

D&A TRANSFORMATION TALES

# Annihilating Vulnerabilities with Data Lakes



## BACKGROUND

A leading data analytics and market research company catering to consumer-packaged goods, retail and healthcare industries approached Aspire to help them create a Data Warehouse in Azure Synapse to store and process their data safely and securely from various sources. The customer wanted to:

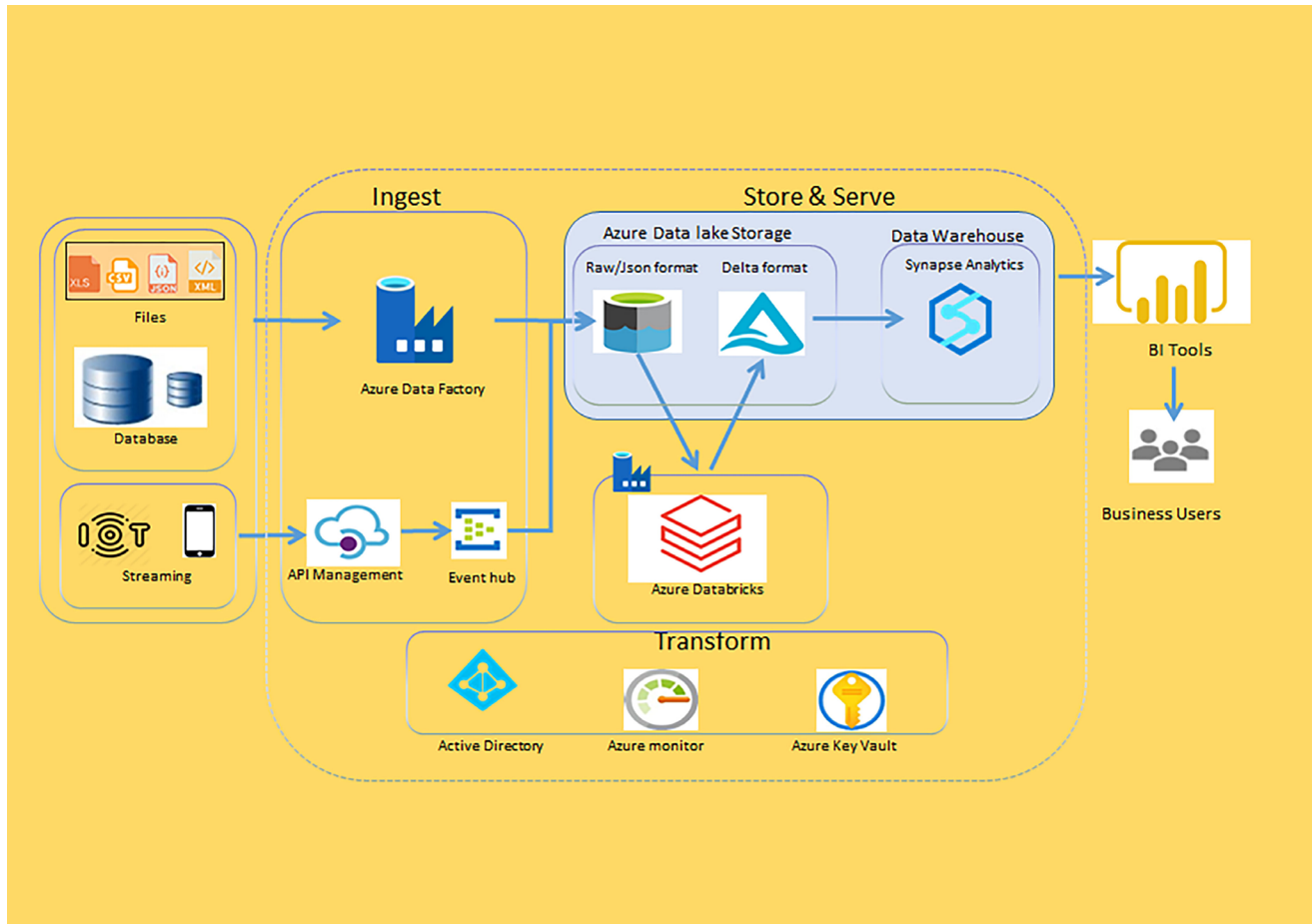
- Manage a single version of truth for reporting requirements
- Develop a data repository in Azure Synapse for easy extraction of data for BI reporting and dashboard
- Proper handling and management of historical and current data
- Diligently schedule and monitor daily, weekly, monthly data load
- Filter and clear our redundant data
- Enhance documentation support for various sources
- Reduce data volume for incremental load
- Define and simplify data received for merge/overwrite/append
- Streamline data ingestion tool based on input source
- Manage, identity and secure access from external resource



## THE SOLUTION

After taking the client's needs and challenges into consideration, Aspire's Data and Analytics team curated a well-crafted plan to build Data Lakes that would help the client process, store, clean, share, analyze, model and monetize their datasets using BI and ML solutions. The solution involved:

- Azure Data Factory to load Data from Flat file and Database to Azure Data Lake
- API and Eventhub to ingest data from streaming to Azure Data Lake
- Azure Data Lake used to store raw data as well as delta Data
- Raw data was processed through Databricks and Delta data to be stored again in the Data Lake
- All notebooks of Databricks were generic/re-usable.
- Data bricks Container workspace was used for ETL processing & writing data into data warehouse
- BI report was visualized using Azure synapse Data
- Daily /weekly/Monthly load job Scheduled and monitored through ADF and Azure monitor





## TECHNOLOGY SNAPSHOT

- Enterprise Cloud Platform – Azure
- ADF for Orchestration and data ingestion
- Enterprise storage – Data Lake storage for raw data and processed data sets
- Databricks for in-memory data
- Enterprise Warehouse– Azure Synapse
- Cloud Identity and security - AAD, Azure key vault
- Cloud monitor- Azure Monitor



## BUSINESS BENEFITS

- Enhanced data quality and consistency
- Increased the power and speed of data analytics
- Seamlessly syncs with on-premise and cloud
- Enhanced query performance and improved historical insights
- Involved heterogeneous sources to produce insightful reports
- High return on investments with quality reports





Aspire Systems is a global technology services firm serving as a trusted technology partner for our customers. We work with some of the world's most innovative enterprises and independent software vendors, helping them leverage technology and outsourcing in our specific areas of expertise. Our core philosophy of "Attention. Always." communicates our belief in lavishing care and attention on our customer and employees.

For more info contact: [info@aspire.com](mailto:info@aspire.com) or visit [www.aspiresys.com](http://www.aspiresys.com)

**USA**  
+ 1 630 368 0970

**SINGAPORE**  
+65 3163 3050

**INDIA**  
+91 44 6740 4000

**BELGIUM**  
+ 32 3 204 1942

**NETHERLANDS**  
+ 31 (0)30 800 92 16

**POLAND**  
+48 58 732 77 71

**MEXICO**  
+52 222 980 0115