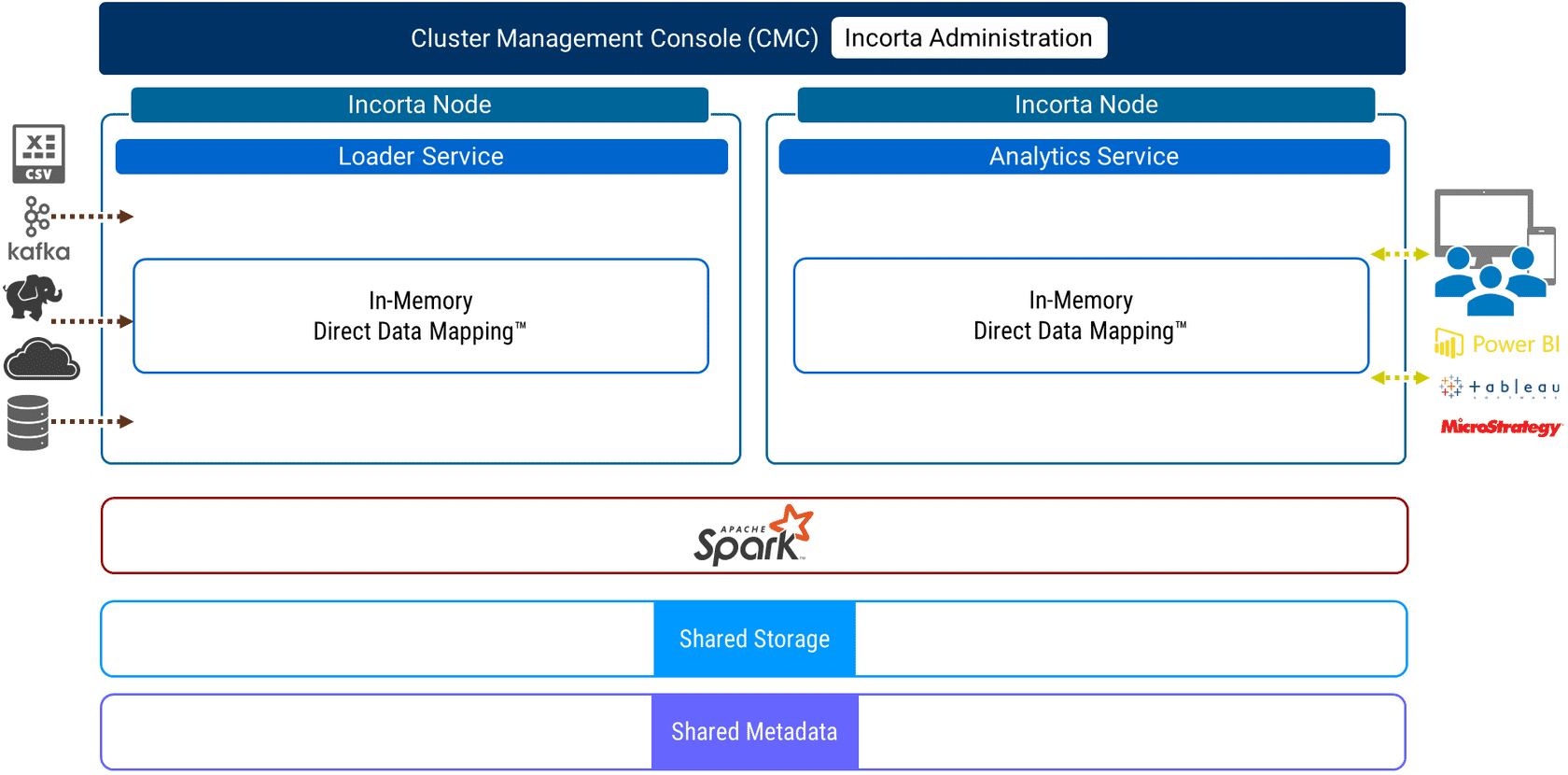
**Incorta Architecture**

## **High Level Architecture Overview**



Incorta consists of several applications and services:

