Machine Learning at the Edge

Introducing SWIM EDX

Everyone wants to discover hidden insights in their streaming data, and even better if those insights can predict future performance. But unless you're a data scientist comfortable standing up the infrastructure to collect, store, clean, move, and analyze never-ending streams of data, you're out of luck. And batch-style analysis and learning doesn't permit a real-time response.

SWIM EDX software gets you to insights fast – as fast as data arrives. SWIM EDX learns insights directly from noisy data, enabling users to respond in real-time. SWIM EDX learns on-the-fly, short-circuiting cloud-centric, batch-based, big-data analytics and learning. This short paper provides a quick technical overview of our solution.

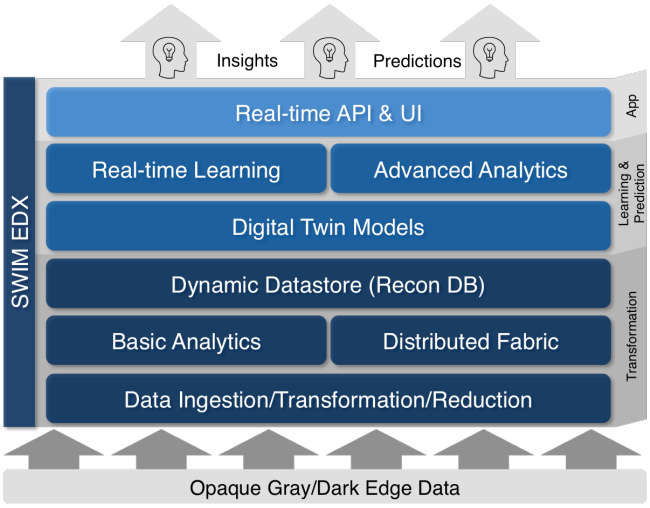

Figure 1. SWIM EDX transforms noisy device data into real-time insights using self-training machine learning at the edge.

Drowning in Data

Cloud vendors want to store your data and run your IoT, big-data, ML and analytics apps. That's fine for cloud-native apps like Twitter, but if you already operate infrastructure that delivers streams of data, a cloud-first architecture won't cut it. Bandwidth, storage & processing sound cheap, but costs add up fast. And despite the appeal of big-data analysis and machine learning services, data cleaning and labeling, and the development and training of models require deep domain expertise. On-prem solutions are no better. They demand substantial investment ahead of any return and saddle you with maintenance, and you still need a data-scientist with domain expertise to make sense of it all.

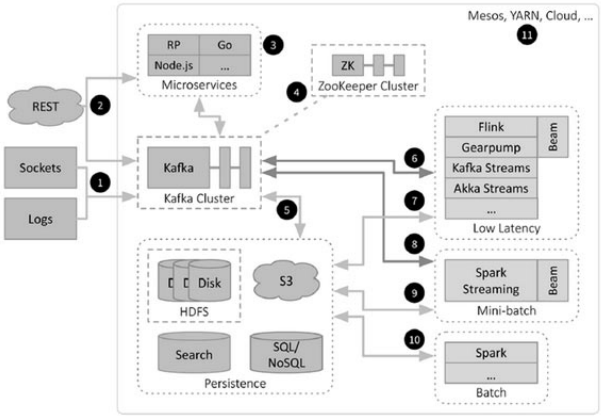

Figure 2. Cloud architectures are optimized for batch processing, making them inflexible and unsuited to real-time data.

SWIM EDX Edge Intelligence

SWIM EDX offers a powerful, disruptive approach to edge learning and analytics. It inverts the big-data "collect, clean, analyze and learn" cycle. Instead it learns on-the-fly as data is created, at the edge, on existing edge hardware. The are three core innovations underlying SWIM EDX: the Fabric, digital twins, and edge learning.

SWIM EDX: The Fabric

The SWIM EDX Fabric creates a single application plane, which enables different edge devices to share information with each other in real-time, regardless of location.

(continued on next page)

SWIM EDX: The Fabric (cont'd.)

- SWIM EDX builds an autonomous, resilient edge data processing fabric that spans edge devices. The Fabric supports the distributed actor framework for stateful, distributed, actor-based edge computing – and includes real-time analysis and learning. Actors are stateful "active" objects / services. The Fabric supports true parallelism, actor migration and replication, automatic load-balancing, and is resilient to failures.
- SWIM EDX hides the distribution of actors across the fabric: Actors interact with each other as though all are local. Communication is consumer driven, and back-pressure regulated. The fabric is eventually consistent, and state updates are efficiently gossiped between instances.

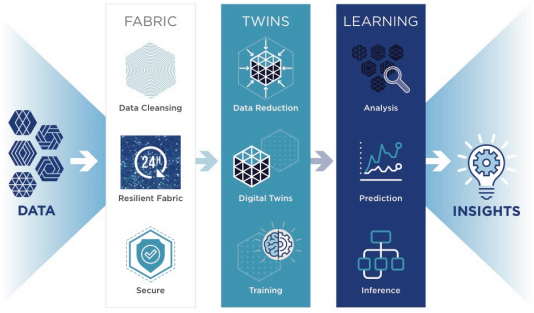

Figure 3. SWIM EDX transforms streaming edge data into digital twins, enabling real-time analysis and predictions.

SWIM EDX: Digital Twins

SWIM EDX uses the Fabric to automatically instantiate digital twins:

- SWIM EDX represents each real-world system (a data source) as a stateful actor that executes on the SWIM Fabric. Each actor is a stateful "digital twin" of the real-world system it represents.
- Digital twins are automatically detected in the data and dynamically instantiated. Each is a stateful representation of the system it represents. It analyzes and learns from its own time-series real-world data. It learns much as you do, observing the world, and continually re-computing its analysis and predictions.
- Each twin efficiently gossips its state to other twins via a pub-sub API. This allows digital twins to share data to solve problems that require data from multiple twins – effectively enabling a dynamic "join" on their combined state.
- Each digital twin offers a real-time API that makes it easy to visualize its state and the results of its analysis and its predictions. Application developers use these APIs at the digital twin level to quickly create custom logic for multi-actor joins to combine, correlate, analyze, or learn and predict from multiple twins.

SWIM EDX: Edge Learning

SWIM EDX Edge Learning enables each digital twin to learn and predict its behavior:

- Each digital twin consumes data from the real world, analyzes it using a rich set of analytical functions, and trains a deep learning model for the real-world system. Each sample is matched against predicted behavior, and errors are back-propagated to improve predictions.
- Edge-optimized implementations (for GPUs and CPUs) of the twin-centric deep learning algorithm gives SWIM EDX an almost 100x speedup over cloud-based ML, and is key to its success on low power edge devices. EDX learns a specific model for each twin, rather than a system-wide model. This has many advantages: The models are robust, contextually rich, and permit local responses in real-time.

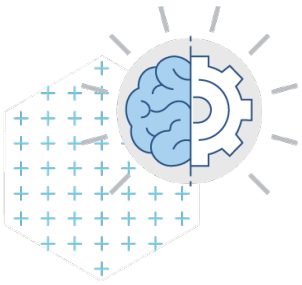

Figure 4. SWIM EDX learns a specific model for each digital twin, and then aggregates twins to represent entire systems.