ScaleOut Software

ScaleOut Digital Twins™

Overview

March 2024

Dr. William Bain, Founder & CEO,
wbain@scaleoutsoftware.com

About ScaleOut Software

ScaleOut Software

- Develops and markets software for **in-memory computing**:
  - Scales application performance and
  - Provides real-time analytical insights using
  - In-memory data storage, computing, stream processing, and simulation.
- Deep domain expertise:
  - Dr. William Bain, Founder & CEO. Bell Labs, Intel, Microsoft
  - Over 19 years in the market
  - Consistent track record of innovation and technology leadership
- Flexible business model to meet diverse needs:
  - Fully supported software releases
  - Dedicated to ease-of-use to minimize training and lower TCO
  - On-premise or cloud
  - Choice of licensing models: perpetual, subscription, cloud-hosted

Microsoft Partner
Silver Application Development

amazon web services Partner Network
TECHNOLOGY PARTNER

Windows Azure

cloudera CERTIFIED

Gartner 2015
Cool Vendor

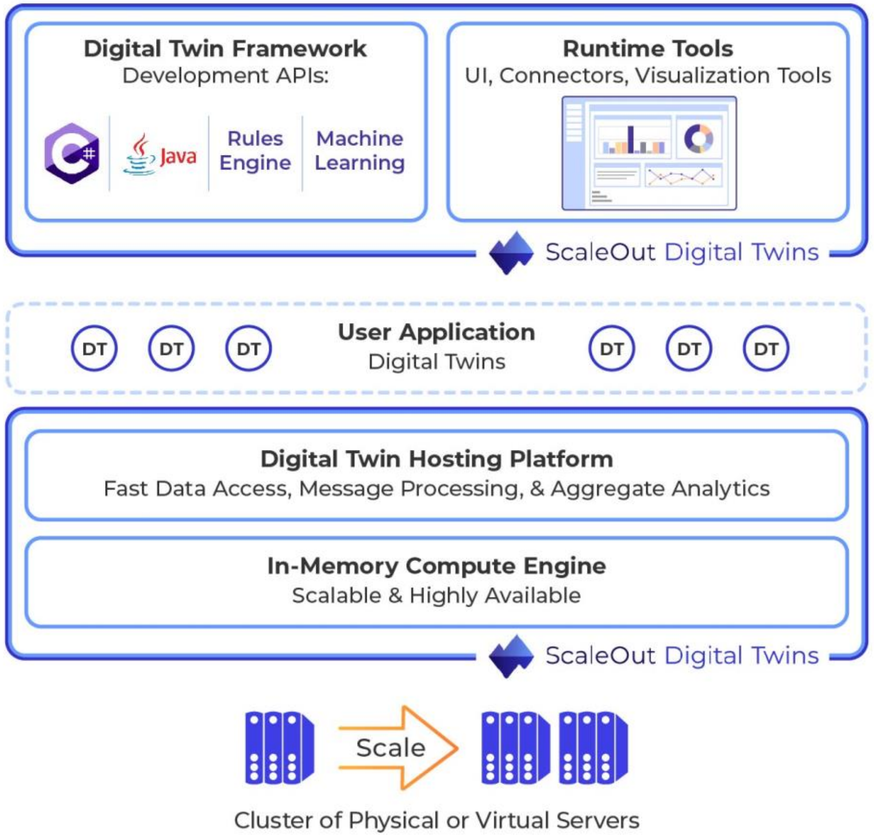
ScaleOut Digital Twins

ScaleOut Software

Real-Time Monitoring and Predictive Modeling with Digital Twins

A software framework for building digital twin applications and a fast, scalable hosting platform:

- Runs in the Azure cloud or on-premises
- Includes development tools in C#, Java, using business rules and machine learning.
- Hosts digital twins on ScaleOut's patented, in-memory computing platform.
- Has a UI for control, query, and visualization of aggregate analytics.
- Connects to messaging hubs, Kafka, REST, alerting services, database servers, and more.



Challenge: Power Grid Security & Disaster Response

How track a geographically distributed power grid with thousands of nodes for intrusion or disruption?

- Where are the threats?
- How significant are they?
- How are they moving?
- How should we react?

