



# Teradata Analytics Platform

Sri Raghavan  
Senior Global Product Marketing Manager



## Agenda

- Driving Business Value with Teradata Analytics
- Evolving from a Database to an Analytics Platform
- Drilling down into Technology
  - Analytic Engines
  - Tools
  - Languages
  - Data types, formats and sources
- Futures



**Customer**  
Experience



**Finance**  
Transformation



**Product**  
Innovation



**Risk**  
Mitigation



**Asset**  
Optimization



**Operational**  
Excellence



**Customer**  
Experience



**Finance**  
Transformation



**Product**  
Innovation



**Risk**  
Mitigation



**Asset**  
Optimization



**Operational**  
Excellence

# Multi-Channel Account Closure Detection: Global Bank

Reduce account closure incidence

Customer account closure and its impact on satisfaction requires constant model tweaking. Highly manual.

CHALLENGE



Multi-Channel data consolidation of events of interest.

SOLUTION



- Sessionize
- nPath
- Dynamic Time Warping Analysis

ANALYTICS



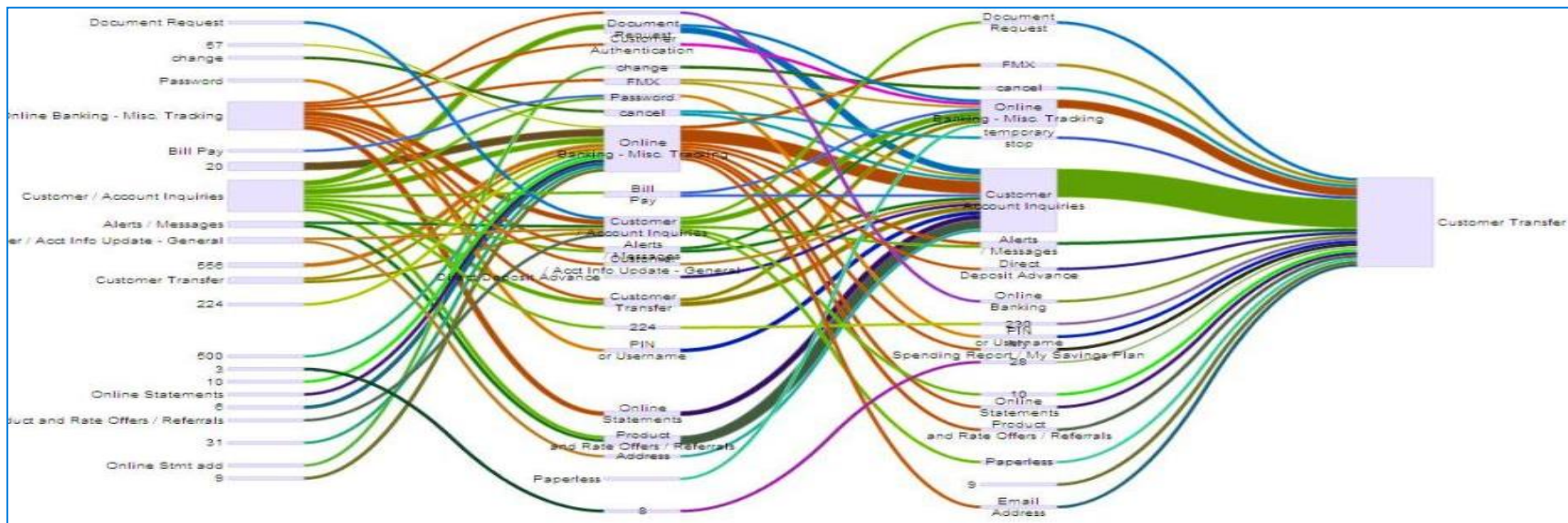
Developed a constantly evolving toolkit to detect new closure patterns

OUTCOME

New closure event patterns detected the following day

# Account Closure Analysis Outcome

## Account Closure Paths



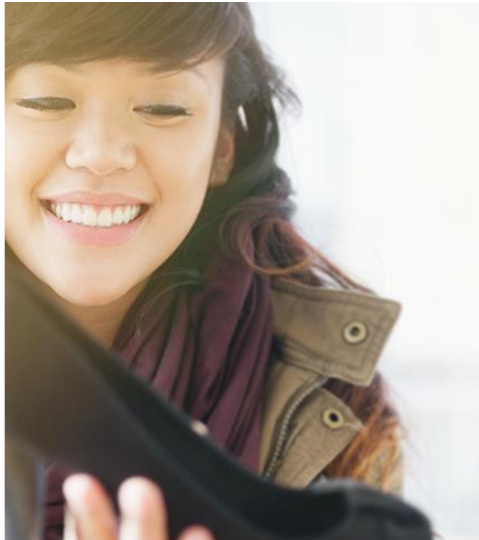
**Customer**  
Experience

**Finance**  
Transformation

**Product**  
Innovation



**Risk**  
Mitigation



**Asset**  
Optimization



**Operational**  
Excellence



# Customer Experience: Major Retailer

Understand a customer's experience

Customers interact through many channels but the impact of each channel in a customer's journey is not clearly known.

CHALLENGE



Multi-Channel data consolidation of events of interest.

SOLUTION



- Sessionize
- nPath
- Attribution
- Customer Satisfaction Index
- Text Analytics

ANALYTICS



Delivered a repeatable set of experiences that ensures that customers attain their objectives at the least cost in the shortest amount of time

OUTCOME

New channel activities are constantly measured throughout the journey



# Advanced Customer Experience Analysis

SQL Engine

+

Graph Engine

+

Machine Learning  
Engine

= Best in  
Class CX  
Analytics

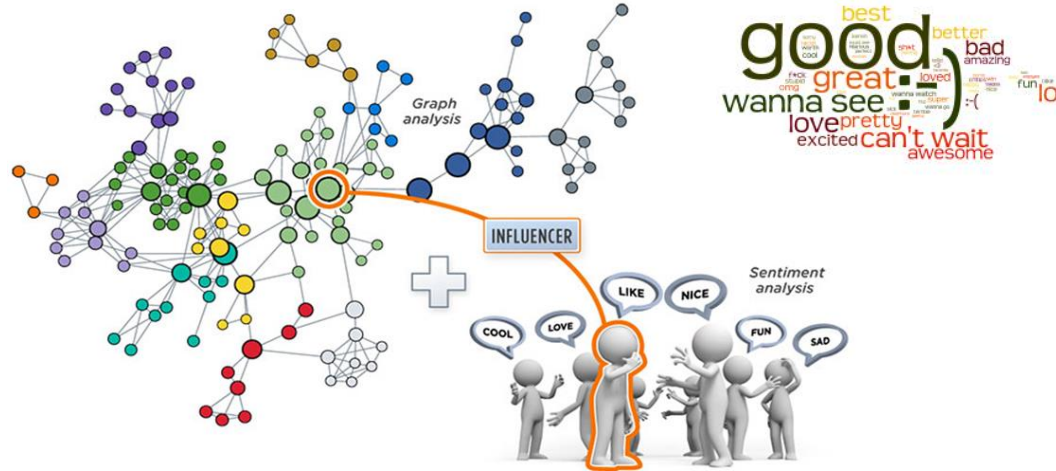
Statistical  
Analysis

Influencer  
Analysis

Sentiment  
Analysis



Multi Channel  
Behavioral Path  
Analysis



TERADATA



**Customer**  
Experience



**Finance**  
Transformation



**Product**  
Innovation



**Risk**  
Mitigation



**Asset**  
Optimization



**Operational**  
Excellence

# Product Innovation (IoT): Major Manufacturer

Detect parts failure

Need to understand equipment usage and behavior to prevent down time or unplanned outages across many support groups

