dremio

CASE STUDY

InCrowd Sports Uses Dremio to Provide Clients with Deeper Insights to Create Personalized Fan Experiences

CUSTOMER
INCROWD
incrowdsports.com

GEO
United Kingdom

INDUSTRY
Sports/Entertainment

OBJECTIVES
Create fast, secure access to a single 360° customer view across multiple data sources, offering insights and personalized experiences for their fans.

CHALLENGES
- Data warehouse did not provide the flexibility or performance to add new data sources
- Governance/data lineage was a tedious, manual process
- It took hours to produce aggregate views of data through batch jobs
- Unable to provide a singular view of a fan's interactions with the existing platform

SOLUTION/DATA ENVIRONMENTS
The elements of the solution include:
- **Primary Data Source: AWS S3**
- **Secondary Data Sources: MongoDB, Redshift, PostgreSQL & MySQL, CSVs**
- **Compute: Dremio**
- **BI Tools: Tableau**

RESULTS
- Increased revenue and loyalty through personalized fan experiences
- Flexible access to new data sources increased competitive advantage
- Fast data access means more actionable insights
- Data governance provides security and control
- Extended direct self-service data access to clients
- Lowered the cost structure for data analytics by reducing dependency on data warehouse

Summary

InCrowd Sports uses Dremio to create a single 360° customer view across multiple data sources so they can offer their sports clients deeper insights and personalized fan experiences.

The Challenge

InCrowd recognized that their data warehouse was too slow and inflexible to meet the needs of the evolving and increasingly demanding requirements of their clients. The system prevented InCrowd from providing fast enough access to fan insights as it took a significant amount of time to run batch jobs to get aggregate views of data. Inevitably this meant that it was becoming more and more difficult to meet client expectations for immediate insights.

Their data warehouse also did not provide the flexibility to easily add new data sources like clickstream data to provide deeper insights into their fans. They could not provide clients with an accurate, real-time view of all of a fan's interactions with the platform, resulting in missed opportunities for fan engagement. It was also extremely time-consuming and labor-intensive to keep data secure because tracking data lineage was a tedious manual process.

To be competitive, InCrowd knew they needed a new approach. They decided to leverage an open data lake architecture because they didn't want to be constrained by (and locked into) a proprietary data warehouse architecture, so they started looking at data lakecentric solutions.

The Solution

The new solution is an open data lake using AWS S3 data lake storage, with their customer analytics powered by Dremio. "We were quite lucky as we stumbled into Dremio relatively early. It really was a light bulb moment," says Ciaran Fisher, InCrowd Chief Technology Officer. "The technology got us excited, so we brought Dremio Community Edition in to test it for ourselves. The technology started delivering tremendous value almost immediately, and it is now a key part of the future of InCrowd."

InCrowd can directly access, curate and query the majority of their data held in AWS S3 as well as Redshift. This data includes reference data and CSVs from clients' own data sets or exports from their own systems. InCrowd can join that data with additional sources such as:
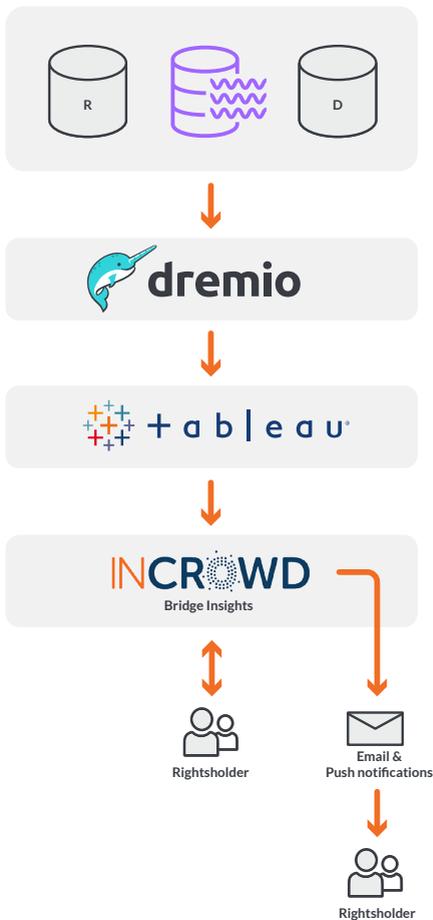
- **Clickstream usage data in AWS Redshift from their analytics tool Snowplow**
- **Structured PostgreSQL database on AWS containing datasets from different Python ETL scripts running on Kubernetes**
- **Ticketing data held in a PostgreSQL database from live feed integrations with ticketing partners, as well as from manual CSV imports**

Each of these previously siloed data sources is represented inside Dremio as a physical data set. These physical data sets are then joined together within Dremio as a virtual dataset, enabling InCrowd's hosted instance of Tableau Server to connect directly to Dremio.

By joining the datasets in Dremio, InCrowd can develop 360° fan profiles to power personalized email, push and other communication tools. Exposing only the Dremio-based virtual datasets within Tableau allows data analysts to quickly understand the data structure and produce relevant dashboards and interactive insights.

After InCrowd used Dremio Community Edition for about ten months, they were confident that it worked for them and then upgraded to Dremio Enterprise Edition to take advantage of the product's enterprise-class security features and support.

