



About Couchsurfing

Couchsurfing, established in 2004, is a community based social travel network focused on creating real and meaningful travel experiences.

Over the past 10 years, Couchsurfing grew to be a global social network of over 10 million travelers in 120,000+ cities. How it works is, say you are planning a trip to Paris. Using Couchsurfing, you can either search for individuals that look interesting and directly ask them to host you. Or you can publicize your trip and have people offer to host you. Stays are generally a few days, include hanging out and being shown around by your host, and are free. In addition to travelling, over 100,000 members attend social events using the Couchsurfing platform.

“Demand for our service continues to grow but the original architecture wasn’t allowing us to roll out new features to the community.”

Jim Nist, Couchsurfing’s VP of Engineering

“The Problem” – Relying on Mythology, Not Data

Couchsurfing has a unique technological history of being “community built” and evolved over the past 10 years. What began as a small company had grown much larger than their infrastructure was designed to handle: hundreds of users had become thousands and then millions.

While the original site has great features and a large user base, the original architecture was built without consideration for data analysis.

The challenges:

- The Product Group was unable to analyze user behavior and was using mythology instead of data to drive product decisions

- A large amount of application data spread over 46 MySQL databases on 25 different servers making analysis difficult, if not impossible
- The environment was plagued with unacceptable delays, query timeouts and dirty data

“Our data analysts were hamstrung. If analysis was possible, the majority of their time was spent locating and downloading data for offline analysis. Now all data is accessible in one online datastore, streamlining analysis.”

Jim Nist, Couchsurfing's VP of Engineering



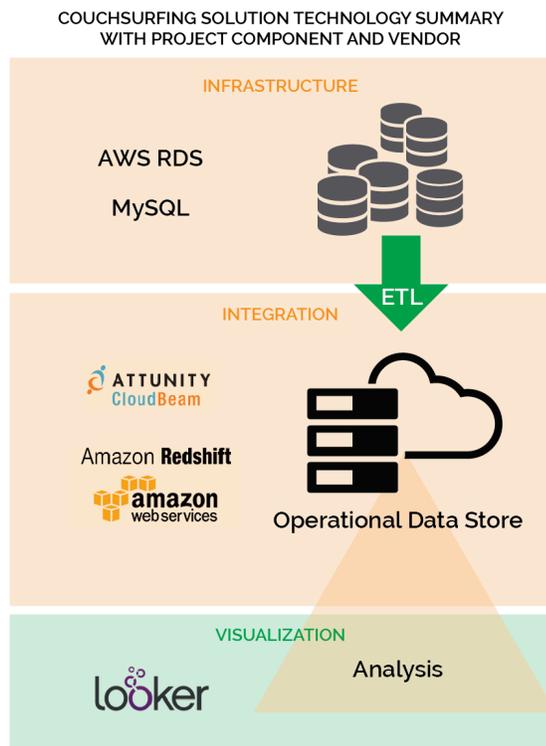
Couchsurfing was looking to re-architect their entire platform under the guidance of a new team led by Jen Billock, CEO and Jim Nist, VP of Engineering. They had a vibrant and loyal user base but they had also inherited a lot of technical debt.

The new management team knew that to be successful re-architecting the platform they

needed data insight into their current user population, user engagement analysis and historical data trends. All of this data existed but there was no way to access it.

“The Solution” – Real-time Actionable Insight

The Solution Architects at Bytecode IO worked with Couchsurfing management to create a scalable Data Intelligence solution to deliver actionable data insight quickly and overcome the limitations of the original data architecture.



The decision was made to combine all the disparate application data sources into a centralized operational data store (ODS). Couchsurfing was already using AWS RDS as their transactional database platform so AWS Redshift was a natural choice for a data warehousing solution. Additionally, Couchsurfing could continue to work within the familiar AWS ecosystem. AWS Redshift is very inexpensive to get up and running as it is offered as a service and can be implemented and maintained with little operational overhead.

Using AWS Redshift, Bytecode IO successfully built a Level One ODS, which is a real-time transactional copy of production systems.

For the movement of the original production data, Bytecode IO researched emerging technologies via the AWS Marketplace and decided on using Attunity CloudBeam as the ETL platform. Attunity CloudBeam does end-to-end database loading from RDS MySQL into AWS Redshift. Using traditional ETL methodologies would have taken months to develop, but with Attunity the development time was reduced to hours. Attunity CloudBeam also replicates data into AWS Redshift in near real-time by performing Change Data Capture (CDC) on the RDS MySQL binlog. In other words, the

46 MySQL databases would be integrated into a single system and production data would be replicated in near real-time into the ODS.

Finally, Looker was chosen for the visualization and analysis layer. Using Looker's data modeling language, LookML, the development cycle to get from "having our data centralized" to "able to explore it" was minimized. Information was visualized within minutes for business consumption at all levels.

"A month into the Bytecode IO's implementation of our new data intelligence solution, things changed for us. Data started to drive our product meetings and accelerate our development cycle with informed decision making. Meetings now start with our target KPIs presented with Looker up on the screen, then we dig in."

Jim Nist, Couchsurfing's VP of Engineering

"The Results" – Relying on Data, Not Mythology

Couchsurfing has been given the ability to analyze their users behavior and anticipate their needs.

Before implementing an Intelligence Architecture that would centralize all of Couchsurfing's operational data, there was no way to answer questions beyond hunches and theory. After implementation, actionable analysis was driving business decisions.

"Our data analysts were hamstrung. If analysis was possible, the majority of their time was spent locating and downloading data for offline analysis. Now all data is accessible in one online datastore, streamlining analysis." Jim Nist, Couchsurfing's VP of Engineering.

In addition, the gained insight into the historical data has allowed Couchsurfing to design and re-launch their product in a scalable responsive fashion. See link: <http://blog.couchsurfing.com/whats-changed-on-couchsurfing>

The scalable architecture of AWS Redshift combined with the flexible nature of Looker as a data modeling and visualization tool has Couchsurfing up to full speed facilitating travel around the world. Management is seeing information in a way that months before wasn't possible.

Couchsurfing is now relying on data instead of mythology.

Technologies List



Attunity CloudBeam enables organizations to simplify, accelerate, and automate data transfer to, from, and across AWS regions.

<http://www.attunity.com/products/attunity-cloudbeam>



Looker lets organizations access, control, describe, explore, and see their data.

<http://www.looker.com/product>

Amazon Redshift



Amazon Redshift is a fast, fully managed, petabyte-scale data warehouse solution that makes it simple and cost-effective to efficiently analyze all your data using your existing business intelligence tools.

<http://aws.amazon.com/redshift>

About Bytecode IO

Bytecode IO designs, builds and supports BI, data warehousing and data integration solutions. We believe the best business decisions come from clearly visualized information. We are determined to help our clients make the best use of their data, whether it comes from a single source or from dozens, and confidently guide their business through an ever-changing future.



1855 Olympic Blvd. Ste. 210
Walnut Creek, CA 94596

+1 (415) 704-3230
info@bytecode.io
www.bytecode.io