CHALLENGE



Multi equipment sensor and performance data

SOLUTION



- Log Parsers
- Generalized Linear Model
- Weibull analysis
- Naive Bayes

ANALYTICS

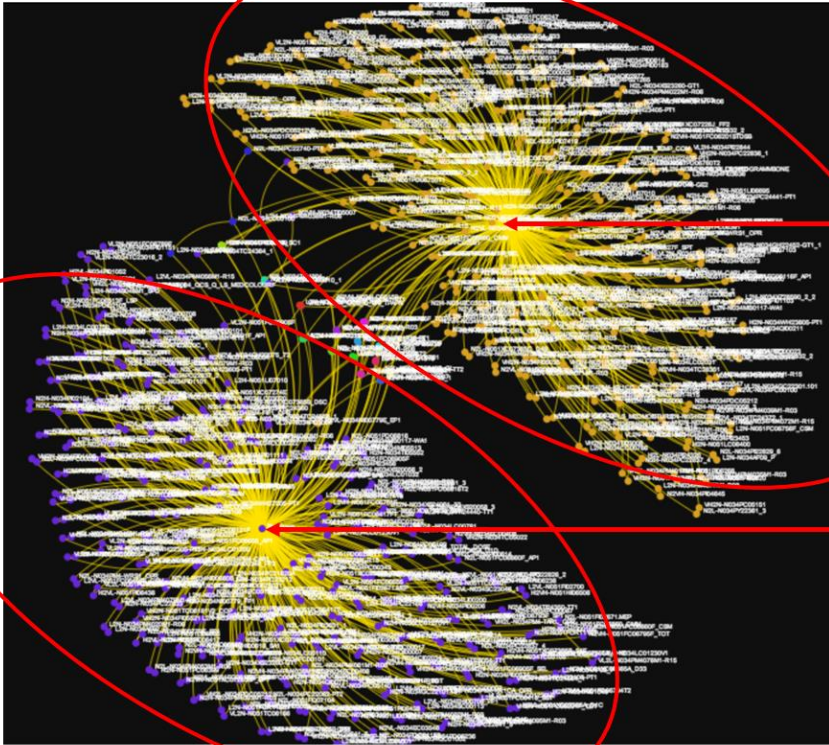


Isolated key indicators of performance bottlenecks that helped remove soon to be defective equipment from circulation for proactive maintenance

OUTCOME

Sensor data continuously ingested and analyzed for pattern changes and performance triggers

# Product Innovation (IoT): Event paths leading to shutdown



Shutdown event  
and sensor paths

**versus**

Stable running event  
and sensor paths

# Evolution from a Database to an Analytics Platform

# Teradata Everywhere

FROM

Structured data warehouse

On-premises, proprietary hardware

Limited purchasing options

Inflexible, lock-in

TO

Analyze **Anything**

Deploy **Anywhere**

Buy **Any Way**

Move **Anytime**

De-risk Buying Decisions

# Teradata Analytics Platform

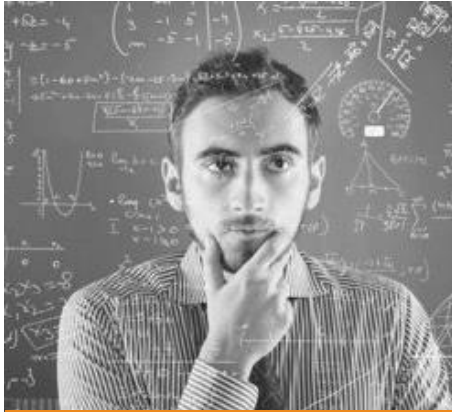
Analyze **Anything**

Deploy **Anywhere**

Buy **Any Way**

Move **Anytime**

# Teradata Analytics Platform — Bridging the Divide



**Data Scientist**  
Ben

**“Access to  
the toolbox  
of choice”**

Ben, the data scientist, wants the freedom to choose the best tools to create business outcomes through data



**CIO**  
Carole

**“Provide  
agility without  
anarchy”**

Carole, the CIO, wants to de-risk IT investments while operating at scale within SLAs and leveraging existing data and analytics investments to augment/grow current work



