



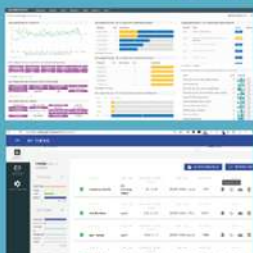
Apache HAWQ integration with Leading Data Privacy platform



The JDBC driver returns the requested data for the scan

Connector App request the data to datasource and make datasource call through JDBC driver

PII data Scanning Application



Connector App request the data and process it to customer platform



Connector App

Customer triggers the call to initiate data scan on Apache HAWQ



Customer

Customer is a leading Personal Data Privacy and Protection provider.

It enables advanced machine learning and identity intelligence to help enterprises better protect their customer and employee data at petabyte scale.

It identifies all PII across structured, unstructured, cloud & Big Data.



Requirement

Customer requested for a connector app to integrate their platform with Apache HAWQ. Connector app will parse data from data store and normalize it in the required format.



Technology Solution

- Format Apache HAWQ is an advanced enterprise SQL on a query engine and analytic database. It combines the key technological advantages of MPP database with the scalability and convenience of Apache Hadoop
- Sacumen developed the connector app to integrate HAWQ using java. The connector app performs the following actions:
 - Set up the prerequisites
 - Amazon Centos instance with apache hadoop file system install on that along with postgresQL 8
 - Authenticate is done by PostgreSQL database user, which will be created while configured while installing postgresql
 - Tables and schemas data are parsed, and Normalized

