Simplifying and Accelerating Data Analytics
TCS Connected Intelligence Platform

BACKGROUND

Today’s businesses are inundated with customer, product, and operational data from internal and external sources, including streaming, IoT, and real time data from sensors, devices, mobile apps and enterprise data centers. As the volume, variety, and velocity of data continue to increase, businesses must find new and better ways to harness the value of the insights hiding in their data, as well as develop easier and more cost-effective ways to access, store, and manage data.

Across all industries, organizations are discovering how data lakes reduce strain on data warehouses, as well as how data analytics can power their business models.

To accelerate and simplify these initiatives, the TCS Digital Software & Solutions Group developed the TCS Connected Intelligence Platform (CIP) – a scalable enterprise insights platform with unified data management that enables business and technical stakeholders to harness and monetize all their data to gain a competitive advantage faster, and at a lower cost than any other approach.
CIP delivers the connected intelligence required for the actionable, predictive and prescriptive insights that have become essential to compete in today’s world. With CIP, business stakeholders and IT organizations can simplify and accelerate their data and analytics initiative across customers, products, and operations while gaining a robust and scalable platform to handle tomorrow’s use cases as well.

Overview

The TCS Connected Intelligence Platform is a scalable enterprise insights platform that accelerates time to insights by providing low code self-service capabilities. It is pre-integrated with unified big data management, stream processing and an AI/ML engine that fulfills multiple customer needs including data lakes, data warehouse, data wrangling and advanced analytic modeling.

Organizations can utilize these connected insights to create new strategies that meet the ever-changing needs of their businesses. These transformative insights are designed to break through the physical and digital barriers between products, customers, and operations. Organizations in all industries can use CIP to fast track their digital transformations.

CIP accelerates development of data analytics solutions and supports rapid deployment of analytical applications and use cases. CIP streamlines end-to-end insight application development with a proven step-by-step process.

CIP bundles traditionally separate technologies and tools into a single unified platform for organizations to collect, process and store big data, as well as derive insights using machine learning libraries and data visualization, and deliver actionable insights using business rules and predictive scores. Built-in data management and transformation capabilities means less time is spent integrating data and more time is spent innovating.

Enterprise-grade security and control help ensure compliance with corporate security standards and regulatory environments. CIP is an enterprise-grade data and analytics foundation that fuels the TCS Connected Intelligence & Insights family of software for Retail, Banking and Communications as well as the TCS Intelligent Urban Exchange.

Key Features and Capabilities

- **Data Modeling**: Model and activate data using Excel-based template
- **Enterprise-grade Hadoop**: Ingest, model, store and manage all data
- **Data Pipeline Processing**: Drag & drop user interface
- **Dataset & Services**: SQL interface to expose the data from both analytics & operational databases through REST APIs
- **Data Visualization**: Build reports and dashboards using low-code visualization framework or use existing visualization tools
- **ML-based Insights**: Multiple options to operationalize analytic models such as PMML based ML model, Python based ML, and deep learning
- **Decision Management**: Built-in business rules and predictive analytic models scoring engine
- **Workflow Automation**: Schedule data jobs for periodic execution
- **Project Workspace**: Build and manage multiple projects with controlled user access
- **Centralized Administration**: Web-based console with integrated user role management and Hadoop management.
Use Cases

CIP helps companies accelerate big data and analytics initiatives, such as:

**Data Lake Management**: CIP helps organizations build, manage, and deploy an enterprise or departmental Data Lake using the following platform features:

- Open source Hadoop distribution to provision, manage and monitor the Hadoop clusters with enhanced security management.
- Built-in batch and real-time data connectors to automate the data ingestion from multiple data sources into Hadoop.
- A common drag & drop data pipeline canvas to design batch and real-time data pipeline jobs and automate the jobs for periodic and ad-hoc execution to prepare and curate data for data analysis.
- Operational and analytic data stores to store and manage different kinds of data requirements.
- SQL interface to expose the curated data as Datasets for data analysis and data visualization.

**Data Science**: CIP enables organizations to derive and deliver deeper insights from raw data using the following platform features:

- Drag & drop data pipeline canvas to integrate data from multiple data sources and prepare the data for analytic model training.
- Leverage the open source Apache Zeppelin notebook or any other existing data science notebook to build and train machine learning models using the prepared and curated data.
- Machine learning algorithms exposed as XML based configuration to train models using highly scalable Apache SparkML engine.
- Deploy the trained machine learning models into production in batch and real-time model scoring mode.

**Streaming Analytics**: CIP enables organizations to derive and deliver operational insights from real-time or streaming data using the following platform features:

- Industry standard connectors to capture and integrate data from real-time data sources.
- Drag and drop data pipeline canvas to design and automate real-time data pipelines to deliver operational insights.
- Highly scalable database to store and access real-time data.
- Operationalize the analytic models in streaming analytic data pipelines.
- Data visualization framework to develop visually appealing operational reports and dashboards.