# Teradata Analytics Platform — Bridging the Divide

## Data Scientist

Quickly get to actual analysis

---

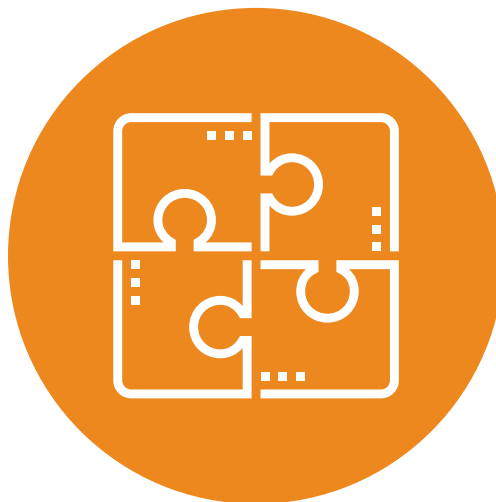
Access preferred tools and languages

---

Integrate multi-format data

---

Make analytics easy to repeat



## CIO

Accelerate time to deliver environments

---

Reduce operational costs

---

Support end user demands

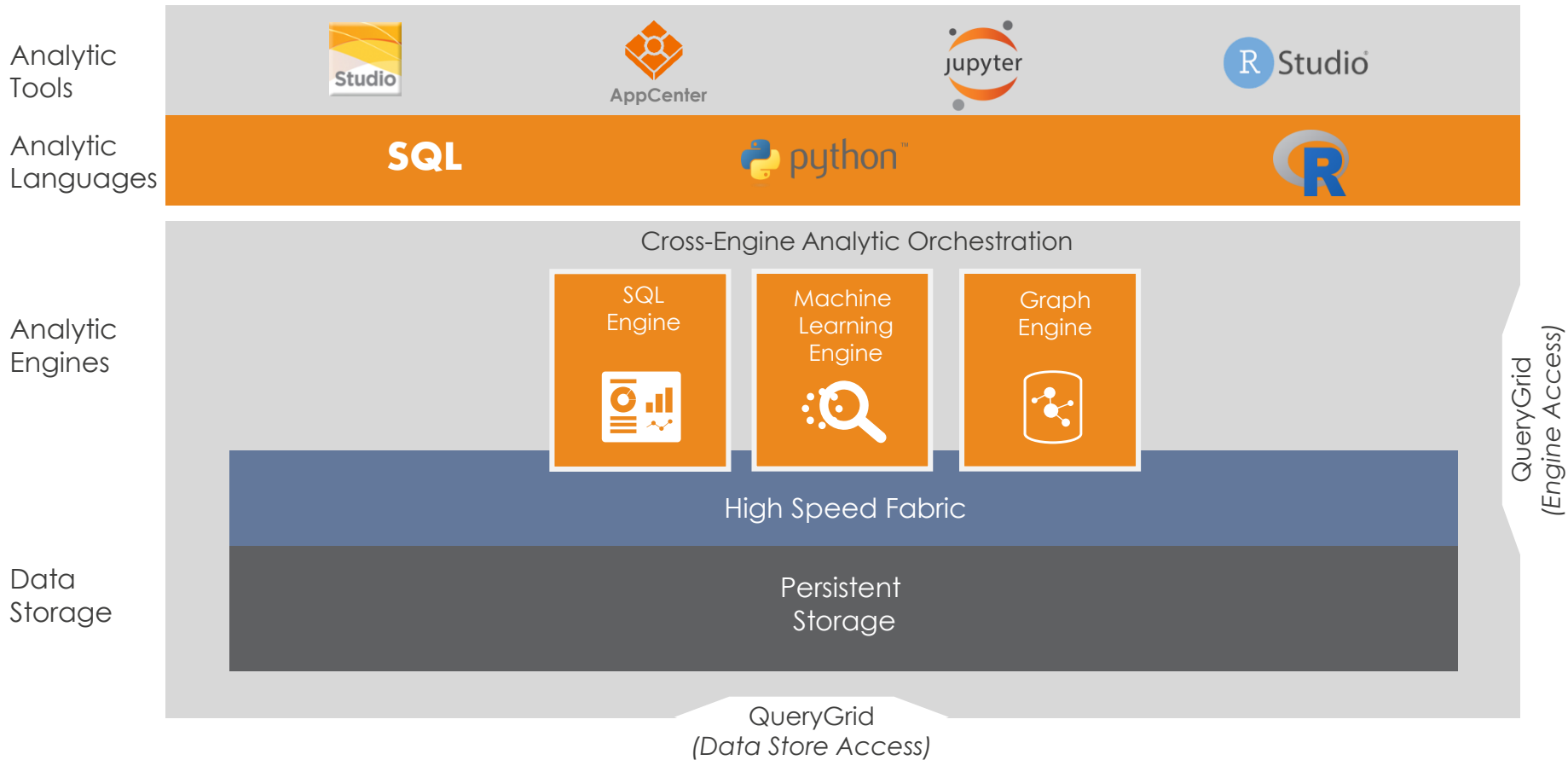
---

Provide leading-edge security

---

Leveraging existing data and analytics investment

# Teradata Analytics Platform

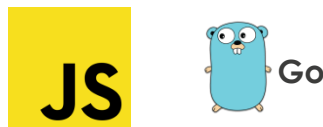


# Our Vision: Teradata Analytics Platform

Use your favorite analytic tools



Leverage the most suitable languages



Access a larger variety and volume of data, including new sources



Use the best analytic functions & leading analytic engines

TERADATA.



TERADATA.



**The Best Analytic  
Functions and Engines**



Preferred Tools  
and Languages



Support for  
Multiple Data Types

TERADATA ANALYTICS PLATFORM

# What is an Analytic Engine?

- **Definition**

- A comprehensive framework that includes all the software components that are well integrated to deliver advanced analytics functionality that can be implemented by a well defined set of user personas

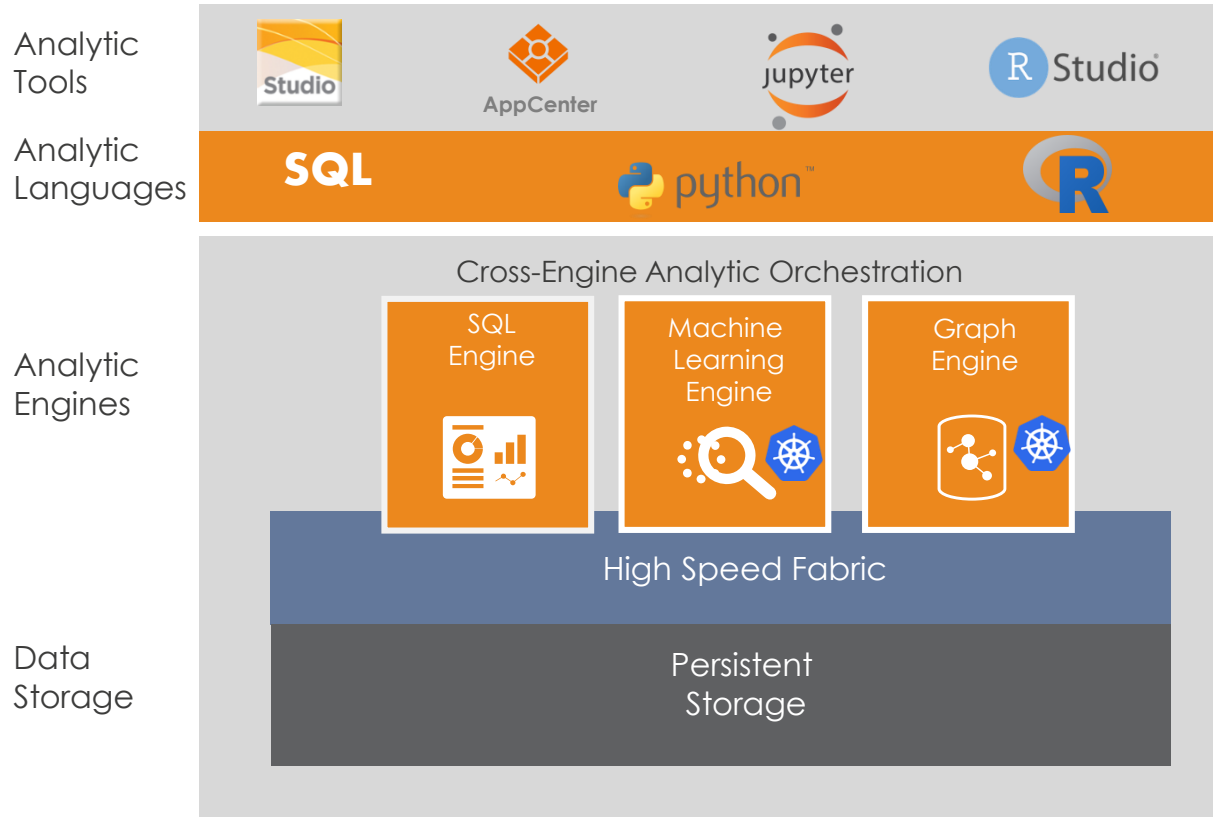
- **Components of an Analytic Engine**


- Advanced Analytics functions (e.g., Machine Learning, Graph, Time Series)
- Access points to data storage that can ingest multiple data types
- Integration into visualization and analytic workflow tools
- Built in management and monitoring tools
- Highly scalable and performant environment with established thresholds

- **Advantages of an Analytic Engine**

- Self-contained (containerized) analytic compute environment that can be separated from data storage
- Analytic engines can be tailored for access and use by specific personas (e.g., DS, Business Analyst)