* Administrators use the Cluster Management Console (CMC) to manage the Incorta Nodes, Services, and Tenants in a standalone or cluster environment. The CMC is a Tomcat web application.
* An Incorta Node manages one or more Incorta Services. Apache Zookeeper helps manage Incorta Node agents.
* The Loader Service is analogous to a data warehouse and the Analytics Service is analogous to a data analytics service. Both the Loader and Analytics Services use in-memory Direct Data Mapping. Both services run as Tomcat web applications.
* Apache Spark (optional) supports creating materialized views and advanced analytic queries, including machine learning for predictive analytics.
* Incorta uses shared storage for both the Loader and Analytics Services.
* Incorta stores all objects, such as schemas and dashboards, in a Shared Metadata database.

## **Loader Service: Connectors**

Graphical user interface

Description automatically generated

Incorta allows you to connect to a variety of data sources using various **connectors.** You simply create a Data Source in Incorta or upload Data Files. Incorta saves the descriptive properties of both the Data Sources and Data Files in the Incorta Metadata database.

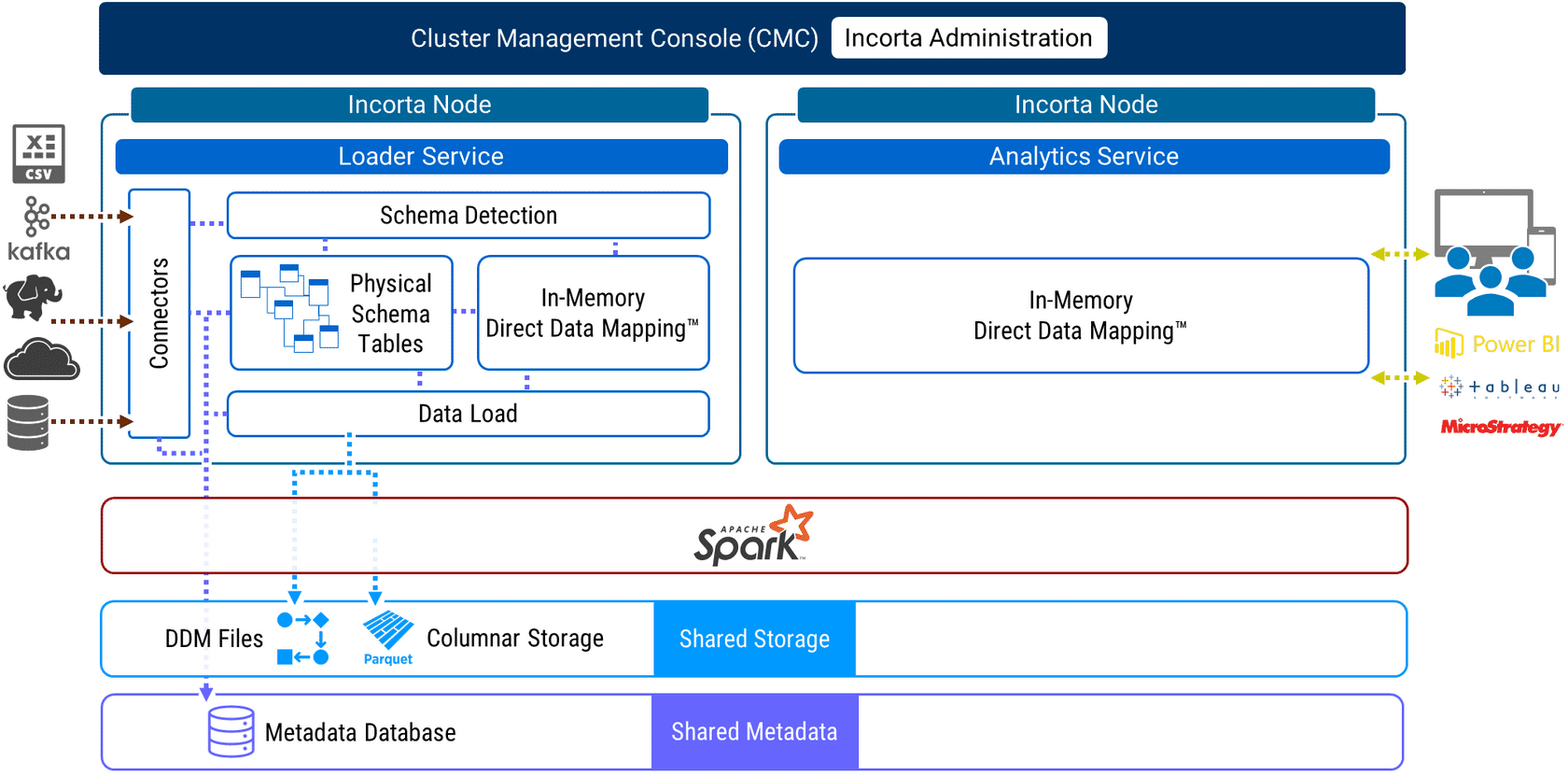
## **Loader Service: Physical Schema Tables**