**Augment Enterprise Data Warehouse (EDW)**: CIP’s highly scalable and flexible architecture can integrate and augment the existing EDW environments to:

- Offload or migrate the detailed transactional level data from existing data warehouse environments into cost-effective and a highly scalable Hadoop environment. This helps reduce the workload on the existing EDW into cost effective Hadoop.
- Offload or migrate the time-consuming data processing jobs on scalable architecture to reduce time to insights and integrate the processed data back into the EDW environment.
- Augment the EDW environment with CIP’s real-time databases to store operational data and insights for operational decision making.
CIP’s drag & drop Data Pipeline Canvas allows users to design the data processing flows using drag & drop canvas.

CIP enables users to import PMML and catalog the models.

CIP allows platform administrators to create and manage project workspaces.

**Business Benefits**

CIP helps organizations accelerate the development and deployment of connected intelligence solutions with an all-data and analytics architecture. CIP is the backbone of TCS Customer Intelligence & Insights products for Retail, Banking and Communications as well as TCS Intelligent Urban Exchange software for smarter cities.

- Enables organizations large and small to leverage all available data to understand their customers
- Accelerates time to value for big data, data lake and data science related efforts
- Reduces the burden on your Enterprise Data Warehouse by migrating or off-loading transaction data and processing jobs to a simplified Data Lake
- Taps into IoT to broaden your data perspective and make faster, better business decisions that improve customer experiences
- Uses machine learning and data science to generate insights in real time from streaming and IoT data
Value Factory Model Approach

Build connected insight applications to turn data into actionable insights for accelerated business value:

1. Data Modeling
   - Define the data model for domain subject areas through Excel-based template or reverse engineer the data model invocation from existing databases.

2. Operational and Analytic Data Stores
   - Physically locate the data model on the operational & analytic databases for both real-time and analytic data needs.

3. Data Source Catalog
   - User interface to catalog the batch and real-time data feeds to collect data from diverse data sources.

4. Data Pipeline Processing
   - Data pipeline canvas to design batch and real-time data pipelines i.e. to merge, blend, transform and store data with in-built scheduler to automate the jobs.

5. Data Set Catalog
   - SQL interface to expose the data from both operational and analytic data stores for visualization, analytics, other applications.

6. Data Visualization
   - Visualization framework to develop canned reports & dashboards using collection of graphical charting library.

7. Data Science
   - Prepare data, train ML algorithms and make analytic models into production for both batch or real-time scoring.

8. Decision Management
   - Configure & automate the decision moments using the in-built business rules engine & real-time predictive or ML model scoring.

9. Workflow Automation
   - Scheduler to define all the data pipeline jobs for periodic execution, with job monitoring and management.

10. Administration
    - Provision project workspace; self-service platform users and define access control to various platform modules; secure & protect data.

Functional Architecture

TCS CONNECTED INTELLIGENCE PLATFORM

DATA & INSIGHTS ACTIVATION
- Data & Insights APIs
- Visualization
- Decision Management

DATA LAKE ANALYTICS
- Machine Learning & Model Scoring
- Streaming Analytics

DATA LAKE MANAGEMENT
- Data Modeling
- Data Pipeline
- Dataset Catalog
- Data Privacy

ADMINISTRATION
- User Role Management
- Workflow Automation
- Logging & Audit Management
- Metadata Management
- Apache Ambari & Ranger

HTTP/S | MOTT | IMS | CSV | XML | JSON | OData2

HADOOP DISTRIBUTION
- Open Source
- Cloudera
- Hortonworks
- AWS
- Azure
About TCS Digital Software & Solutions (DS&S) Group

TCS Digital Software & Solutions Group delivers on the promise of Connected Consumer Intelligence. Our experience working with the world’s most successful enterprises drives the development of integrated software that helps them meet the higher expectations of today’s wired consumers and citizens. With TCS DS&S Group software, organizations can design experiences that logically connect every touch point of the consumer’s digital and physical journey.

To learn more
Visit: dss.tcs.com
Email: tcs.dssg@tcs.com

About Tata Consultancy Services Ltd (TCS)

Tata Consultancy Services is an IT services, consulting and business solutions organization that delivers real results to global business, ensuring a level of certainty no other firm can match. TCS offers a consulting-led, integrated portfolio of IT and IT-enabled infrastructure, engineering TM and assurance services. This is delivered through its unique Global Network Delivery Model, recognized as the benchmark of excellence in software development. A part of the Tata Group, India’s largest industrial conglomerate, TCS has a global footprint and is listed on the National Stock Exchange and Bombay Stock Exchange in India.

Extensive industry experience

All content / information present here is the exclusive property of Tata Consultancy Services Limited (TCS). The content / information contained here is correct at the time of publishing. No material from here may be copied, modified, reproduced, republished, uploaded, transmitted, posted or distributed in any form without prior written permission from TCS. Unauthorized use of the content / information appearing here may violate copyright, trademark and other applicable laws, and could result in criminal or civil penalties.

Copyright © 2020 Tata Consultancy Services Limited