# Teradata Analytics Platform

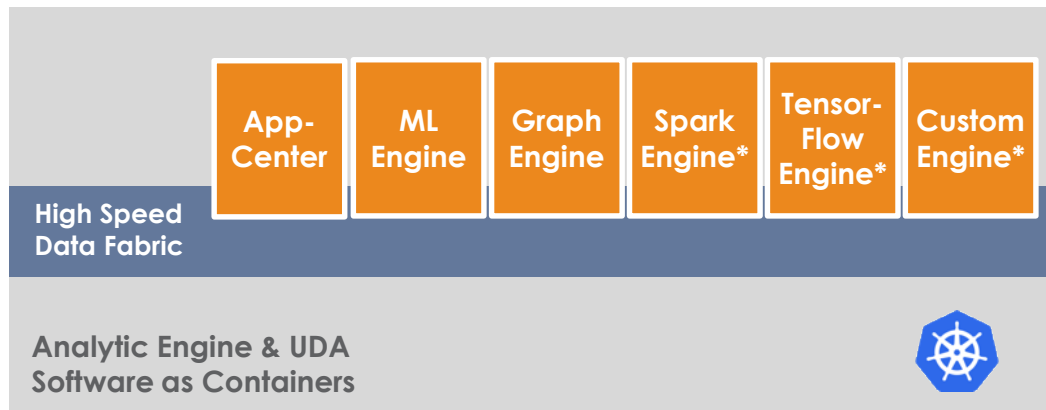


Kubernetes Managed Nodes.  
Requires additional hardware 

# Implemented Through Containerization



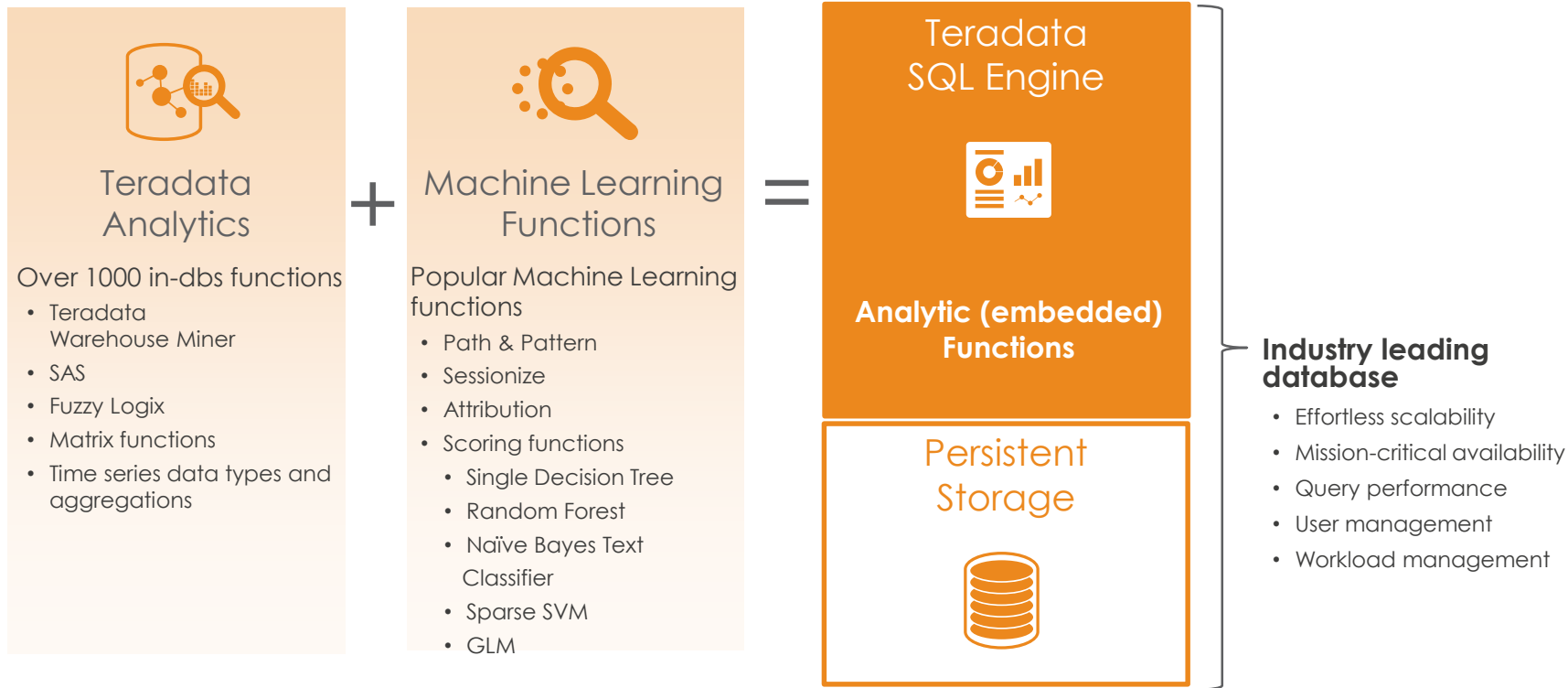
- Integrating Docker containers in Teradata Analytics Platform
- Containers enable
  - Add new analytic engines faster
  - Minimized downtime for customers
  - Automatic scale out of engines
  - Improved SLA through compartmenting end-users workload
- Type of Containers: Analytics, UDA, languages, open source, commercial products and custom engines



Teradata Analytics Platform Docker Containers

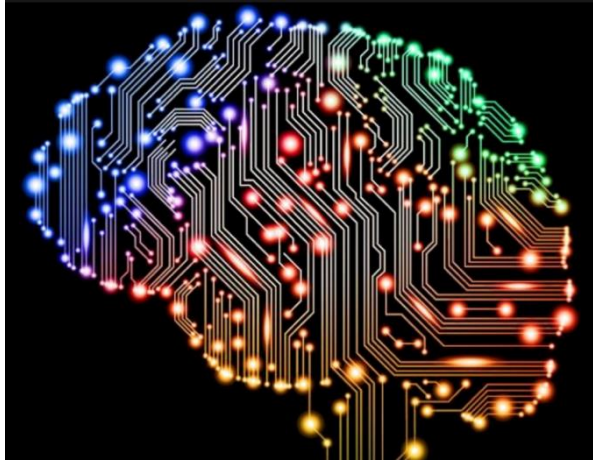
\* Anticipated future capabilities

# What is the Teradata SQL Engine?



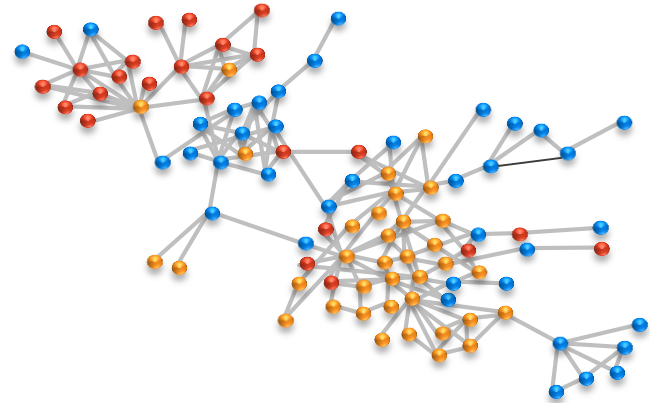


# What are the Machine Learning & Graph Engines?



- The Machine Learning Engines consists of a wide range of analytic capabilities that span analytic genres such as Artificial Intelligence, Statistics, Text & Sentiment determination in addition to Data Preparation and unstructured data Parser functions
- Examples of Machine Learning Engine analytics include Outcome Categorization, Sentiment Strength Assessment, Customer Path Behavior, Fraud Detection, and more

- The Graph Engine enables the analysis of relationships and how they influence outcomes in networks of people, products, or processes or really any entities that are connected in a network
- Examples of Graph Engine analytics include Network Threat Detection, Market Basket Rendering, Social Media Influencer Behavior, and more



# Machine Learning Engine Functions

# Graph Engine Functions

## Statistics

AdaBoost  
Approximate Distinct Count  
Approximate Percentile  
CMAVG  
ConfusionMatrix  
ConfusionMatrixPlot  
Correlation  
CoxPH  
CoxPredict  
CoxSurvFit  
Cross Validation  
Distribution Matching  
EMAVG  
Enhanced Histogram  
Fmeasure  
GLM  
GLMPredict  
Hidden Markov Model  
Histogram  
KNN  
LARS Functions  
LinReg  
LRTEST  
Non-linear Kernel SVM  
Percentile  
Principal Component Analysis  
Random Sample  
ROC Curve  
Sample  
Shapley Value  
SMAVG

## Statistics (cont.)

Support Vector Machines  
VectorDistance  
VWAP  
WMAVG

## Data Transformation

Antiselect  
Apache Log Parser  
Fast Fourier Transform  
FellegiSunterTrainer  
FellegiSunterPredict  
IdentityMatch  
IpGeo  
Inverse Fast Fourier Transform  
JSONParser  
Multicase  
MurmurHash  
Number as Categories  
OutlierFilter  
Pack  
Pivot  
PSTParserAFS  
Scale Functions  
StringSimilarity  
Unpack  
Unpivot  
URIPack  
URIUnpack  
XMLParser  
XMLRelation

## Path, Pattern and Time Series

Arima  
ArimaPredictor  
Attribution  
Burst  
ChangePointDetection  
Causality Detection  
DTW  
DWT  
DWT2D  
FrequentPaths  
IDWT  
IDWT2D  
Interpolator  
Path\_Analyzer  
Path\_Generator  
Path\_Start  
Path\_Summarizer  
SAX  
SAX2  
SeriesSplitter  
Sessionization  
Shapelets  
TimeSeriesOrders  
Unsupervised Shapelets  
VARMAX

## Visualization

CfilterViz  
NpathViz

## Association

Basket\_Generator  
Cfilter  
FPGrowth  
KNN Recommender  
WSRecommender

## Cluster

Canopy  
Categorical Clustering  
Gaussian Mixture Model  
KMeans  
KMeansPlot  
Minhash

## Decision Tree

XGBoost  
Forest\_Drive  
Forest\_Predict  
Forest\_Analyze  
Single\_Tree\_Drive  
Single\_Tree\_Predict

