Solution Brief

Open FPGA Stack

Deploying a Real-Time Data Management Solution Using Intel® FPGA-based Boards and Fully Open-Source Code

Open FPGA Stack (OFS) and Intel FPGA-powered acceleration boards provide the infrastructure needed to expedite the development of FPGA-based workloads

Overview

- SigmaX develops an incredibly fast end-to-end data flow from data ingestion to the consumer, where decisions can be made in near real time
- SigmaX extends and accelerates the Apache open-source software using Intel® Xeon® processors, Intel Agilex® FPGAs, and Open FPGA Stack (OFS)
- SigmaX achieves 100X lower latency and 20X increased data ingestion with FPGA-based acceleration compared to a CPU-only based approach

Executive Summary

SigmaX tackles the latest challenges in solving enterprise data management problems at scale. Customers benefit from their data management solution without the obligation of vendor lock-in due to the use of open-source Apache software and FPGA development resources such as Open FPGA Stack (OFS). The SigmaX solution powered by FPGA acceleration enables data to flow at incredibly fast rates, enabling users to make decisions in near real time.

- The SigmaX data flow solution, based on Apache Pulsar* and Apache Arrow*, reduces latency by 100X[1]
- Using Apache Pulsar alone, SigmaX benchmarks at an immediate 250% speed increase for streaming vs. the tcompetition using Apache Kafka*[1]
- The integration of Apache Pulsar with Apache Arrow results in 20X increased throughput, scalable to thousands of nodes[1]

Background and Challenge

A data broker is a collection of event-streaming data, either public, private, or both, served as subscriptions. The data broker will process, clean, and structure published data and serve the data to other businesses or consumers within the business. A data producer refers to the root source of data, whether it be a user interface, service, or edge and Internet of Things (IoT) device. Millions of data producers can concurrently send information to data brokers. Consumers can then retrieve recent data from the data broker once it has been processed, cleaned, and structured. These data brokers can scale up to thousands, processing immense volumes of data in parallel. Industries such as 5G, autonomous vehicles, predictive maintenance, and other edge computing and transport platforms are tackling these increasingly larger data sets that can scale to thousands of data brokers and producers.

Authors

Robert Morrow
Founder and CEO of SigmaX

Tamara Lin
Product Marketing Specialist
Intel Programmable Solutions Group

In these industries, making decisions at real-world speeds and reacting almost simultaneously is often crucial.. The autonomous vehicle industry alone has been estimated to generate anywhere from 4 TB up to 40 TB of data per hour. Alongside this huge demand in data processing, new types of data structures and forms of data representation have also emerged, with untapped performance advances in CPU, GPUs, and parallel processing.

Apache Arrow is a standard language-agnostic software framework commonly used to improve the speed of data analytics by creating a standard columnar memory format at all wall-clock savings of 80%. It is frequently used in contexts with large data sets generated by sensors at the edge, IoT, and large-scale applications. Apache Arrow combines the benefits of columnar data structures with in-memory computing that CPUs, GPUs, and FPGAs can use to enable seamless and efficient data sharing across platforms without copying or moving data.

By taking advantage of heterogenous processing, along with open-source tools such as Apache software and OFS, SigmaX delivers a solution that improves the conversion of data formats from JSON to Apache Arrow at 100X lower latencies and 20X increased data ingestion rates when compared to scaling out using Intel Xeon processors alone.

Solution

SigmaX has significantly improved the efficiency of data being produced to a central broker by coupling their data management stack with Intel FPGAs and OFS. The following is a breakdown of the SigmaX solution:

Step 1: Bolson converts sensor data to the Apache Arrow format

JSON sensor data is first received by Bolson running on an Intel Agilex FPGA using the open-source OFS infrastructure. Bolson then converts the JSON sensor data to the universal data format, Apache Arrow. By placing data into Apache Arrow, the brokers become faster and more responsive by orders of magnitude. Thus, this pathway stacks latency and throughput benefits on top of the benefits of an only CPU-based approach.

