**Universal Big Data ETL Platform**

**Introduction**

The Big Data ETL platform leverages the inherent fault tolerant and reliable big data infrastructure (Hadoop and Spark) to extract, transform and loading any complex data types such (binary, XML, JSON, Comma Separated Value (CSV) files) for Big Data analytics processing.

Enables users to define the whole ETL process totally from business logic aspect, leaving the complex big data related technical implementation and optimization to the platform.

This unique solution eliminates the difficult Hadoop learning curve by providing out of the box configurable ETL commands, including but not limited to transformation and aggregation, file loading, XML processing, data backup, database load, and shell script execution. Using these building blocks, expensive and rare Hadoop expertise is not required – complete ETL flows can be developed by any network engineer.

The Big Data ETL platform converts the configuration flows into runtime big data jobs (Hadoop MapReduce Jobs/Spark Jobs) taking full advantage of the Hadoop/Spark built-in distributed processing, resource mangement and scheduling functions.

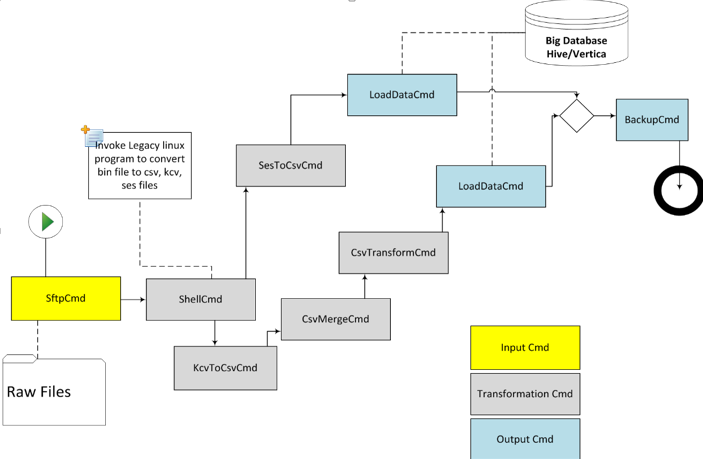
The key benefits of this solution are greatly reduced delivery time, cost, and a solution based on an industry accepted open source platform. In recent deliveries to a Tier 1 North American carrier, HPE was able to deliver an analytics solution in quarter the time and at one quarter of the cost of the competition using the Universal Big Data ETL platform and has mandated the HPE solution be used for all future network element analytics.

**Business Need**

Entering the Big Data world, there are two primary issues. Firstly, there is a large number of different types and formats of data that need to be analyzed. Secondly, Terabytes to Petabytes of data need to be connected in real-time and loaded and then analyzed. The Universal Big Data ETL platform was built to solve these industry constraints. The solution enables users to define his/her data processing flow by using a library of commands (different protocol connectors, various processing commands, etc.) and easy to follow graphical tools. Without the specific Big Data technology background, users can still define the flow with ease to be executed on a cluster of hundreds and thousands of nodes to process Terabytes of data. The Universal nature of data ingestion capabilities from the big data ecosystem (Hadoop/Spark/etc) provides the promise that any Telco data source ever created can be transformed by this platform.

## Graphical Flow

This graphical flow designer describes how…. The Universal Big Data ETL requires no Hadoop expertise given that is a configuration based platform allowing for out of the box tools such as transformation, aggregation, file loading, XML data processing, data backup/ load, and shell script execution. These tools give the platform a high production grade of processing close to 700 million records per day. No limitation on the number of records can be processed, this scales linearly with the hardware.



**Use Cases**

Femtocell Analytics

The Femto Access Point is a Samsung hardware device located the customer’s premise that interfaces with mobile devices over the air radio interface. A Tier 1 US Communication Service Provider has contracted with HPE to provide Femtocell analytics using the HPE SPS analytics system. The Analytics reports include

Location Based Services

The Universal Big Data ETL system can convert PDE binary files and create analytics such as identifying M2M devices that are generating the greatest amount of traffic and correlate additional data to include device manufacturers and serial numbers

Virtualized Mobile Terminated Core Nokia Analytics

A Tier 1 US CSP is deploying new Mobile Terminated Call (MTC) Nokia virtual equipment, the SGS-IWF and vSMSC. Nokia could not provide the KPI’s and data to integrate to customer’s Operations reporting environment. Instead, the carrier has contracted with HPE to provide the analytics for these components. The Universal Big Data ETL system will convert XML input to CSV format

SPC Device Enablement Gateway Analytics

A US carrier has deployed the Subscriber Profile Controller (SPC) to provide subscriber authentication and application authorization of 4G services hosted by ASPs. The SPC Management Complex uses data from the SPC to provide robust reports with graphical displays that are interactive and allow the user flexibility to view the report based on their needs. The CSP has contracted with HPE to migrate all of their SPC Analytics from the current Oracle system to the SPS. The SPS system will provide a 40 X improvement in report performance.