## Naïve Bayes

naiveBayesMap  
naiveBayesReduce  
naiveBayesPredict

## Deep Learning

Neural Networks

## Text

Chinese Text Segmentation  
LDA Functions  
Levenshtein Distance  
Named Entity Recognition (CRF Model)  
Named Entity Recognition (Max Entropy Model)  
nGram  
PoStagger  
Sentenizer  
Sentiment Extraction Functions  
Text Classifier  
Text\_Parser  
TextChunker  
TextMorph  
TextTagging  
TextTokenizer  
TF\_IDF

## System Functions

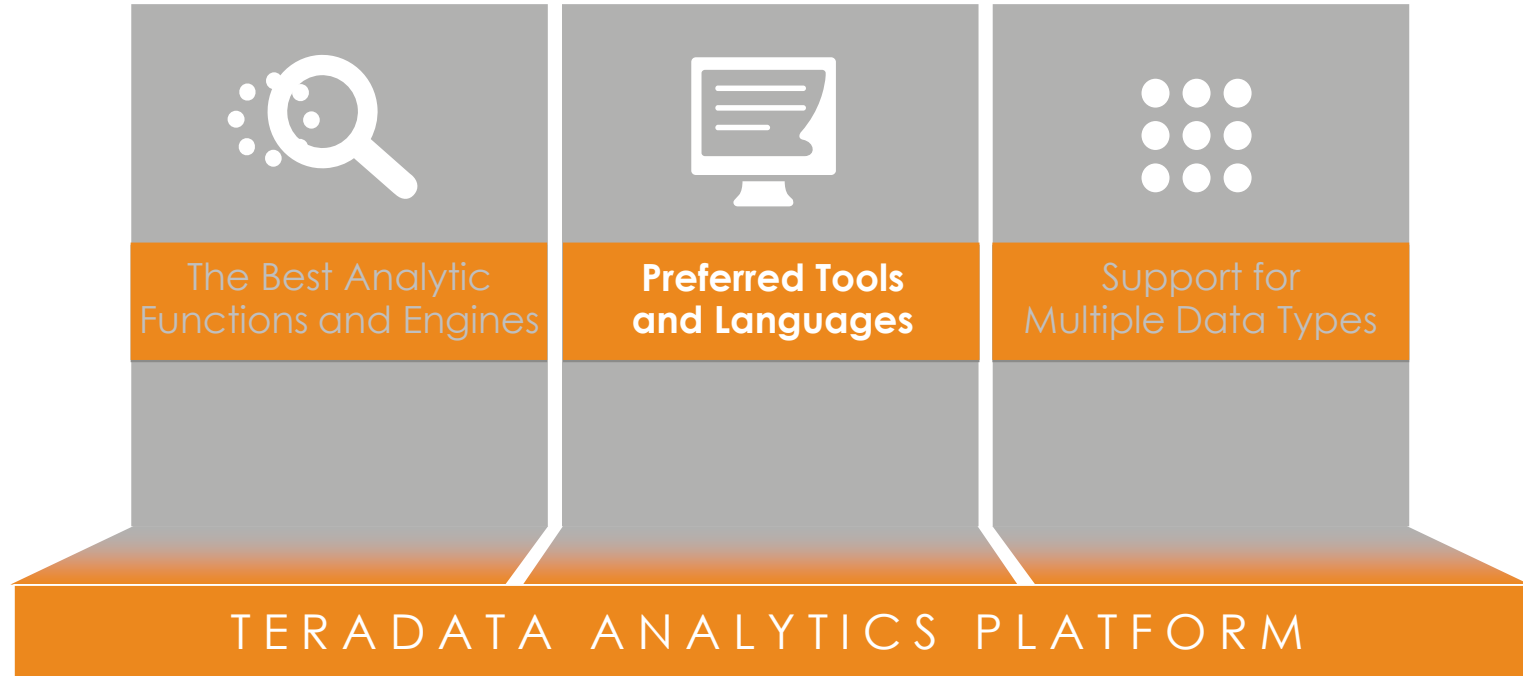
nc\_skew  
nc\_relationstats

## Location Analysis

LoadGeometry  
PoinfinPolygon  
GeometryOverlay

## Graph

AllPairsShortestPath  
Betweenness  
Closeness  
EigenvectorCentrality  
gTree  
LocalClusteringCoefficient  
LoopyBeliefPropagation  
Modularity  
nTree  
PageRank  
PersonalizedSALSA  
RandomWalkSample

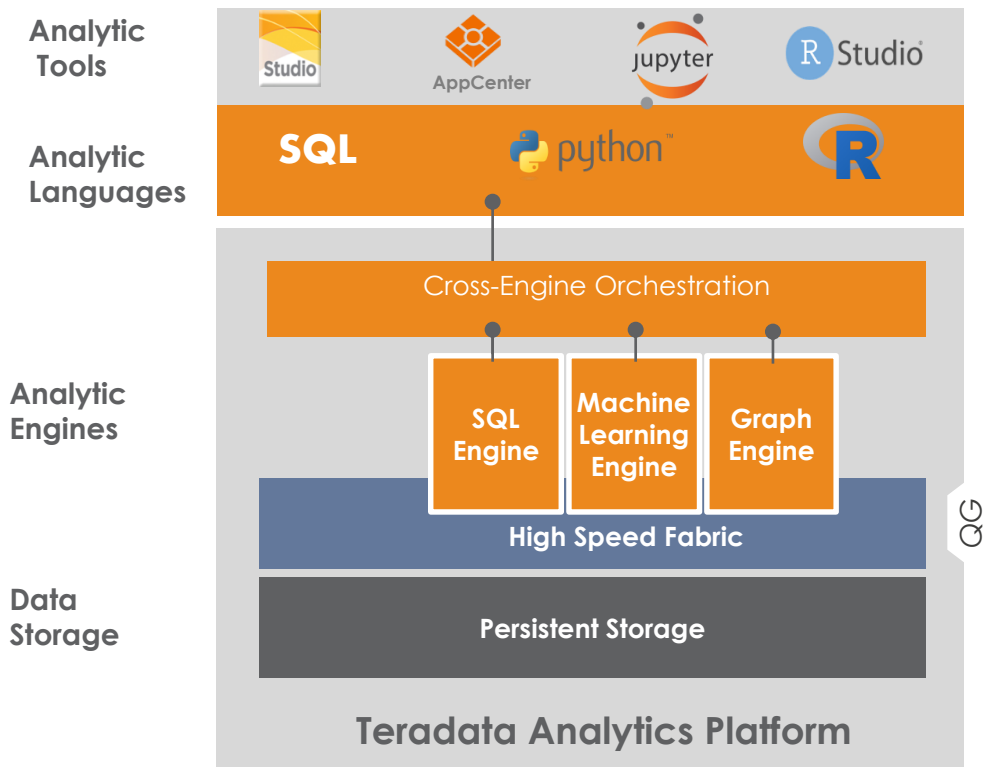


\*Future

# Analytic Tools Support



- Teradata Studio or any SQL editor
  - SQL
- AppCenter
  - SQL & Python
- Jupyter Notebooks
  - SQL and R client package
  - AppCenter
- Rstudio
  - Call R functions via R library



# Analytic Languages Integration



- SQL
  - Native and AppCenter
- R
  - R client package, embedded process and AppCenter
- Python
  - Embedded process and AppCenter