Graphical user interface

Description automatically generated

With the Schema Wizard, Incorta can detect the nature of an underlying Data Source or Data File.  Schema detection details the various tables, the relationships between tables, and for each table, the columns and associated data types. For each column, Incorta determines if the column is a Key, Dimension, or Measure.  
  
You create your own physical schemas in Incorta as well as modify any schemas created by the Schema Wizard. The very creation of a physical schema triggers Incorta's in-memory Direct Data Mapping.

## **Loader Service: Data Load**

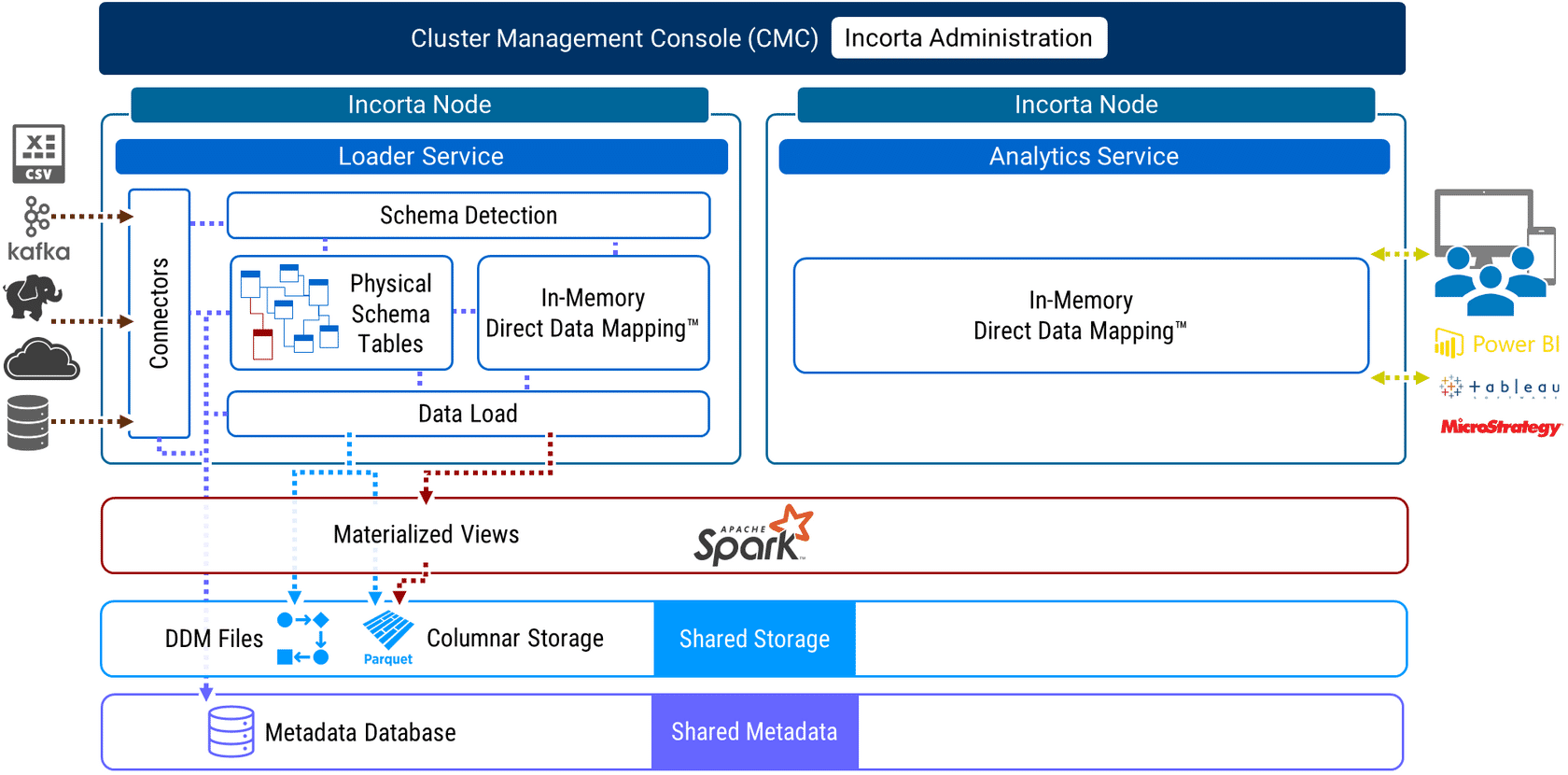