Challenge: Logistics & Telematics

How track the safe distribution and delivery of millions of time-critical items?

- Where is each item/vehicle right now?
- How are delays or issues (e.g. temperature) affecting its safety?
- Which vehicles are most in need of assistance?
- Is there an emerging widescale problem that needs a strategic response?

A New Vision for Digital Twins

ScaleOut Software

> *A digital twin is a virtual representation of real-world entities and processes, synchronized at a specified frequency and fidelity. … Digital twins use real-time and historical data to represent the past and present and simulate predicted futures. …   -- as defined by the Digital Twin Consortium*

- Digital twins were conceived to help design and test complex new devices (PLM).
- More recently, operational digital twins are used in small numbers to track telemetry in production for preventative maintenance.
- The next step: use digital twins for **real-time monitoring** and **predictive modeling** of large systems:
  - Vehicle fleets
  - Logistics systems
  - Large infrastructures
  - Ecommerce shoppers

Designing a Jet Engine

Monitoring an Industrial Robot

Tracking the US Railway System

The Need for Granular Real-Time Analytics

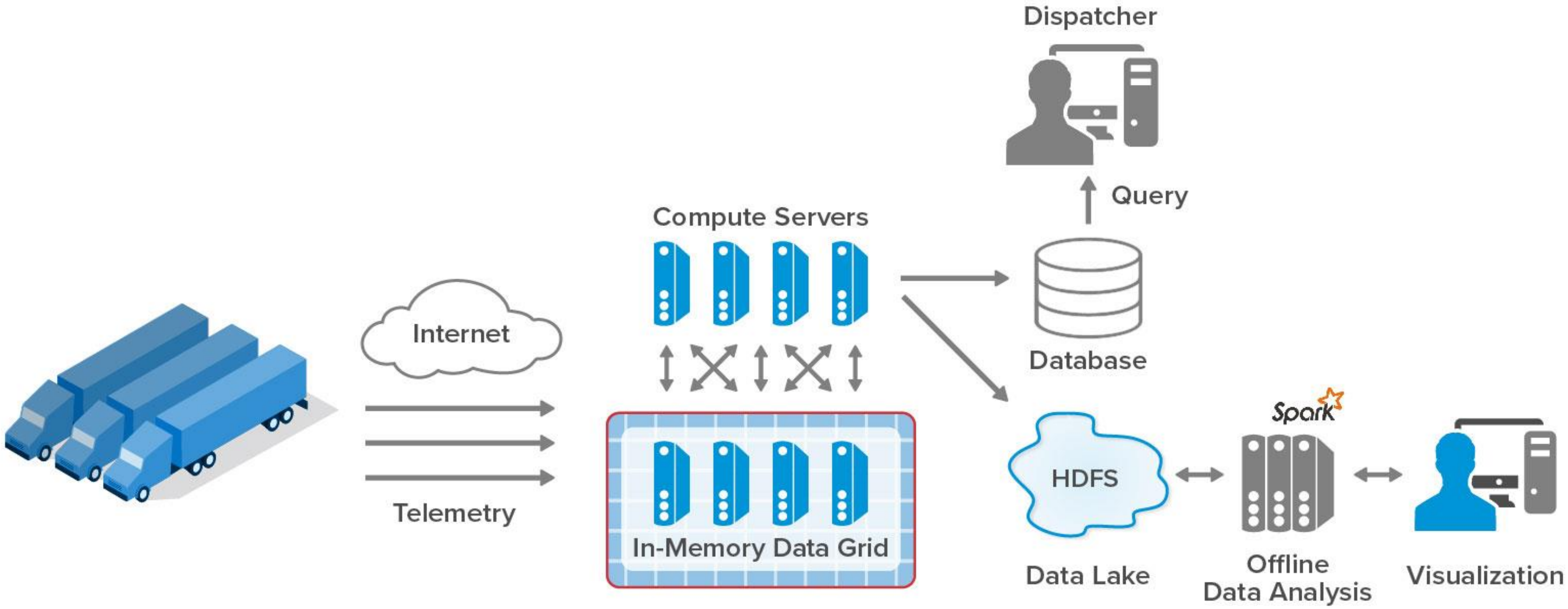
ScaleOut Software

Today's streaming analytics do not:

- Track data sources *individually*.
- Perform aggregate analytics online.