Analytic Tools



Analytic Languages

SQL



python



Analytic Engines

Cross-Engine Orchestration

SQL Engine

Machine Learning Engine

Graph Engine



High Speed Fabric

Data Storage

Persistent Storage

Teradata Analytics Platform



The Best Analytic  
Functions and Engines



Preferred Tools  
and Languages



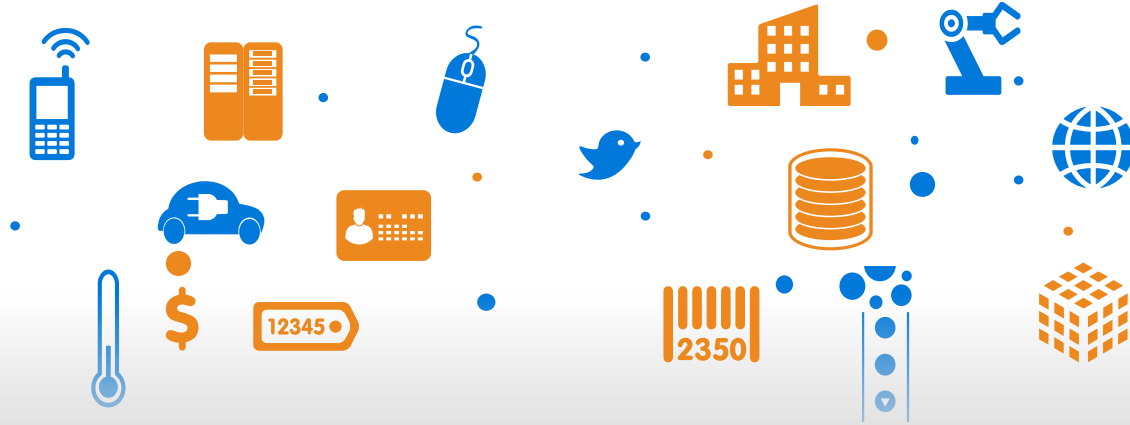
**Support for  
Multiple Data Types**

TERADATA ANALYTICS PLATFORM

\*Future

TERADATA.

# Teradata Support for Multiple Data Types, Formats and Sources



## Data Formats

JSON, BSON, AVRO  
CSV, XML, PDF, Voice,  
Video and Images

## Data Types

Geospatial, Temporal  
and Time Series

# Business Outcomes through Time Series Analysis



**Optimized delivery and pricing plans** by geography, product type, and other variables



**Smarter maintenance protocols** to catch equipment degradations well before a full blown crisis



Monitoring of prescriptions and pharmacological disbursements to **prevent abuse and theft**



**Optimizes usage patterns** by providing efficiency recommendations and **prevention of unauthorized or illegal access** to power systems



# Teradata Database – Time Series Capabilities

## Teradata Database 16.20

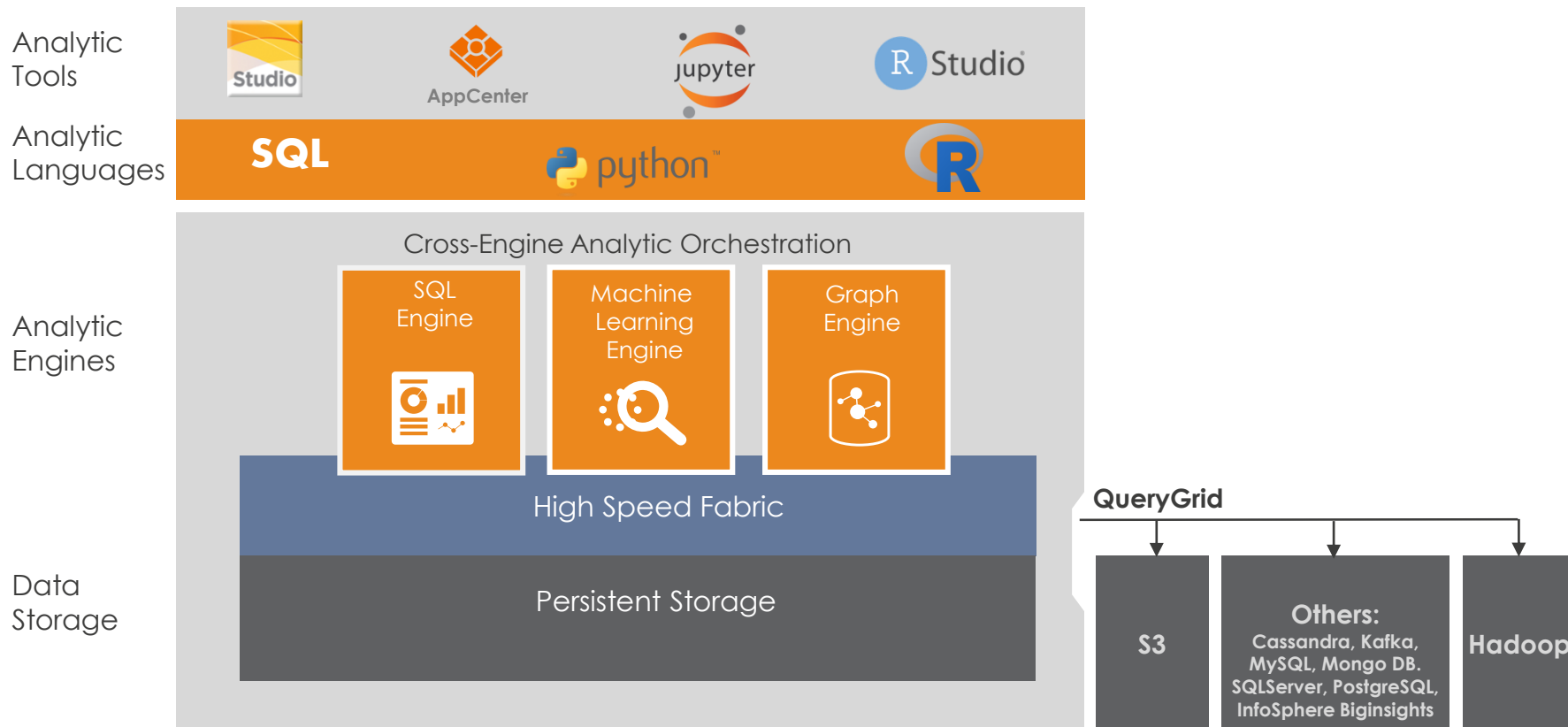
### Agile Analysis enabled by Time-Aware Functions

- Time period aware aggregations
- Work with ANY time component data
- Impute missing values
  - Ignore, removed, update with constant

### High-Performance enabled by Primary Time Index (PTI)

- Supports time sensitive decisions
- Fast access through:
  - Hash distribute by time bucket
  - AMP-local processing
  - Sequenced data

# Data Sources and Storage





Teradata IntelliSphere™



# The Reality

Wide range of deployment choices

More analytic tools & techniques

Proliferation of departmentalized analytic

Dynamically changing analytic workloads

Large upfront investment

**Constant Change** is the norm

“

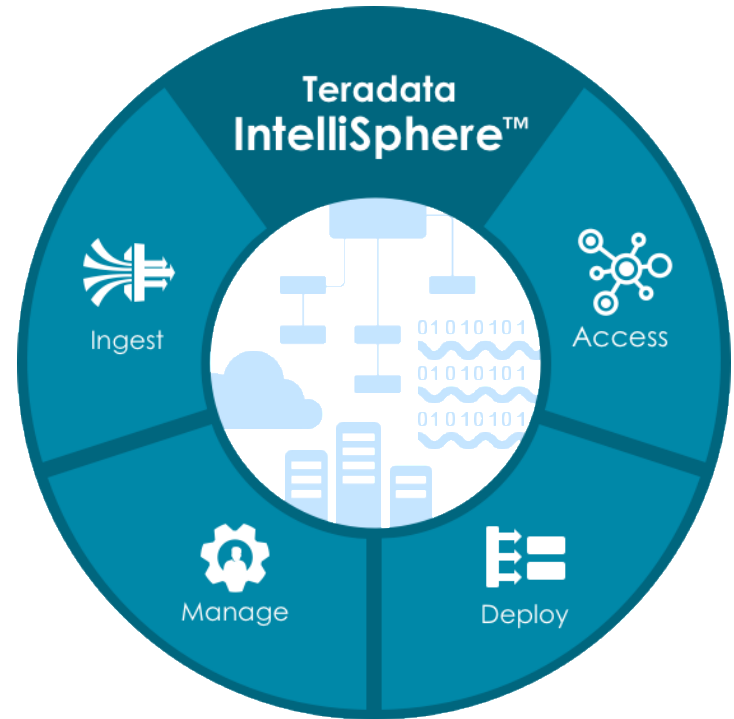
*It's extremely difficult to understand what products to buy that create value for me, especially around the data movement products because there are so many to choose and understand*

