

Priceline Bolsters its Recommendation Engine and Reduces Data Storage Costs by 5X-10X with Starburst

priceline.com®

The global travel brand Priceline recently launched a major data transformation initiative to leverage vast amounts of streaming and historical datasets stored across different cloud and on-prem systems. The company wanted to democratize data access and give its users a high-performance, cost-effective way to transform all that raw information into novel insights.

The solution Priceline envisioned would need to provide the necessary security guardrails, maintain agility and flexibility, and support the company's move to a decentralized Data Mesh infrastructure. After evaluating its options, Priceline selected Starburst Enterprise.

Background

Priceline was founded in 1997 with ten people, a small office, and a brilliant idea: What if you could fill empty hotel rooms and airplane seats by harnessing the power of the Internet and allowing customers to name their own price?

Today, Priceline is part of the world's largest travel company, Bookings.com, and it helps people save more than \$1B per year. Given its prominence, Priceline remains a lean operation, with only 1,000 employees. The company emphasizes diversity and inclusion, and promotes a smart, focused work culture. Priceline also leverages data science and analytics to deliver personalized recommendations to the millions of travelers who visit its sites each month.

Today, its data is stored in multiple systems:

- Oracle RDBMS on-prem warehouses
- Google Cloud Storage (GCS) data lake
- Google BigQuery
- CloudSQL instances in Google Cloud Platform (GCP)

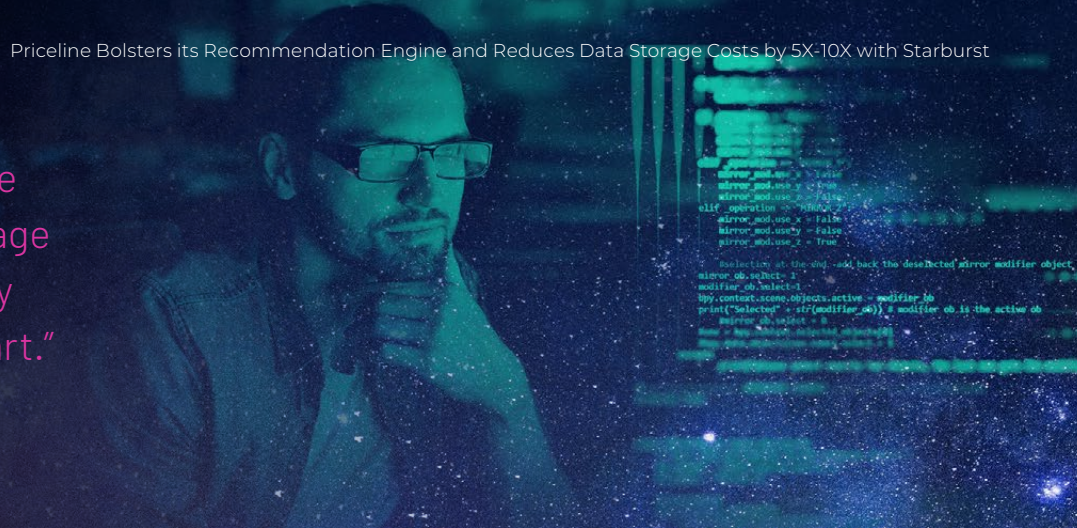
Generally, users run ad hoc queries or generate reports and dashboards through Tableau. Although the company has already demonstrated its success turning data into profit, the Data team decided they could do even better. "Historically, it is only the curated, transformed, refined data stored in warehouses that people get access to," notes Priceline's Senior Director of Data Sachin Gopalakrishna Menon. "But that only provides one view of what the user is asking for - it doesn't give them the larger view of the data. We wanted to democratize access to our raw data, including historical data sets in various unstructured formats, and we wanted our users to be able to access them using queries."

Furthermore, Priceline wanted a solution that would be secure, low-cost, and fast. "Our core intention was to enable our end users to use the raw data we have in a much faster way, rather than waiting for developers or data engineers to provide it," Menon notes. "That is how we finally decided on Starburst."



