing that is faster and more flexible than existing methods.

BENEFITS OF CLOUD ETL

Using cloud-based ETL, HotelTonight has created a fully aggregated dataset that is more accurate, more up-to-date, and more reliable. Turning raw data into knowledge faster than their competition gives their analysts the ability to make quicker updates on hotel room prices and availability.




"The team at Iron.io has been a great partner for us for building our ETL pipeline. Their worker platform gives us a quick and easy mechanism for deploying and managing all our Ruby workers. They have excellent dashboards for reporting worker status, giving us great visibility over the current state of our pipeline."

Harlow Ward, Lead developer at HotelTonight

HOW HOTELTONIGHT'S ETL PIPELINE WORKS

HotelTonight's ETL pipeline gathers external data from multiple sources and brings it together into Amazon Redshift, a managed petabyte-scale data warehouse solution from

Amazon Web Services. Amazon Redshift acts as a “Unified Datastore” and makes use of the SQL query language to connect a variety of platforms using a Postgres Adapter. Custom Ruby scripts power the HotelTonight ETL process, connecting their Business Intelligence team to an SQL Workbench interface which front-ends their Amazon Redshift clusters. The dashboard lets anyone in the organization query the data and extract information for use in their initiatives.

DISTRIBUTED ETL WORKFLOW

HotelTonight deploys a custom worker for each external data source. This loosely coupled approach means that aggregation of each data source is independent of other sources.

“IronWorker’s modularity allows for persistent points along the lifetime of the pipeline. It also allows us to isolate failures and more easily recover should data integrity issues arise.” according to Harlow Ward, lead developer at HotelTonight. “Each worker is responsible for its own unit of work and can kick off the next task in the pipeline.”

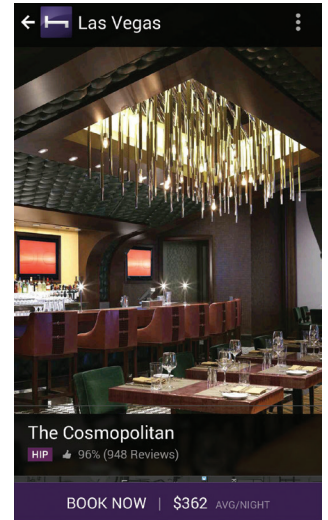
MONITORING AND ORCHESTRATION

In addition to solving the challenge of quick and easy deployment of independent workers, the Iron.io dashboard provides current status of tasks to HotelTonight developers, giving them instant visibility and insight on the state of their ETL pipeline. Users can control settings for the workflow, including increasing or decreasing concurrency, retrying tasks that may have failed in prior attempts, or modifying job schedules.

NET RESULTS

Now that HotelTonight’s business intelligence data is consolidated within Amazon Redshift, HotelTonight can run SQL queries to combine and correlate data from multiple platforms into a unified dataset. Prior to this solution, HotelTonight’s data analytics consisted of multiple exported CSVs from each data source, merged into a single pivot table with lots of “magic” applied to make sense of it.

IronWorker makes it possible for HotelTonight to streamline and automate their entire ETL process and bring together all their disparate data sources into a single flexible datastore. HotelTonight can rest easy with the assurance that in using IronWorker, their data pipeline into Amazon Redshift is in excellent order and can be updated or amended at will.



Turning raw data into knowledge faster than competition gives HotelTonight’s analysts the ability to make quicker updates on hotel room prices and availability.



Behind the scenes at HotelTonight lies a massive business intelligence system that uses a sophisticated cloud-based ETL platform that collects, converts, and stores data from multiple external services.