”

Global Chip Manufacturer

# Teradata's Comprehensive Software Portfolio that Enables the Orchestration of an Analytic Ecosystem

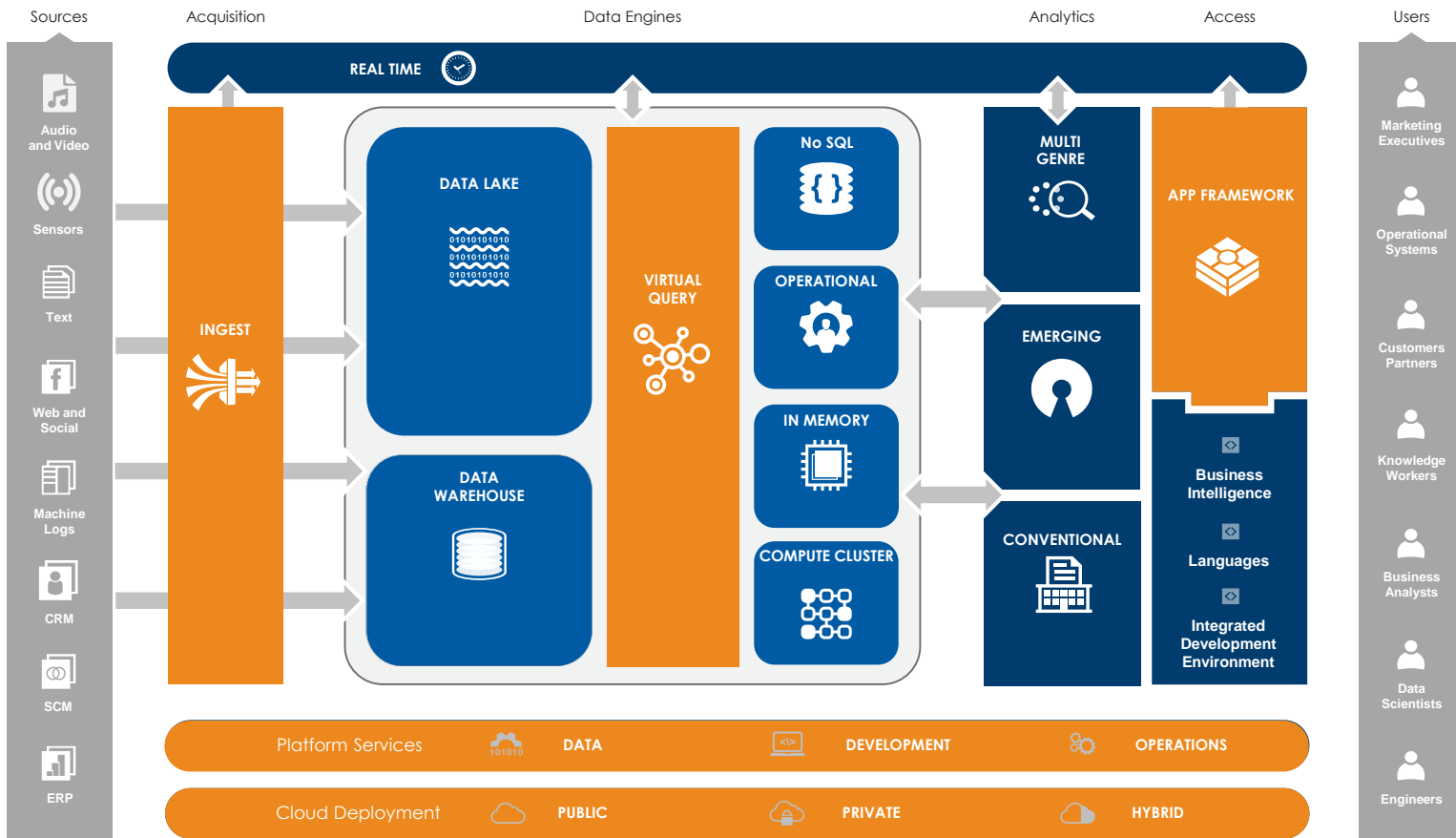
- Key enabler of the **Teradata Everywhere** strategy
- Core software required to **Ingest, Access, Deploy, and Manage** a Unified Data Architecture
- **One simple, subscription-based software license bundle**



# Key Takeaways of Teradata IntelliSphere

- IntelliSphere™ offers capabilities for complete solutions to solve challenges **through apps and use cases**.
- The Teradata IntelliSphere portfolio is a **single purchase** that covers **all deployments** within your organization.
- Built on a **modern stack** that provides **evolving capabilities** through **continuous innovation**.

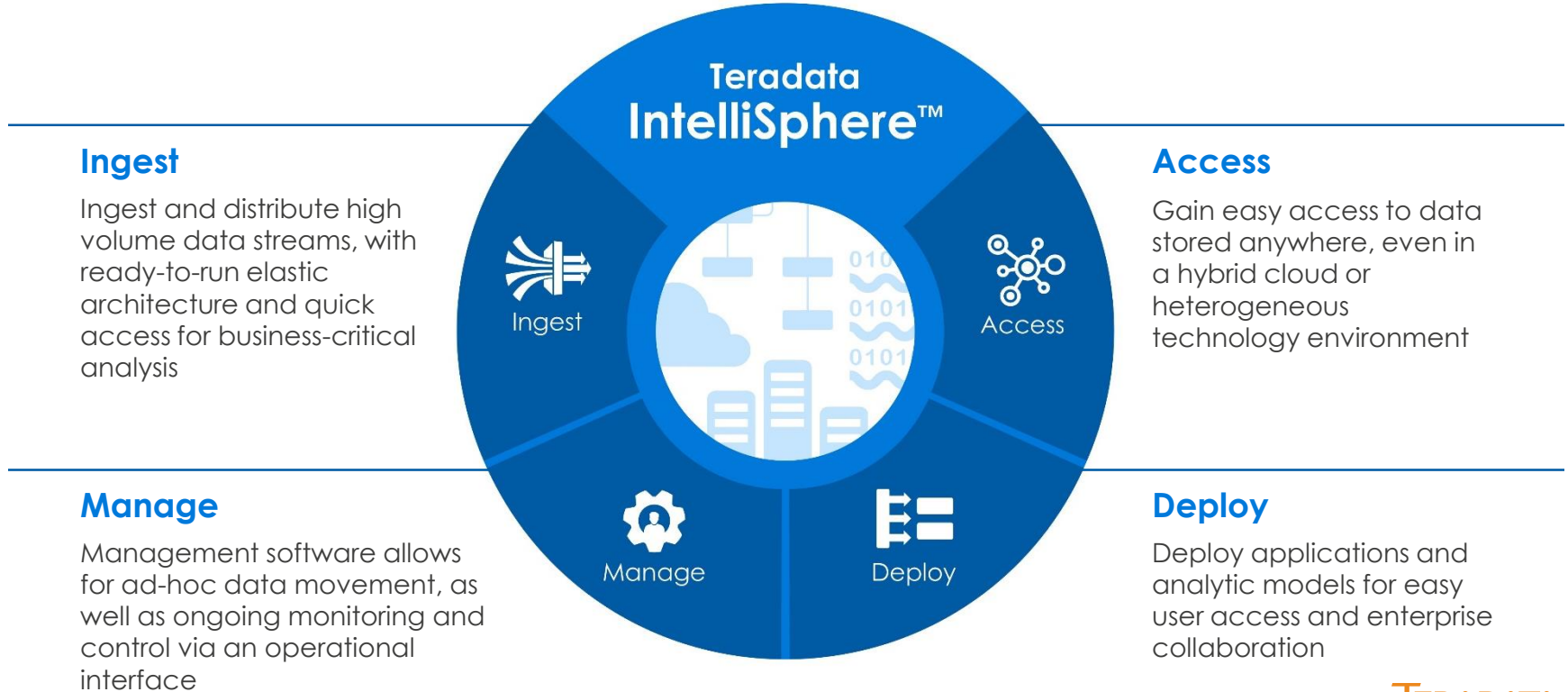
# We know that no single analytic system can meet all customers' needs





