The Sigma^{X} Stack

Accelerated Edge Data Ingest

Apache Arrow

Apache Pulsar

Heterogeneous Processor Support

Big Data Scale

Connected, Cutting-edge, Open Source
Data Transport and Processing

SigmaX.ai

React Faster

Process More

Open Source Big Data Edge Ingest Software

Ingest structured data from millions of edge data sources with **higher throughput**, **lower latency**, and **higher efficiency** than ever before.

- Apache Arrow: Support for hardware accelerated data format coercion at ingest
- Apache Pulsar: Support for Intel OPTANE accelerated performance
- GPU Engine for Apache Pulsar

Eliminate:

- Multiple copies of data
- AVRO data serialization / de-serialization
- Different in-memory and durable data formats
- Exaggerated cache misses
- Choosing between fast read speeds or fast write speeds

Data Engineered

- Data from Edge to AI+ML
- Easily lever Hardware Acceleration technology
  - Pulsar -to- FPGA
  - Pulsar -to- GPU
- Data to AI+ML clients independent of locality

Stack Elements

APACHE ARROW

- Columnar in-memory data / Ingest
- Apache Arrow$^{\Sigma}$
- Apache Arrow Flight$^{\Sigma+}$
- Arrow Dive$^{\Sigma+}$

PULSAR

- Publish & Subscribe messaging system$^{\Sigma}$
- Real time event stream processing$^{\Sigma}$
- GPU Engine$^{\Sigma+}$

presto

- Easy to use SQL layer for Pulsar query

THE APACHE SOFTWARE FOUNDATION

- HDFS support
- Unstructured data support

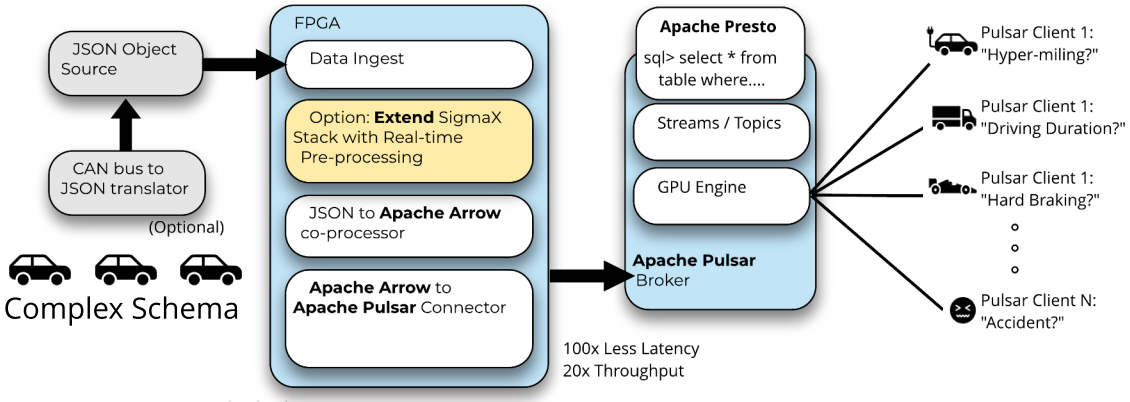
Σ Hardware Acceleration Supported Now

Σ^{+} Future Hardware Acceleration Support

www.sigmax.ai

Faster answers from AI+ML applications with cutting-edge Data Engineered software

Example Dataflow



- Write durably to storage
- Transmit over networks without the need for de-serialization at the point of receipt (a key difference when compared with Avro or Parquet)
- Stream in record batches of structured data rows
- Read with zero copies

- Native data geo-replication
- Larger message sizes
- Streaming and queuing
- Pub/sub native
- **Tiered persistent storage native**
- Linear horizontal scale-out
- **Multi-tenant**
- **Cloud-native & Multi-cloud**
- Real-time streaming
- Distributed
- Stateless
- End-to-end encryption

Insurance by-the-mile example:

Insurer needs to answer questions about their customer's driving behavior.

Millions of sensors stream JSON data to FOG ingest. FPGAs transform data in real-time to Apache Arrow.

Latency is able to keep up with, for example 5G wireless network performance.

Data is connected to Apache Pulsar and placed into streams and topics. Bandwidth enhanced with hardware acceleration to support high throughput and complex schemas.

Sigma^X Partners

Based in and Supported from the USA

Sigma^X is based in Virginia and has made significant investment in US based development and support.

Contact Us!

Sigma^X is a US company.

Telephone: (443) 717-2522

Email: info@sigmax.ai

or visit us at https://www.sigmax.ai