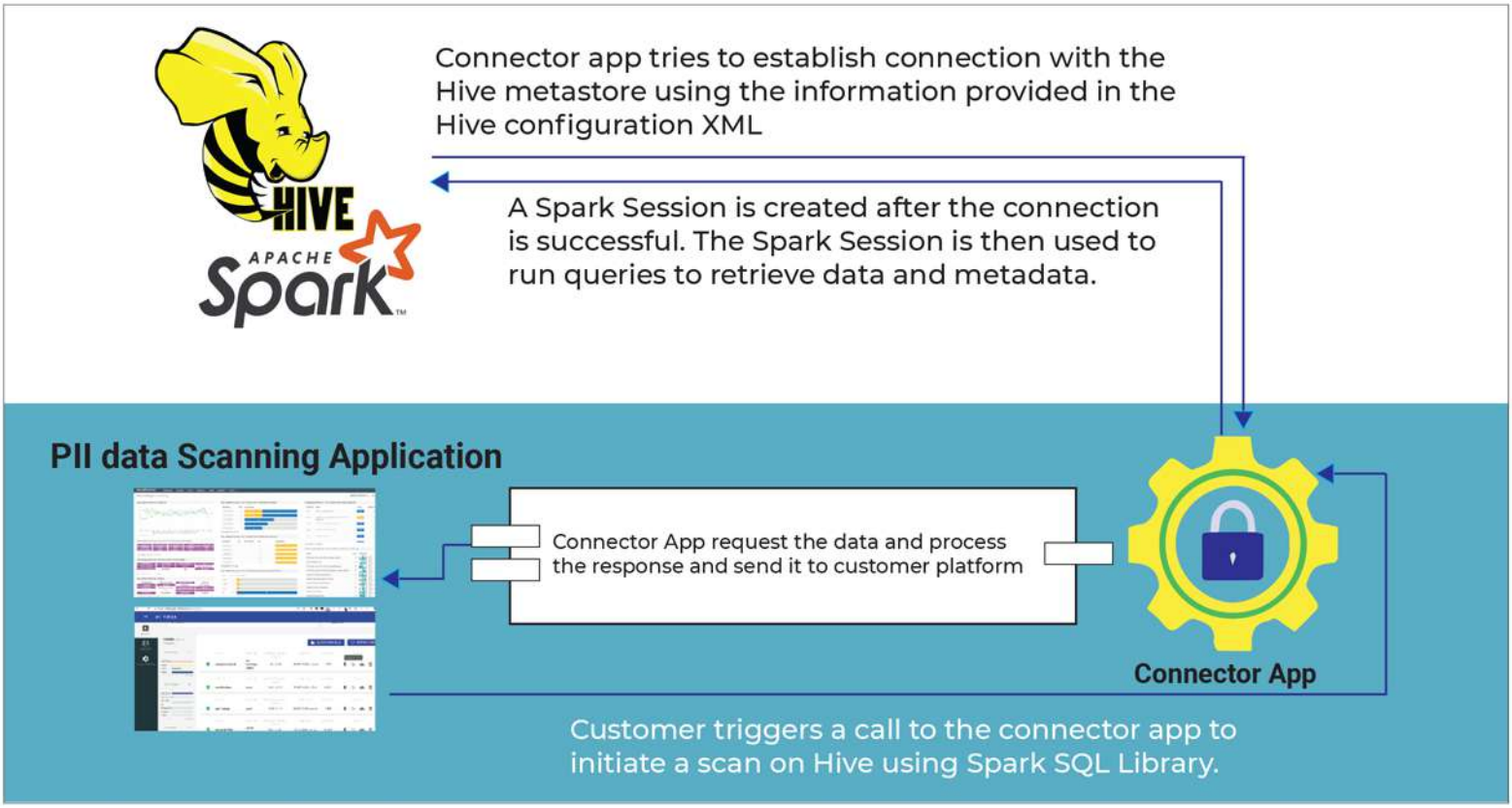
SparkSQL integration with Leading Data Privacy platform



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.

Customer demanded a connector app to integrate their platform with SparkSQL. Connector app will parse data from SparkSQL and normalize it in the required format.

Requirement

Technology Solution

- Spark SQL brings native support for SQL to Spark and streamlines the process of querying data stored both in RDDs (Spark's distributed datasets) and in external sources. Spark SQL conveniently blurs the lines between RDDs and relational tables.

  Sacumen developed the connector app to integrate Spark SQL with Hive using java. The connector app performs the following actions:
- Set up the prerequisites
  - Install Hive and setup a metastore
  - Create databases and tables in Hive
  - Configure Hive settings using the hive-site.xml
  - Use the custom parameters in the DS form to test connection

  Scan the schema/ tables and normalize the data.

SACUMEN Security Execution Partner

Learn More : www.sacumen.com/services/connector-development

www.sacumen.com

Sacumen is an award winning pure play security product development services company. We are engaged with 50+ security product companies such as Symantec, Palo Alto Networks, Varonis, AlienVault, IBM, CA Technologies, ThreatConnect, SecurityScorecard, ForgeRock, Code42, BigID, Flashpoint etc in the areas of Connector Development, Connector Support and Product Engineering.

info@sacumen.com

+1 408 215 1253