THE POWER OF DIGITAL TWINS

Traditional stream-processing and complex event processing systems focus on extracting patterns from incoming data with stateless applications. While these applications maintain state information about the data stream itself, they don’t generally make use of information about the data sources or their context. This makes it much more difficult to fully analyze what incoming telemetry is saying. For example, an IoT predictive analytics application attempting to avoid an impending failure of a medical freezer must look at more than just trends in temperature readings. It needs to evaluate these readings in the context of each freezer’s operational history to get a complete picture of the freezer’s condition.

That’s where the power of the digital twin comes in. While digital twins have been used for several years in product life cycle management, their ability to enable stateful stream processing has only now been made possible by advances in scalable, in-memory computing. The digital twin model provides a simple, intuitive technique for organizing important, dynamically evolving, state information about each individual data source and using that information to analyze incoming event streams in real time. This enables deeper introspection than previously possible and leads to significantly more effective and timely feedback within operational environments.

Digital twins also offer compelling performance advantages. Most stream-processing platforms need to pull state information from remote data stores, such as database servers and distributed caches; this creates delays and network bottlenecks. Digital twins enable streaming events to be delivered directly to their associated state information for deep analysis — without delay.

Key advantages of the digital twin model include:

- Automatic correlation of incoming events by data source
- Immediate access to dynamic state information for each data source
- Intuitive, object-oriented encapsulation of real-time analytics for each type of data source
- Fast, real-time alerting combined with real-time, data-parallel analytics across all data sources

MANY APPLICATIONS

Digital twins can enhance the ability of any stream-processing application to analyze the dynamic behavior of its data sources. Here are just a few examples:

- **Financial Services**
  - Ex: Portfolio tracking, wire-fraud detection, stock back-testing
- **Internet of Things**
  - Ex: Device tracking for manufacturing, vehicles, mobile devices
- **Healthcare**
  - Ex: Real-time patient monitoring and alerting, health device tracking
- **Logistics**
  - Ex: Real-time inventory reconciliation, manufacturing flow optimization

...and much more
BRINGING THE DIGITAL TWIN TO LIFE

The ScaleOut Digital Twin Builder™ is a software toolkit for creating stateful stream-processing applications using the digital twin model and running them on ScaleOut StreamServer™. This toolkit defines APIs in Java and C# (with JavaScript support to be released soon) that let application developers easily define digital twin models for several types of data sources, deploy them, and then process event messages sent from data sources or messages sent by other digital twins.

Using the ScaleOut Digital Twin Builder, developers construct a digital twin model for each type of data source using well understood, object-oriented techniques. A digital twin model consists of a data object that describes the state information to be maintained for each unique data source and a message-processing method, which encapsulates the application’s real-time analytics for incoming messages, as illustrated in the following diagram depicting an instance of a digital twin:

CONNECTIVITY TO MANY DATA SOURCES

The ScaleOut Digital Twin Builder includes APIs for connecting ScaleOut StreamServer to numerous message sources, including Microsoft Azure IoT Hub, Kafka, and a REST web service, with more connectors to be released soon. In addition, Java and C# applications can directly send messages to digital twins to assist in development and testing.

PERFORMANCE ADVANTAGES

With its highly scalable in-memory data grid and fast event-processing, ScaleOut StreamServer offers an ideal platform for hosting digital twins and performing stateful streaming analytics. Algorithms that previously only could be run offline in batch processing can be executed in real time, dramatically improving the quality of feedback that can be provided to operational systems. Here are some of ScaleOut StreamServer’s unique capabilities for delivering high performance:

- Scalable, highly available in-memory data grid capable of hosting millions of digital twin objects
- Fast message delivery to digital twins using ReactiveX observers integrated into the grid
- Automatic message correlation by data source for analysis within digital twins
- Immediate access to digital twin state objects by message-processing code without network bottlenecks or database access delays
- APIs for data-parallel analysis across all digital twins in real-time
- Scalable message-processing throughput of more than 8K messages/second per server with message-processing latencies under 4 milliseconds