As a result, they cannot:

- Predict emerging issues for each data source.
- See important trends in real time (seconds).



Typical Telematics Architecture for Streaming Analytics

The Solution: Real-Time Digital Twins

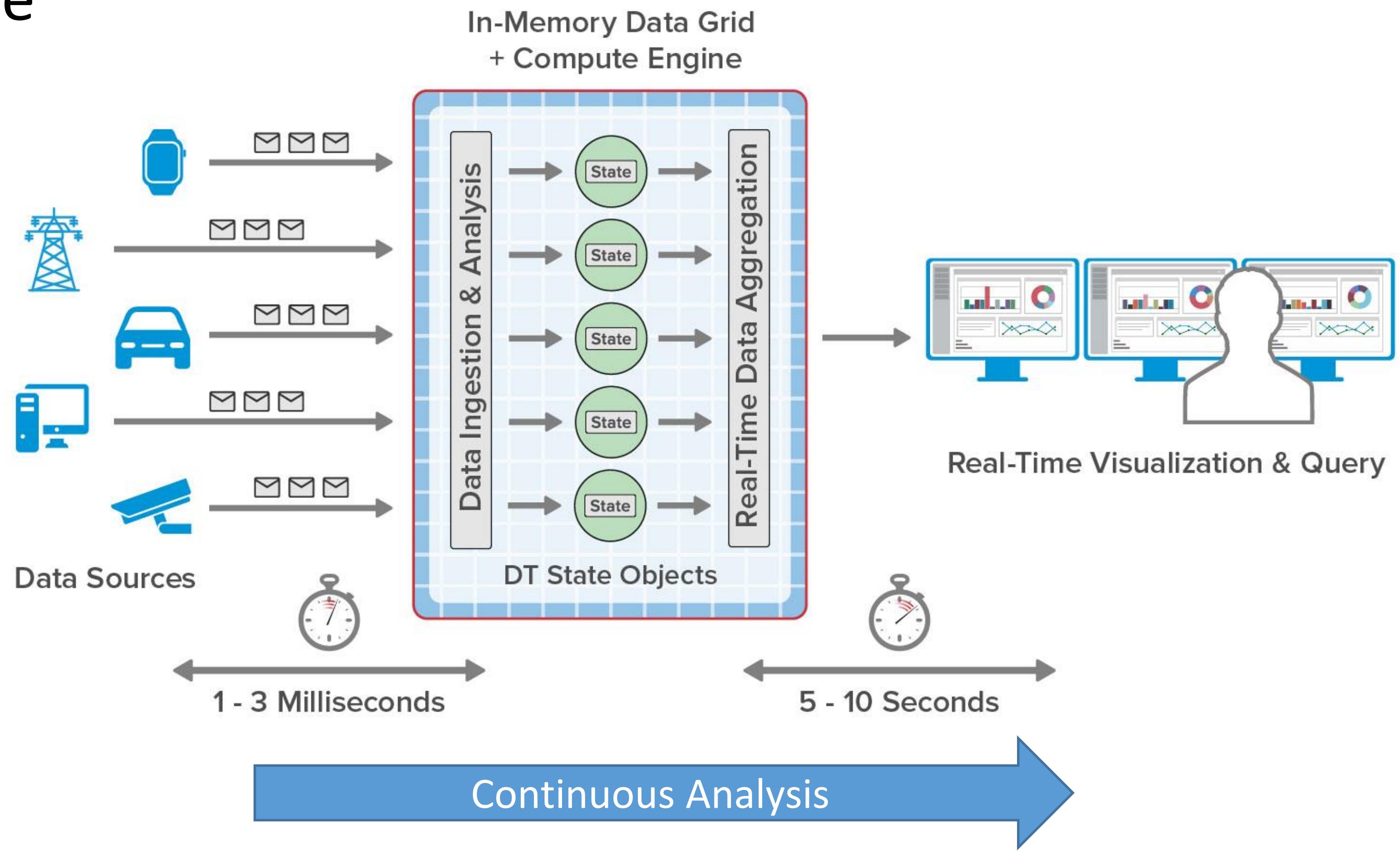
ScaleOut Software

A new in-memory technique for streaming analytics:

- Separately analyzes telemetry for *each* data source.
- Maintains dynamic state for each data source.