“The ability to access the data in Google Cloud Storage is where Starburst really stands out and stands apart.”

Sachin Gopalakrishna Menon
Senior Director of Data, Priceline



Priceline + Starburst Enterprise

Today, Starburst Enterprise serves as the secure, high-performance analytics engine and Data Mesh enabler at Priceline. Plus, as Priceline migrates more data into its data lake, Starburst will maintain access to data, minimizing the disruption of analysts. Menon highlights several features and advantages.



Federated Access

Starburst functions as a federated query layer that lets users tap into multiple heterogeneous data sources. Production-tested, high-performance connectors to Oracle RDBMS, cloud data lakes, BigQuery and other systems gives users the wider view of data they'd been waiting for. “Starburst taps into this data for analytics purposes and helps us join these raw datasets and refined datasets, plus metadata and reference datasets in CloudSQL instances in GCP,” Menon notes.



Semantic Layer

With Starburst Enterprise, users don't need to know where or in what format data resides. This complexity is abstracted away, allowing them to focus on what they do best. “We don't want the users to be confused about where data is coming from,” says Menon. “The semantic layer takes care of all that.”



Security & Access Controls

Starburst allows Menon and his team to apply fine-grained access controls to unstructured datasets – a capability the company lacked previously – and strengthens security with regards to their distributed datasets.



Cloud Storage Savings

Priceline can already run fast queries on data in BigQuery, but data has to reside there in a structured format, which drives up storage and compute costs. Starburst separates storage and compute, so Priceline can move a significant volume of that data into an unstructured format and reduce its storage costs by 5X to 10X – without surrendering access.



Performance vs. Cost

Reducing the time from question to insight is essential, so Priceline wanted a query engine that delivers great performance. Starburst Enterprise satisfies that metric and strikes an efficient balance between cost and speed. Noting the Google scenario above, Menon adds, “We may not hit the exact same performance numbers we would have hit in BigQuery, but we'll reduce the storage cost at least by 5 to 10 times, and performance will be almost that same level.”



GCP Integration

Priceline is already able to query data in BigQuery, but Starburst extends this capability to other Google datasets. “The ability to access the data in Google Cloud Storage is where Starburst really stands out and stands apart,” Menon adds.

Business Impact

The Data team's ultimate goal is to increase the success of the business. On that front, Starburst Enterprise is poised to have an impact in several areas.

Enhanced Recommendations

In the short term, Priceline will be able to tap into more historical data when generating recommendations for shoppers. "If someone is looking for a rental car, we could look at their historical preferences, and the cars they booked in the past, and recommend the appropriate type, which will increase success and customer satisfaction," Menon says.

Exploring the Unknown

One of the goals of Starburst is to give data consumers the freedom to be curious, and Priceline is excited about this possibility. "The ability to tap into the unknown and come up with recommendations that will change how we do business and how we provide capabilities to our customers is going to be a big, big change," adds Menon.

Faster Time-to-Insight

Previously, data processing required ETL, so by the time Priceline had the data to make a decision, at least a day or two had passed. Now, Priceline is accelerating the timeline. "We can make decisions faster based on the analytics," Menon says. "Instead of waiting for days, this happens in near real-time."

Future Outlook: Decentralized Data Mesh and Exploration

In the long run, Menon envisions Starburst as a key piece of his company's Data Mesh infrastructure and overall data transformation initiative. "The ability to democratize data and make informed decisions with datasets that we have never explored will be a significant game changer," Menon concludes. "We're on a journey to democratize data as much as possible, because there's so much we don't know and so many elements we have not tapped into. With Starburst, there's so much more we can explore to drive decisions and insights."



"The ability to tap into the unknown and come up with recommendations that will change how we do business and how we provide capabilities to our customers is going to be a big, big change."

Sachin Gopalakrishna Menon
Senior Director of Data, Priceline

Unlock the value of your data with Starburst Enterprise today.
For more information, contact us at starburst.io.