There are several ways to load data into Incorta: Full, Incremental, and Scheduled. You can load an entire schema or a single table in a schema.

Incorta copies the source system data during a data load and stores this data as Apache Parquet files in Shared Storage. Apache Parquet is a modern, columnar storage format.

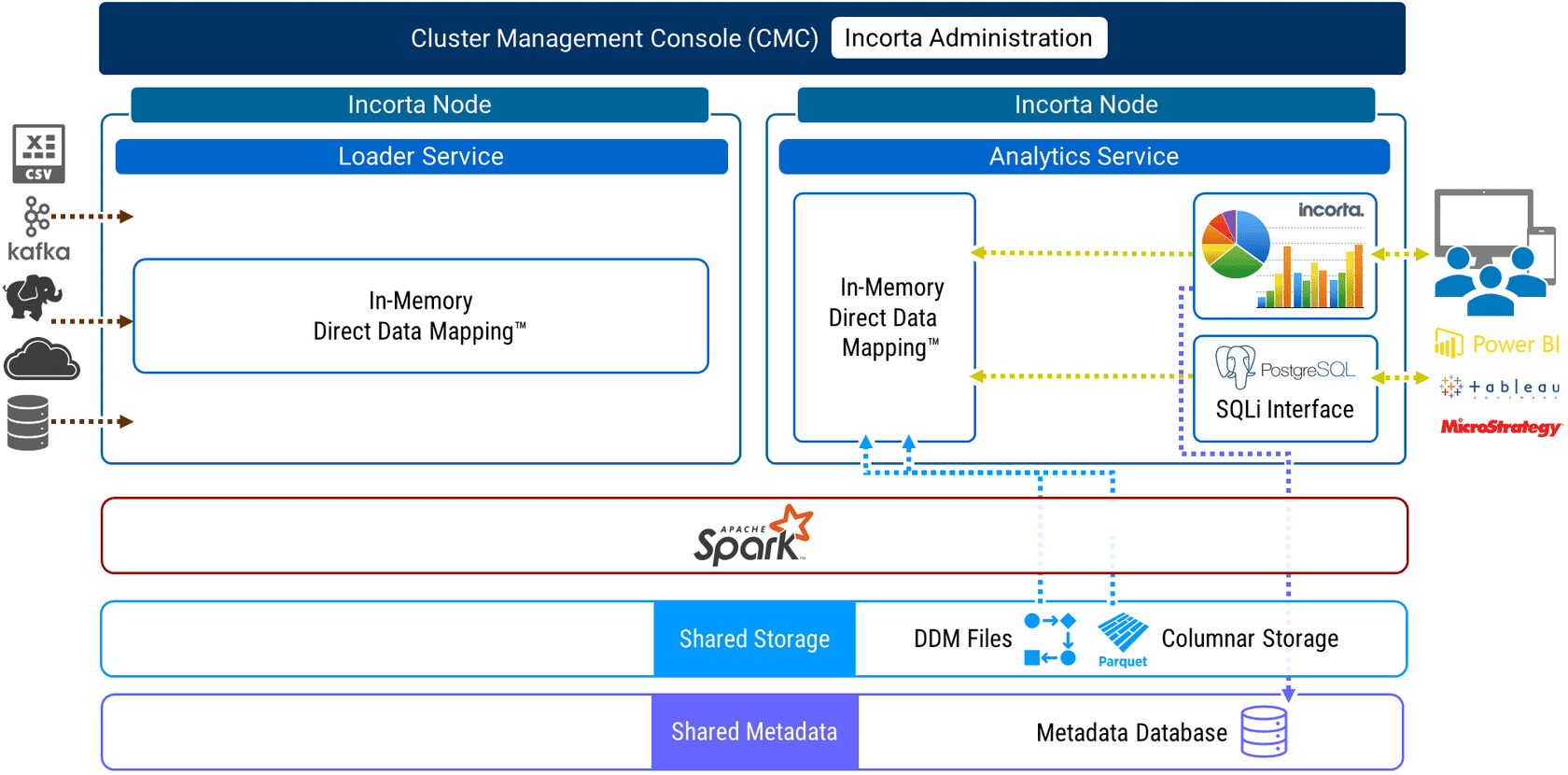
Incorta then creates an in-memory Direct Data Map of the schema and persists it to shared storage as Direct Data Mapping files (DDM files), also referred to as Snapshot files. The Direct Data Map — in-memory and on-disk — describes the schema data, tables, keys, indexes, and relationships between tables.