Key Advantages:

- Introspects in real time on *each* data source.
- Provides data for immediate aggregate analytics.
- Transparently scales.

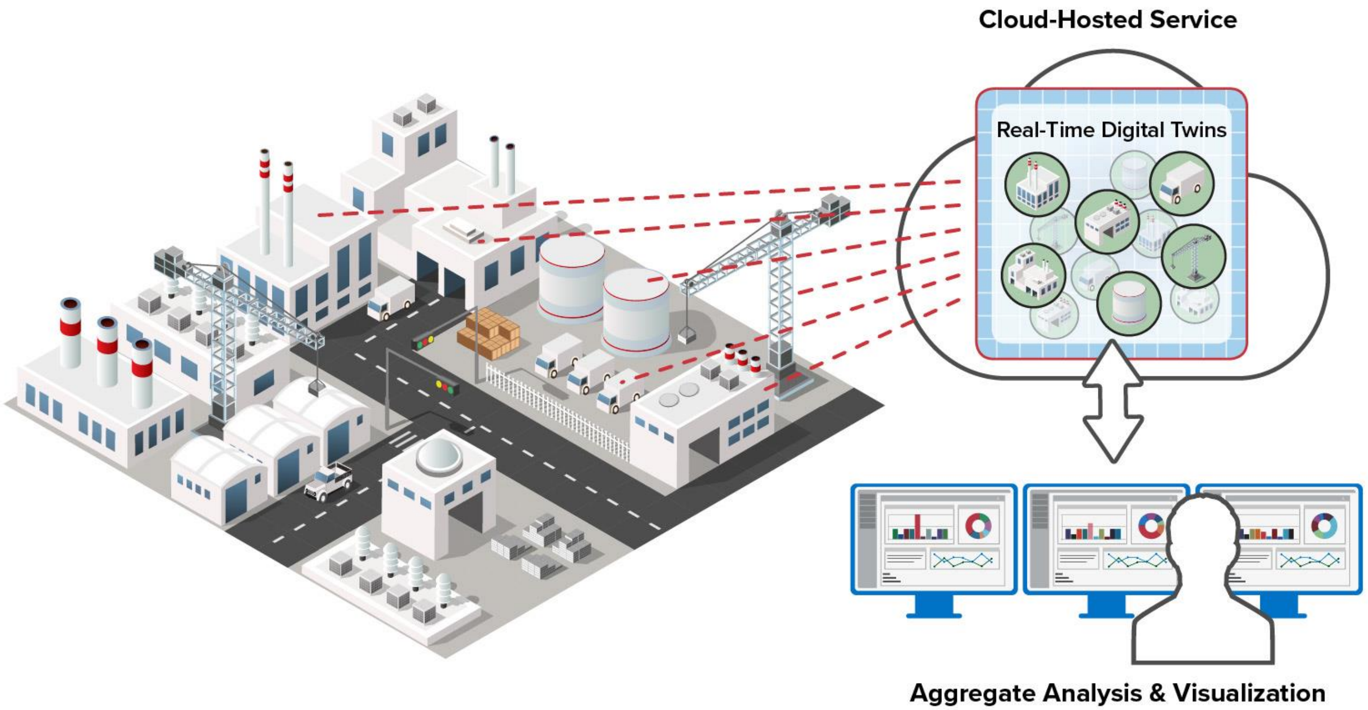


Software Architecture Using Real-Time Digital Twins

Benefits of Real-Time Digital Twins

ScaleOut Software

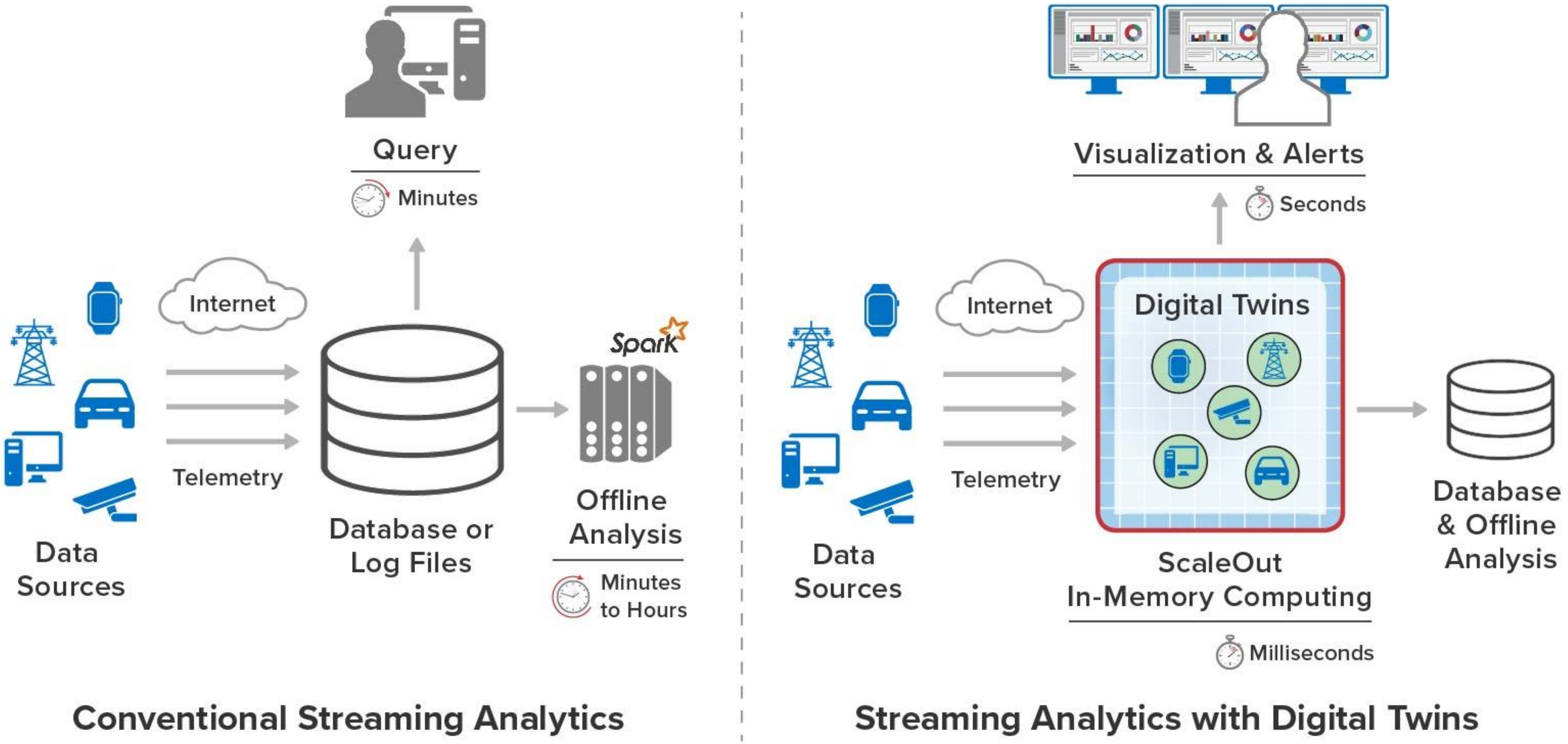
- **Faster responses**: Continuously analyzes incoming telemetry.
- **Deeper introspection**: Tracks and updates information about each data source.
- **Better situational awareness**: Continuously aggregates & visualizes derived state.
- **Transparently scalable**: Seamlessly scales using in-memory computing.
- **Easy to use**: Uses simple, object-oriented APIs.