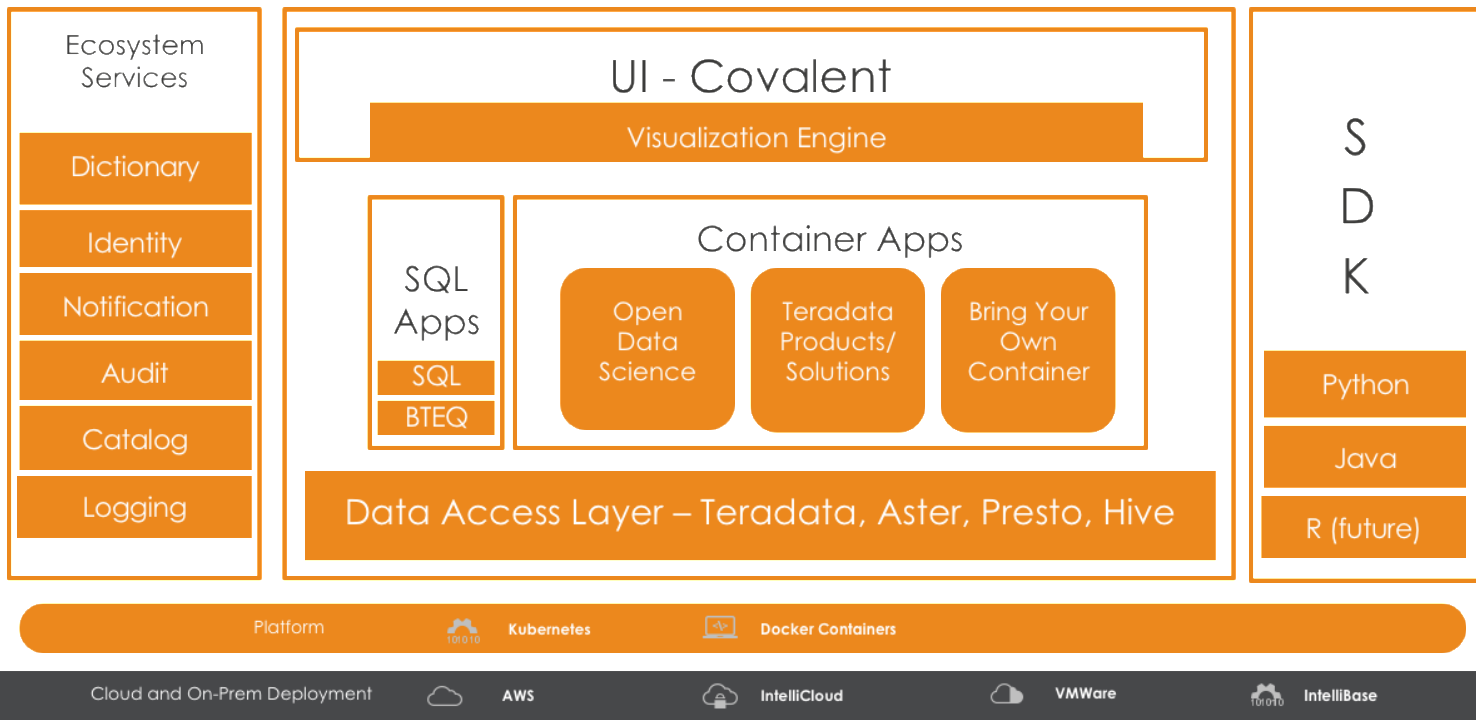
# IntelliSphere Unlocks Several Key Capabilities

All included in one simple subscription-based software license



# Modern Stack

## Teradata AppCenter



# Built On Key Use Cases

Data Lab

Disaster  
Recovery

Dual  
Active/High  
Availability

Dual Systems/  
Workload  
Balancing

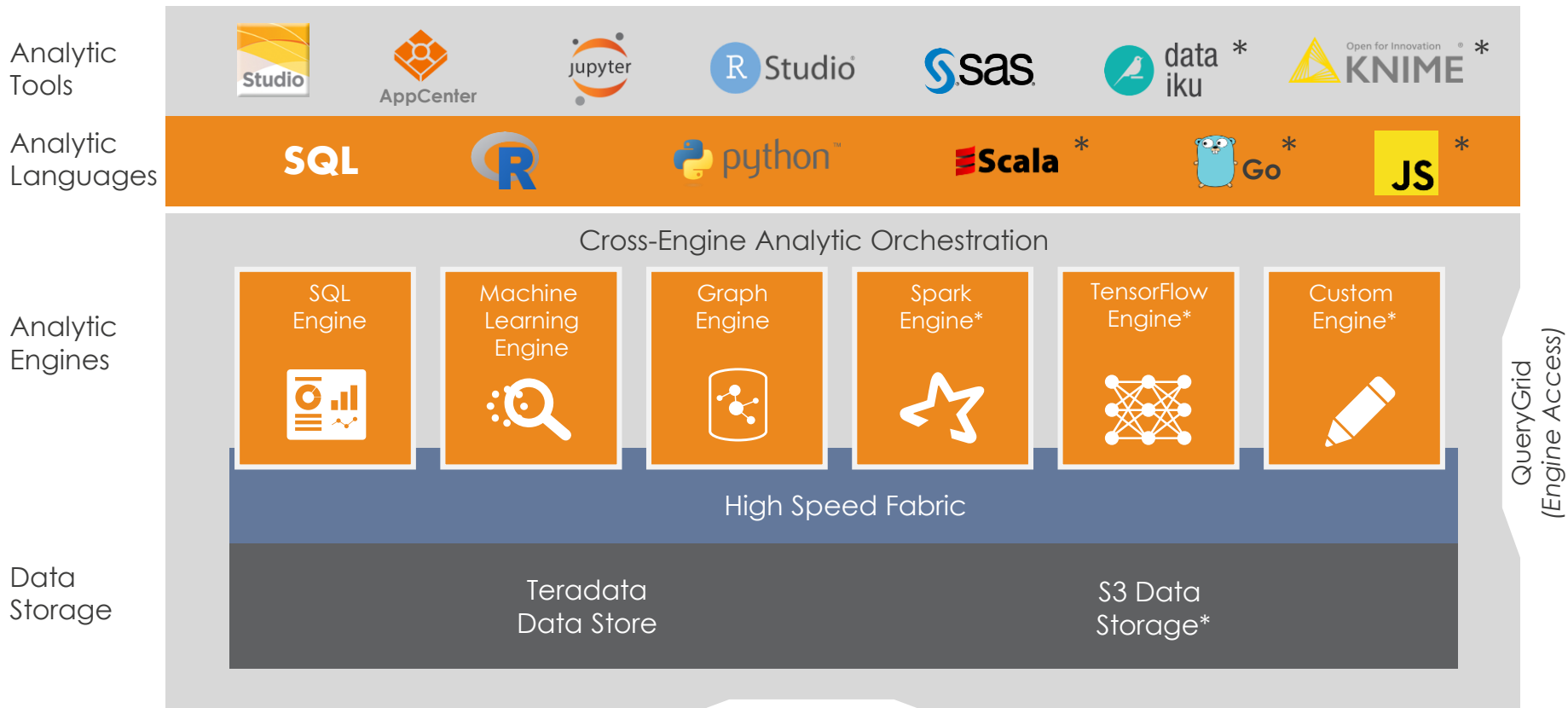
Self-Service Data  
Ingest for  
Analytics

Borderless  
Analytic Apps



**Future Direction**

# Teradata Analytics Platform Future Vision



\* Anticipated future capabilities

QueryGrid  
(Data Store Access)

TERADATA®