



CASE STUDY

Hotels.com Optimizes the Booking Experience with Big Data, Migrating Data Sets to Apache® Hadoop®

Hotels.com® provides a service for reserving hotels, B&Bs and other types of commercial lodgings, helping millions of customers around the globe. It offers 90 localised websites in 41 languages and hundreds of thousands of places to stay in more than 200 countries and territories.

With millions of consumers searching for the ideal hotel room around the world, Hotels.com was finding it difficult to cope with the vast amounts of data coming from its search engine marketing that could be used to enhance the customer journey. With keywords drawing customers in, spreading across a variety of languages, not to mention a large number of devices and potential visitor journeys, it was looking for a technology partner able to deliver data analysis at scale.

PROCESSING DATA AT SCALE

Operating a website on a global scale inevitably brings the problem of user diversity. Each website visitor will have different criteria when it comes to factors such as location, price, and facilities—and can browse hundreds of hotels before booking a room. Hotels.com was looking for a solution that could help customers narrow the gap between query and booking stages, as well as establishing the technology foundation to run multiple data analysis use cases for the organization.

Hotels.com needed to bring together data reaching across both consumer preferences and consumer habits, which would come from a variety of sources. Customer profiles, search criteria, location, device and website behavior are all factors that could generate the valuable data needed for a data lake. If analyzed and cross referenced, these

KEY HIGHLIGHTS

- Migrating data warehousing from relational databases to Hadoop
- Data processing at scale
- Real-time data analytics providing instant insights and benefits for customer experience
- Aggregating customer data from regional website sources to draw universal insights
- Ability to discover, explore and share various datasets to generate valuable insights across the entire organization

datasets can help generate the insight needed to help customers find the hotel they're looking for as soon as possible, vastly reducing time spent searching, filtering and browsing.

It became clear that traditional enterprise data warehousing was no longer suitable to sustain the company in its data journey. The need to extract insights from large datasets required a massive amount of processing and standardization, even to run basic Extract, Transform and Load (ETL) functions. Additionally, data-driven insights based on keyword recognition—a foundational aspect of search engine marketing (SEM)—required a broad system of activation keywords to pick up searches from countries speaking different languages.

TURNING ASSUMPTIONS INTO DATA-DRIVEN INSIGHTS

As a data-driven organization, Hotels.com was unwilling make critical business decisions based on hunches and assumptions. With its former relational data warehouse unable to manage the sheer scale of the data, it decided to move its data management into the scalable world of [Apache Hadoop](#), benefitting primarily from the flexibility, scalability, and openness of the platform.

It also needed a technology partner who could provide the necessary enterprise support, while still maintaining the flexibility to innovate on top of the platform to avoid vendor lock-in. Hortonworks' commitment to the open source community allowed it to stand out against competitors by empowering its customers to utilize a wide variety of open source solutions from one single platform.

Hotels.com deployed a [Hortonworks Data Platform \(HDP®\)](#) cluster on premise to be able to access the data on a granular scale, enabling IT administrators to analyze different types of traffic, different kinds of customers along with the type of devices they use. By using Hortonworks Data Platform as the primary data warehousing solution, it has been able to discover, explore and share various datasets to ultimately generate valuable insights across the entire organization.

"Our back-end data processing involves analyzing the data, joining it, and putting it together to deliver a really good booking experience" said Adrian Woodhead, Principal Engineer, Hotels.com brand.

"HDP is a rich platform, it's a very big ecosystem and it gives us the ability to innovate. There's a wide range of tools on the platform, so it gives our developers lots of choice when it comes to choosing what technologies can help solve their problem" commented Elliot West, Principal Systems Engineer, Hotels.com brand.

THE BOOKING EXPERIENCE OF THE FUTURE

Almost all data processing at scale has been moved over to the Hotels.com internal Hortonworks Data Platform cluster. This has generated an open ecosystem for data analysis that fellow analysts and data scientists can now use across the entire organization, to address any number of problems and draw out efficiencies with clear data reasoning.

Hotels.com has gone from analyzing data queries daily to multiple times a day, with the ultimate objective of eventually running a real-time data streaming ingestion platform. For example, Hortonworks Data Platform enabled the teams to map clicks and reconcile them according to the various marketing channels to assess the effectiveness of a marketing campaign and in future this will inform a live streaming ingestion platform to perform clickstream analysis in real time. The data libraries will also likely be used in many other use cases.

With Hortonworks, Hotels.com has been able to lay the foundation for its development teams to generate value from large datasets. Having started with an initial use case, the company has now grown its commitment to big data analytics to develop hundreds of other applications to support mission-critical objectives across other departments. By bringing a variety of complex datasets into a data lake, Hotels.com is now able to accurately map a profile of a customer and provide the best booking experience accordingly.

"Our back-end data processing involves analyzing the data, joining it, and putting it together to deliver a really good booking experience."

Adrian Woodhead
Principal Engineer,
Hotels.com

"HDP is a rich platform, it's a very big ecosystem and it gives us the ability to innovate. There's a wide range of tools on the platform, so it gives our developers lots of choice when it comes to choosing what technologies can help solve their problem"

Elliot West
Principal Systems Engineer,
Hotels.com

About Hortonworks

Hortonworks is a leading provider of enterprise-grade, global data management platforms, services and solutions that deliver actionable intelligence from any type of data for over half of the fortune 100. Hortonworks is committed to driving innovation in open source communities, providing unique value to enterprise customers. Along with its partners, hortonworks provides technology, expertise and support so that enterprise customers can adopt a modern data architecture. For more information, visit hortonworks.com.

The information in this document is provided "as is" without any warranty, express or implied, including without any warranties of merchantability, fitness for a particular purpose and any warranty or condition of non-infringement.

Apache, Apache Hadoop, and the Apache Hadoop logo are registered trademarks or trademarks of the Apache Software Foundation in the U.S. and/or other countries.

Contact

For further information, visit hortonworks.com

+1 408 675-0983
+1 855 8-HORTON
INTL: +44 (0) 20 3826 1405

