

PUMEX

Data ETL Development and Ongoing Synch Process

International Association of Ironworkers

CASE STUDY



About the International Association of Ironworkers

The International Association of Ironworkers (IAI) is a union helping to secure stronger benefits, better pay, safer working conditions, and a higher level of job security to its 120,000 members across North America. Their members have been involved in major construction projects such as the Golden Gate Bridge, the World Trade Center, and the Freedom Tower.



Executive Summary

IAI needed the data on their legacy systems database (DB) to be synced with their partnership forum: IMPACT and their Salesforce instance. This data contained an initial 2.5 Million records with about 10,000 weekly updates, and the data sync needed to maintain all relationships within the Salesforce objects.

After an unsuccessful implementation of Heroku due to capability gaps and high costs, IAI reached out to Pumex to develop a custom solution.



The Challenge

IAI's Heroku implementation was not able to handle the initial data sync of 2.5M records and the ongoing update of 10K records per week. Source member data was incorrect and didn't have unique identifiers, there was Unicode within text, incorrect country/state values, and invalid date ranges within the data. Important member data was spread across multiple legacy databases and did not utilize a common primary key. The sync to the IMPACT Salesforce forum required that member certifications and course relationships be maintained, members correctly map to the appropriate Union locations, and members are located under the correct District Councils.



The Solution

Pumex implemented a client side ETL layer to extract, load, and store the data into an MSSQL Server 2016 intermediary DB prior to syncing with Salesforce. To manage the business rules, we used a .NET Core application. This DB updates every 10 minutes to ensure all data is properly synced on an ongoing basis. We used Salesforce's Streaming API to associate ID's with the correct record, and Salesforce's API to sync all 2.5M records into Salesforce.



Immediate Results

- **Data sync with refresh** – an initial sync of 2.5M records from IAI's legacy systems and refreshed any changes in the data every 10 minutes.
- **Data records integrity** – all updates and deletes to any records were properly managed and correctly synced with Salesforce.
- **Low sync error rate** – we achieved a sync error rate of .0008% (~2,000/2.5M records)

Future Plans & Projected Results

Pumex developed an additional "Gold Source DB" for IAI. All data will be migrated here once IAI decommissions all of their legacy systems. Over a 3-year period, IAI has projected that our Unity Data Sync work will reduce their ongoing data management operational costs by 40%.

