The Death of ETLs and The Traditional Data Warehouse

Remember the data warehousing promises of the past?

- Data organized for ease of access and understanding

- Data at the speed of business

- Single version of truth

Today nearly every organization operates at least one data warehouse, most have two or more. The goal of fast, easy, and single source still remain elusive. Once the source of pride, Data Warehouses are now seen as troublesome legacy systems.

Past practices for data warehousing struggle to meet the needs of today in the age of big data and fast analytics. Legacy data warehouses don't scale easily, are performance challenged, and are constrained by relational models for structured data. More importantly they are slow to build, slow to deploy, slow to change, and slow to deliver.

The time has come to rethink and remake data warehousing. Despite their challenges, we can't simply decommission data warehouses and rely solely on a data lake. The primary purposes of a data warehouse are integration and reconciliation of enterprise data and retention of time-variant snapshots of business history. Failure to integrate and organize data limits the ability to publish information people routinely use for their job. The publish-and-subscribe model works and people need the information published for them. Lack of time-variant history with uniform intervals inhibits the ability to perform time-series analysis and to understand trends in the business. The big question when seeking to modernize legacy data warehousing is how to realize the benefits of data warehousing without facing the challenges of managing a data warehouse?

A Modern Alternative

Incorta offers a seamless, end-to-end analytical warehouse solution engineered for simple and powerful real-time analysis of massive volumes of data. High-impact graphic representations transform raw data from arcane and inaccessible to clear and meaningful, delivering unprecedented insight into patterns, trends, key performance indicators, and their causality. Unifying multiple data sources into a single secure source of truth ensures that all users are able to analyze, strategize, and collaborate around the latest and most relevant information.

The Incorta Analytics platform removes the dependency on star schemas, pre-aggregation, complex ETL processes, and a large data team. With Incorta's cloud technologies and mobile-first design, even a mobile user can merge external real-time data sets like weather data with large-volume business data, and then make informed time-sensitive decisions based on the results.

Incorta's Value Proposition

- The rapid expansion of data usage through the integration of multiple, disparate systems without the need for traditional warehouse or marts

- Query performance five times faster than high-end data appliance-based systems

- 90% reduction of deployment time and 70% reduction of ETL processes

- 85% reduction of maintenance costs

- Deployment-ready in AWS, Azure, GCP or on-premise using commodity HW

- Over the last 30-years the traditional approach to analytics has followed a fairly standard supply chain path:

✓ Identify data	✓ ETL data Into a staging
✓ Model data and push to a warehouse	✓ Perform secondary modeling in the warehouse
✓ Create summarizations/aggregations	✓ Build a semantic layer
✓ Generate reports/analysis	

The Incorta Architecture
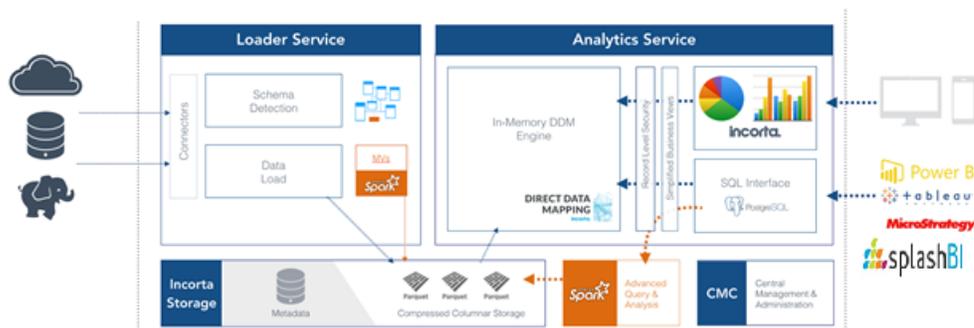
The architecture above is a logical representation of the Incorta Platform. It is a Java-based software solution deployed in a distributed environment to increase concurrency, isolate internal activities to improve throughput or satisfy high availability requirements.

Incorta ingests data through a series of connectors, predominantly using JDBC although other methods are possible using our provided SDK. The specific data sources supported by the connectors grow on a release by release basis. At present, Incorta ships with more than 30 embedded connectors and has access to an additional 200 connectors through partnerships with Cloud Elements, SAP, and Informatica.

At ingest, Incorta replicates the source system data and stages it in Apache Parquet files. Parquet becomes the Incorta persistence layer. Parquet is a compressed and highly efficient columnar database that meets the need to capture and store enterprise data and to collect historical data that is essential for time-series analysis. Parquet offers many advantages for data management, including reduced disk consumption through compression, performance gains with parallel processing efficiencies, and the abilities to store complex nested data structures while breaking away from relational model constraints. Building on the Parquet foundation, Incorta provides exceptionally fast data ingestion from a variety of sources in near real time.

This document is an abstract from a nine-page document; we are happy to send you the entire document. For additional information, or to set up a demo please call Bob Heriford at 719-494-6182 or via email to robert.heriford@solutions4less.tech . Save your organization hundreds of thousands of dollars! There is so much more to say about Incorta's functionality, direct data mapping, high-speed queries and prebuilt analytics content for many ERP system.

We look forward to serving from you!