## **Loader Service: Materialized Views**



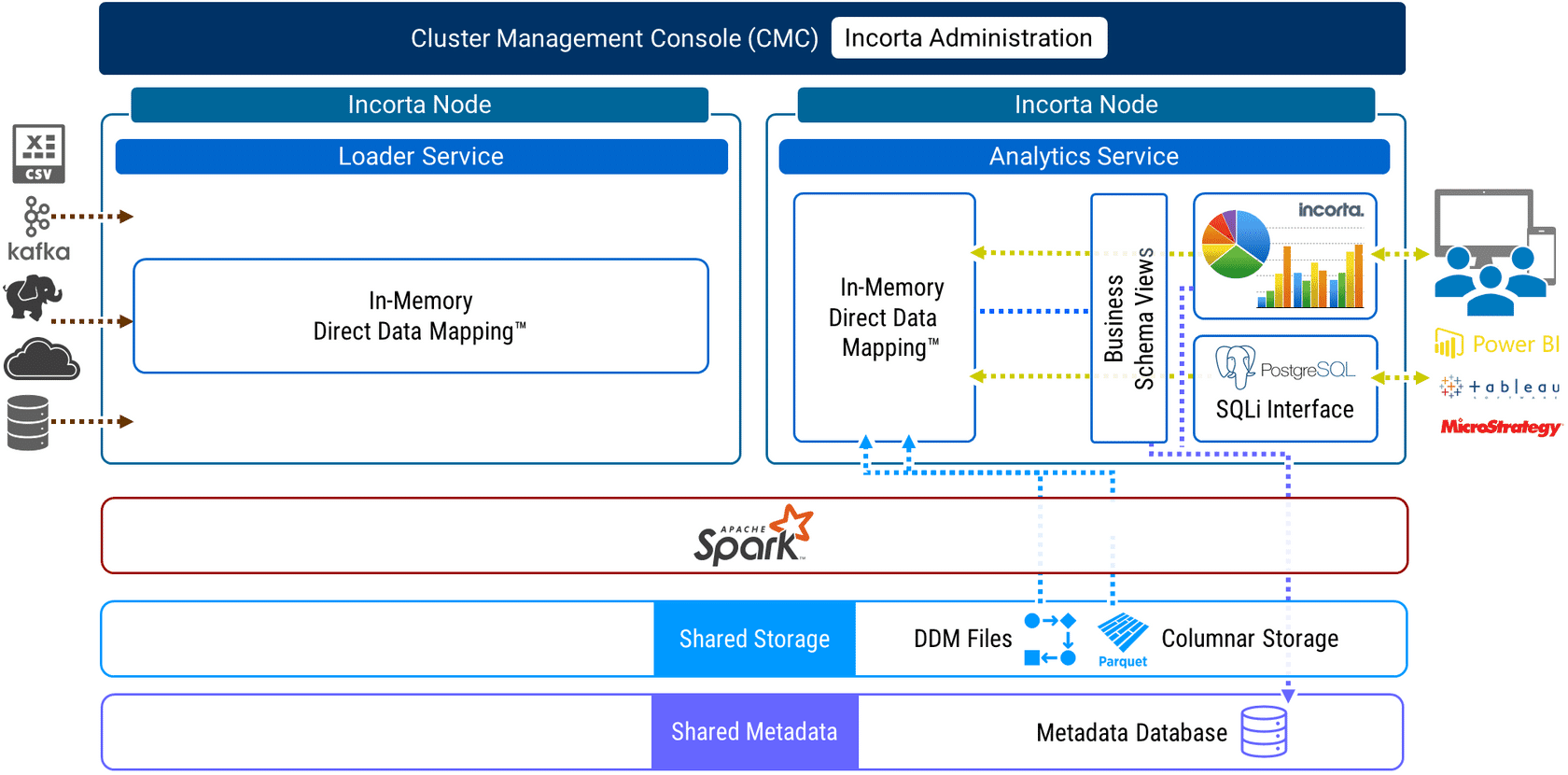
Incorta allows developers to create a special type of table called a Materialized View. Developers create Materialized Views in Physical Schemas using SQL or PySpark.  
  
Incorta creates Materialized Views using Apache Spark. Apache Spark is a modern, distributed compute platform. Incorta instructs Spark to materialize a view and persist the view in Shared Storage (Staging) as an Apache Parquet file.

## **Analytics Service: Dashboards and SQLi Interface**



The Analytics Service is how users access Incorta as both a modern data warehouse and data analytics platform.  
  
The web user interface (Web UI) for the Analytics Service manages all tenant objects, including data sources, data files, schemas, business schemas, dashboards, and users. The Analytics Service stores all object properties in a Shared Metadata database which the Loader Service utilizes for scheduled jobs. Business users, analysts, and developers often create, edit, and share dashboards using the Analyzer, a key feature of the Analytics Service.  
  
Using the SQLi Interface, third-party Business Intelligence applications and tools such as Power BI, Tableau, and Microstrategy, can execute SQL queries directly in Incorta. The SQLi Interface uses the PostgreSQL protocol.