The Results

Dremio enables InCrowd to provide their clients with faster, more secure access to more data—improving insights and fan experiences.

INCREASING REVENUE AND LOYALTY THROUGH PERSONALIZED FAN EXPERIENCES

InCrowd Sports' clients can now get one 360° view of how their fans are engaging, interacting and spending across all digital and transactional touch points. They get personalized analytics so they can provide a better fan experience and make targeted offers to increase revenue spend.

Several clients utilize Android and iOS mobile apps developed by InCrowd for their fans. InCrowd uses clickstream usage data from these mobile apps to create usage dashboards in Tableau, hosted in their client-facing insights tool called Bridge. They are able to join the clickstream usage data from AWS Redshift with information such as football match data from Opta to provide additional end-user insights. Dremio's high-performance data lake query engine, coupled with transparent acceleration features, ensures the underlying datasets can be served to all of their clients with sub-second response vs. minutes or even hours.

> ❝
>
> **We can get more actionable insights from data. Since the processing speeds are so much faster, we can ask more questions and obtain insights quickly.**
>
> **CIARAN FISHER**
> **INCROWD CTO**

> ## We've used Dremio in our environment for almost a year now and see it as a vital part of our infrastructure as we take the next step forward as a business.
>
> **CIARAN FISHER**
> **INCROWD CTO**

FLEXIBLE ACCESS TO NEW DATA SOURCES INCREASES COMPETITIVE ADVANTAGE

Dremio enables InCrowd to be more flexible and agile in how they leverage new data sources and bring them to life with Tableau. Their previous data warehouse could not flexibly access new data sources to provide deeper insights. "We would have to amend the structure of a database table in order to support new data or changes to existing data," says Fisher. Dremio's support for Tableau's native data source format (TDS) makes it easy to create and then publish data sources. Their hosted instance of Tableau Server is only connected to Dremio, so they have a single integration to maintain and secure.

FAST DATA ACCESS MEANS MORE ACTIONABLE INSIGHTS

Previously, InCrowd had been running batch jobs with AWS Redshift to produce aggregate views of data to improve performance for Tableau users. This took hours to run and impacted other systems importing data into Redshift. It also required Tableau users to know which views to connect to, creating a complex, brittle environment.

With Dremio, InCrowd can join data sources with sub-second performance. Dremio uses a range of high-performance, columnar query execution technologies along with transparent query acceleration so users can keep pointing to the same virtual datasets, even as changes are made and updates occur to those datasets.

Dremio's ability to process large datasets with live, interactive performance also means faster insights for InCrowd clients. "We can get more actionable insight from data. Since the processing speeds are so much faster, we can ask more questions and obtain insights quickly," Fisher says.

With Dremio, InCrowd can also join and analyze data in real time from data sources outside the data lake, including OracleSQL, Teradata and MongoDB. This enables users to connect directly to the data and begin their exploration and analysis. For example, users can join large CSV-based data sets in the AWS S3 data lake with data from operational sources such as Oracle and Microsoft SQL Server and provide immediate SQL access via Looker, Databricks, or other client tools.

This ability to create virtual data sets that are mapped and joined against multiple physical data sets provides InCrowd with the flexibility to transparently modernize their underlying data infrastructure and bring more data directly into the data lake storage environment. They are able to provide consistently fast access without impacting end user analysts during the modernization process.

DATA GOVERNANCE PROVIDES SECURITY AND CONTROL OVER DATA ACCESS

In the past, governance and data lineage was a manual process that required a lot of documentation on the views and aggregations in order to show what data came from where. Dremio's data lineage ensures InCrowd knows what data is going where and who should and should not see it.

Dremio's IT-governed, self-service semantic layer provides InCrowd with controls that define how clients interact with the data that InCrowd is collecting for them, as well as the customer and InCrowd's datasets. With this added governance, InCrowd can ensure they remain compliant with all relevant legislation and "do right" by the fan.

EXTENDING DIRECT SELF-SERVICE DATA ACCESS TO CLIENTS

InCrowd plans to provide more sophisticated customer teams self-service access to raw and aggregate data within the Dremio user interface. As they continue to scale out their business and environment, the company plans to allow clients to access their own data with strong data governance rules and faster response times from their dashboards.

LOWERING THE COST STRUCTURE FOR DATA ANALYTICS

InCrowd's CTO has been very satisfied with the Dremio solution and how it has evolved. "We've used Dremio in our environment for almost a year now and see it as a vital part of our infrastructure as we take the next step forward as a business."

Long term, InCrowd will also benefit from significantly reduced data analytics costs. They intend to decommission their old analytics databases as well as transfer a significant volume of data out from Redshift to low-cost S3 storage. While InCrowd intends to leverage the Redshift data warehouse for certain use cases, it will store less than 10% of their data, significantly reducing InCrowd's costs.

dremio

ABOUT DREMIO

Dremio is a fundamentally new approach to data.

We started Dremio to shatter a 30-year-old paradigm that holds virtually every company back. Removing barriers, accelerating time to insight, putting control in the hands of the user. That's Dremio.

Deploy Dremio →

Learn more at dremio.com

CONTACT SALES
contact@dremio.com

dremio.com