LIGADATA

Friday, December 19, 2014

# OnLEP 1.0 Release Notes

## General Notes

- Please be sure you are using the latest version of the COPD reference model from the RTD ReadOnly folder, as it has updates. You may want to compare any models you are testing against this, so that you gain the benefit of those updates.

## 1.0 Install and setup

- Setup Guide - Revised to a single guide including both general and Metadata.API setup. Expanded Config file information. Added multi-node setup
- The Engine and the MetadataAPI both use a single properties (Config) file
- Config file now takes multiple jar paths
- Developer Tools
    - Instructions for building UDF libraries and extracting metadata from them added to "white paper". White paper is in the trunk/Docs folder. The Scala script is found in trunk/Pmml/Scripts/extractUdfLibMetadata project with 'how to' documentation found in the README.md file there
    - SBT dependency extraction tool (see trunk/Pmml/Scripts/sbtProjDependencies project) for pulling project dependencies from sbt projects into multiple formats, including classpath, comma delimited list, et al. See the extractUdfLibMetadata script for example use.
    - Several bug fixes made to the MethodExtractor component that generates the UDF JSON metadata for the d script.

## 2.0 Engine

- Multi-Node functionality added including ability to scale distributed processing
- Added support for primary and partition keys
- Exactly-Once Functionality for each input message – so that processing occurs only once, and output is sent only once.
- Dimensions & messages are "set-and-get" using EnvContext
- Built jar libraries for custom UDFs

## 3.0 Storage

- Node/Cluster/Adapter information is persisted into persistent store
- Support for HBase and Redis has been added
- Dimensions and messages are stored in the single table in EnvContext

## 4.0 Compilers

### 4.1 Message and Container Compiler

- Mapped Messages and Containers support for fields, BaseType, Arrays, ArrayBuffer, Message and Container
- Added Unique Key feature, Add Message and Get Message feature of both Fixed and Mapped message to support Engine.
- Added PartitionKeys and PrimaryKeys to both Fixed and Mapped Message Definition objects
- Utility to generate the message definition json data from CSV data (WILLIAM)

### 4.2 PMML Compiler

- Mapped message support for sparse data
- Integrated transaction use into model framework
- Aggregations added to COPDv1.xml reference model
- Guidance for group by and aggregation for the COPD model added to the PMML Model Whitepaper.
- Extended Compiler, built new UDFs & new types to support aggregation for the COPD model
- Provided examples for SUM in aggregation

## 5.0 Metadata.API

- JarFiles are also persisted into metadata store which enables propagation of jar libraries (corresponding to new models or messages) to every node at run-time
- Added Metadata.API functions to support storage persist
- Added handling for metadata updates to engine sent from Metadata.API through zookeeper

## 7.0 Documentation changes

Added: Completely revised Install & Setup Guide, combining into single guide, clarifying Config file use, adding multi-node setup
Added: Change Summary 12/19/2014
Added: API Documentation at 12/19/2014
Added: Hardware Technical Specifications