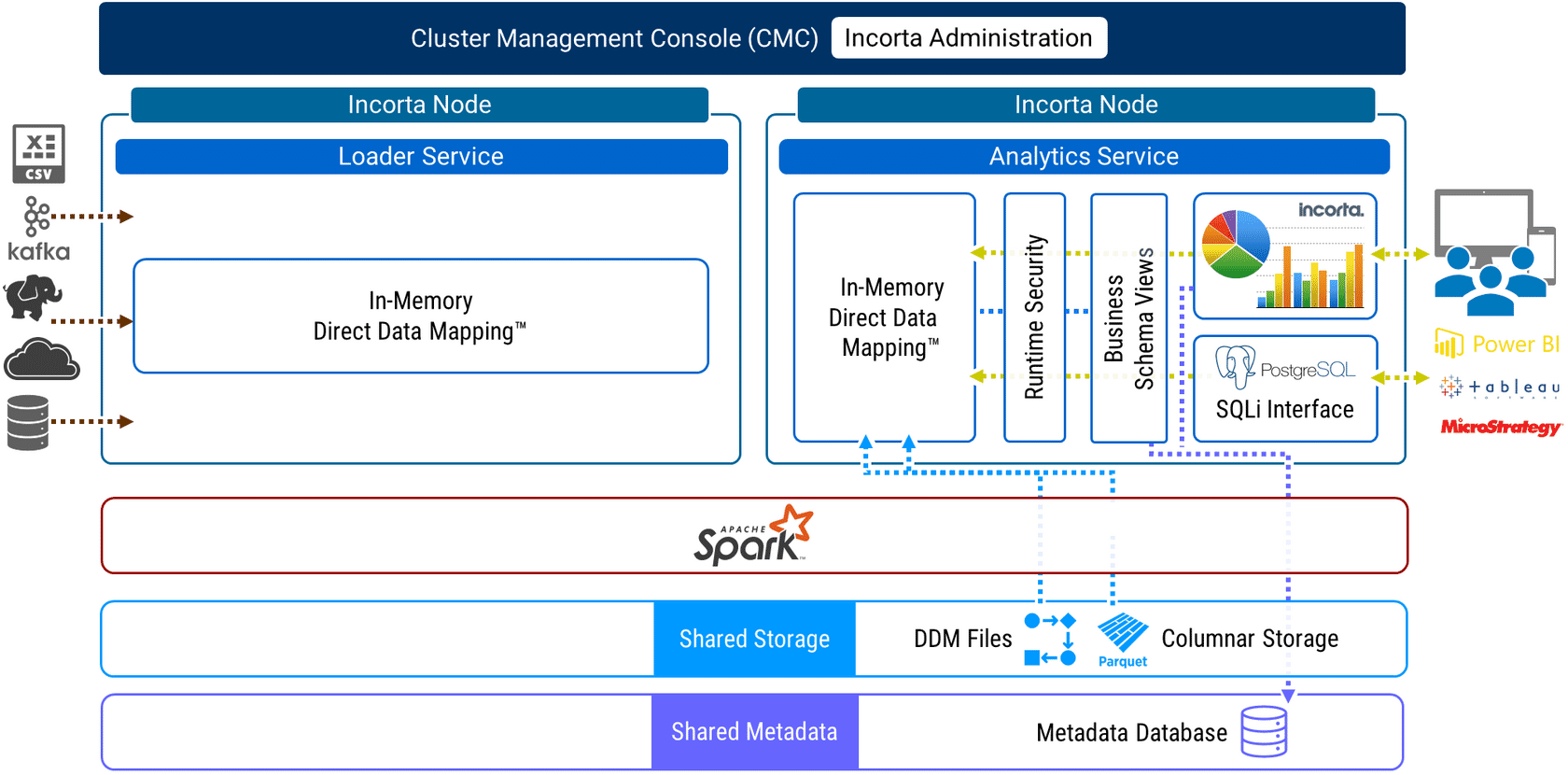
# **Analytics Service: Business Schema Views**



A Business Schema View is a runtime view of one or more physical schema tables.

Developers create Business Schema Views to introduce a semantic layer of both standardized business terminology and standardized formula calculations.  
  
Formulas in Business Schema Views can be runtime aggregations and may reference other formulas.

# **Analytics Service: Runtime Security**



Incorta supports a runtime security model for Row Level Security (RLS).  
  
Administrators and Developers implement Runtime Security Filters on specific tables that either match values in a given table column or use formulas to determine record level access.

# **Analytics Service: Advanced Analytics and Machine Learning**

Graphical user interface

Description automatically generated

For certain query shapes, the SQLi Interface will direct the query to Apache Spark for processing. Typically, these queries contain advanced SQL syntax such as window functions.  
  
Some dashboards in Incorta may use materialized views. Materialized Views may contain PySpark references to Spark's Machine Learning library (Spark MLib). In these cases, Incorta will use the Spark MLib for predictive analytics.