Step 2: Apache Pulsar processes and cleans data

The data broker, Apache Pulsar, will then receive the messages in the Apache Arrow format. Data received by Apache Pulsar is now computable as is. Apache Pulsar will then process, clean, and restructure the data.

Apache Pulsar is a distributed messaging and streaming platform comparable to Apache Kafka. However, it provides significant benefits compared to Apache Kafka, such as improved security, speed, latency, and performance with built-in data reliability features such as geo-replication. It is commonly used in latency-sensitive applications involving high-complexity schemas or real-time data needs.

Step 3: The data broker transfers information to the consumer

After delegation from the broker, data is then transferred to a subscribed consumer application.

SigmaX has validated this data management workload using Intel technology-based hardware. Their open-source stack uses Intel Xeon processors running in the client application and the Hitek Systems HiPrAcc* NC100 board based on an Intel Agilex FPGA to run Bolson.

The HiPrAcc NC100 board is enabled with both OFS and oneAPI. OFS is a key foundational tool that enables FPGA developers to build custom FPGA-based workloads and applications. OFS provides all the hardware and software source code, documentation, reference examples, and tools needed to jump-start FPGA-based development. The software and hardware code for OFS is currently open source on GitHub.

Getting Started with FPGA Acceleration Using OFS

FPGA developers can choose from a range of custom, Intel-provided, or third-party OFS-enabled boards and use the open-source documentation and source code to start building their custom workload.

The following table outlines how a developer can get started with FPGA-based workload development using either an Intel-provided board or a third-party board.

Product Line	Using an Intel board	Using an ecosystem board
Step 1: Choose a board	**Use an OFS reference platform** Reference platforms can expedite evaluation or bring-up but are not required.	**Use a custom or third-party board** Browse the OFS board catalog to see available boards.
Step 2: Evaluate OFS open-source resources	Technical documentation can be found on GitHub.	The board vendor will provide the corresponding OFS technical documentation.
Step 3: Access open-source hardware and software code	Modify or use the provided OFS software and hardware code available on GitHub (OFS).	The board vendor will provide the corresponding OFS software or hardware code.
Step 4: Develop workload using RTL or C/C++ (using oneAPI)	Follow the OFS RTL flow OR OFS enables the compilation of oneAPI kernels. Utilize the oneAPI development flow and build FPGA workloads in C/C++.	

Quote

"OFS enabled us to create our FPGA-accelerated workload by providing all the hardware and software source code, documentation, reference examples, and tools we needed to get started – no in-depth FPGA tinkering required."
- Robert Morrow, CEO of SigmaX

Results

SigmaX's data management workload accelerates data conversion into Apache Arrow using two key Intel technologies – Intel Agilex FPGAs and OFS. Using FPGA acceleration and OFS, SigmaX's data conversion workload is 100X faster with 20X more data. This data management workload can be applied to a broad range of applications, including healthcare, insurance, 5G, predictive maintenance, and more.

Further Information

- Learn more about SigmaX
- Learn more about OFS
- Get OFS open-source code and documentation
- Learn more about the Hitek Systems HiPrAcc NC100 board
- Download the Hitek Systems Solution Brief

[1] Figures published in "Tens of gigabytes per second JSON-to-Arrow conversion with FPGA accelerators." IEEE Xplore. December 2021. ieeexplore.ieee.org/documents/9609833

Test configuration: Design of an FPGA accelerator for JSON parsing that writes the deserialized data into host memory in the Apache Arrow columnar in-memory format. Consists of five stages: Receive JSON documents, Parse JSON documents and deserialize data into an Arrow RecordBatch, Resize the Arrow RecordBatch, Serialize the Arrow RecordBatch into Arrow IPC message, Publish IPC messages to a Pulsar topic via a Pulsar broker. All implementations use a maximum of eight bytes, which gives each parser a peak theoretical input throughput of 1.6 GBps when running at 200 MHz.

Intel technologies may require enabled hardware, software or service activation.

No product or component can be absolutely secure.

Your costs and results may vary.