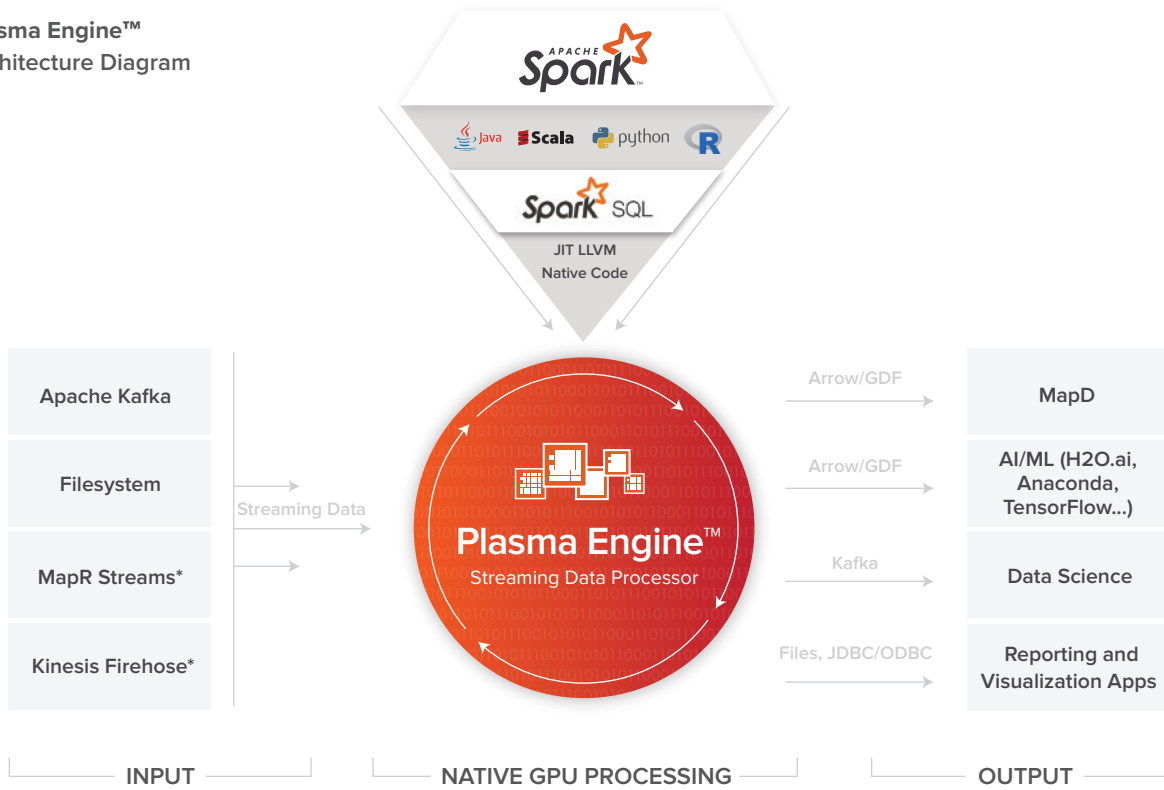


Plasma Engine is the first GPU-native software to fully leverage the massively parallel processing power of NVIDIA GPUs for real-time stream processing of data in motion. Executing over 80% of its operations on GPUs, Plasma Engine runs natively on Apache Arrow, a vectorized columnar in-memory data format that is optimized for analytical data processing.

Plasma Engine™ Architecture Diagram



PROGRAMMING LANGUAGES

Plasma Engine supports any programming languages that can be used with Apache Spark.

- ✓ Java
- ✓ Scala
- ✓ SQL
- ✓ Python*
- ✓ R*
- ✓ CUDA*

DATA FORMATS

- ✓ CSV
- ✓ Apache Arrow
- ✓ Apache Avro
- ✓ JSON*
- ✓ Syslog*
- ✓ Protobuf*
- ✓ XML*
- ✓ Apache Parquet*
- ✓ Apache ORC*

MINIMUM SYSTEM REQUIREMENTS

- GNU/Linux x86_64 with kernel version > 3.10 | Ubuntu 16.04 | Docker >= 1.12
- NVIDIA GPU architecture required: Kepler, Maxwell, Pascal, Volta or above
- NVIDIA drivers 390.x and above

STRUCTURED STREAMING

Plasma Engine executes most Apache Spark Structured Streaming features natively on the GPU.

- ✓ Kafka Input Source
- ✓ File Input Source
- ✓ SQL Window Functions
- ✓ Aggregate Window Functions with Watermarks
- ✓ Basic SQL (*selection, projection, filtering, aggregation*)
- ✓ Managing and Monitoring Streaming Queries
- ✓ Recovering from Failures with Checkpointing
- ✓ Output Mode: append, update, complete
- ✓ Joins
- ✓ ORDER BY
- ✓ Streaming De-duplication*
- ✓ Arbitrary Stateful Operations*

DATAFRAMES, DATASETS & SQL

Plasma Engine executes Apache Spark DataFrame/Dataset API and SQL natively on the GPU.

- ✓ Create Dataframe from file
- ✓ Running SQL queries programmatically
- ✓ User-defined functions*
- ✓ Null-value
- ✓ Global Temporary View
- ✓ User defined aggregation functions*
- ✓ Unions
- ✓ Relational, arithmetic, logical operators, mathematical functions
- ✓ User defined serialization formats*
- ✓ Caching Data In Memory (columnar)
- ✓ Generic Load/Save Functions*

SQL DATATYPES

- ✓ TINYINT
- ✓ SMALLINT
- ✓ INT
- ✓ BIGINT
- ✓ BOOLEAN
- ✓ FLOAT
- ✓ DOUBLE
- ✓ STRING
- ✓ BINARY
- ✓ TIMESTAMP
- ✓ DATE
- ✓ STRUCT

FAULT-TOLERANCE

Plasma Engine supports the same end-to-end exactly-once fault-tolerance semantics as Apache Spark.