Comparison to Conventional Techniques

ScaleOut Software

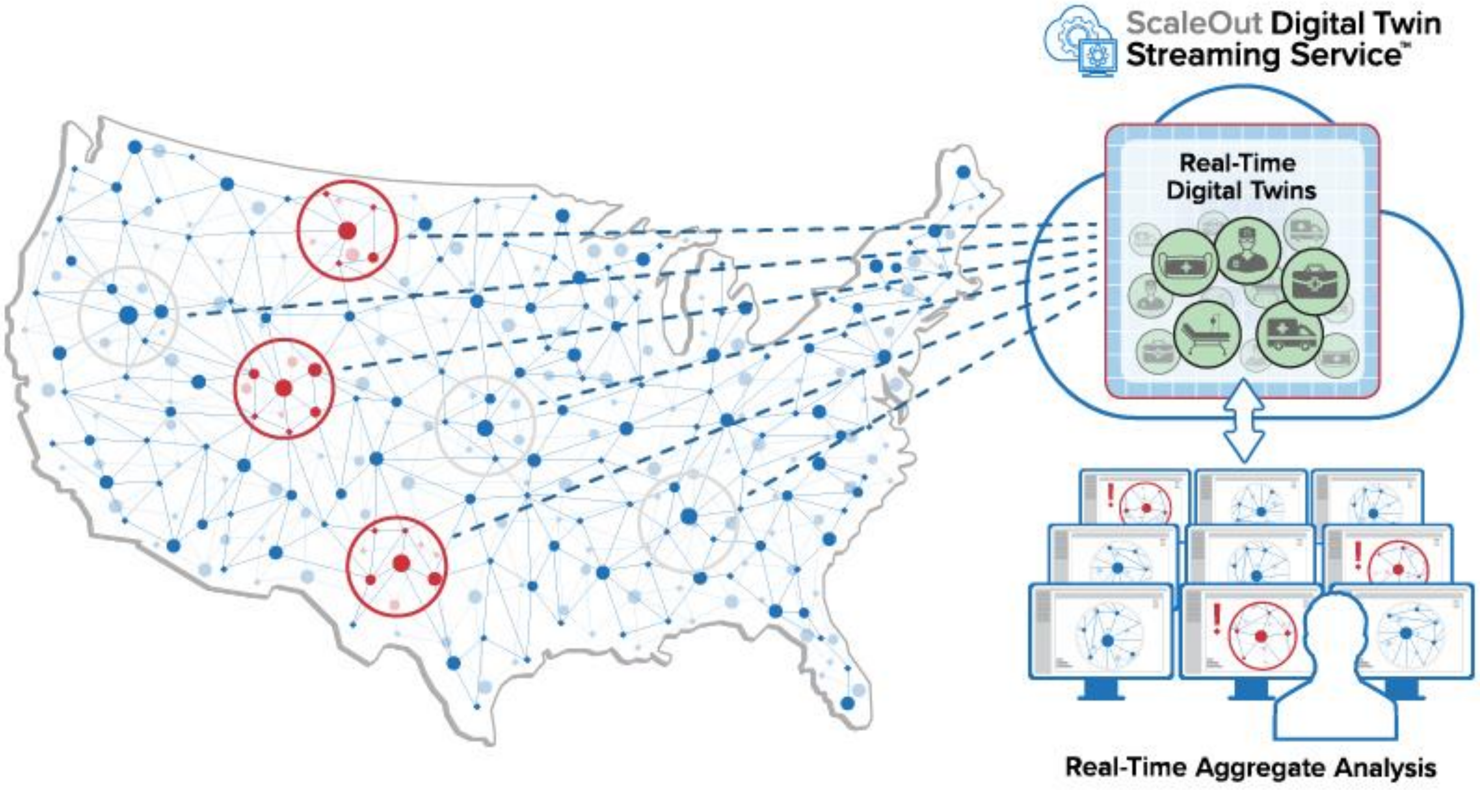
Real-time digital twins immediately extract useful information, generate alerts, and visualize emerging trends:



Digital Twins Enable Aggregate Analysis

ScaleOut Software

- **Simultaneously** analyzes all real-time digital twin instances to identify and visualize widespread trends in real time.
- Boosts **situational awareness**.
- Some examples:
  - Offer flash sale to individual online shoppers *and* tweak it in real time by tracking overall impact.
  - Monitor power stations *and* identify rolling blackout.
  - Track a rental car fleet *and* alert all drivers vehicles approaching an unplanned road closure.
  - Identify an emerging pattern of fraudulent wire transfers.

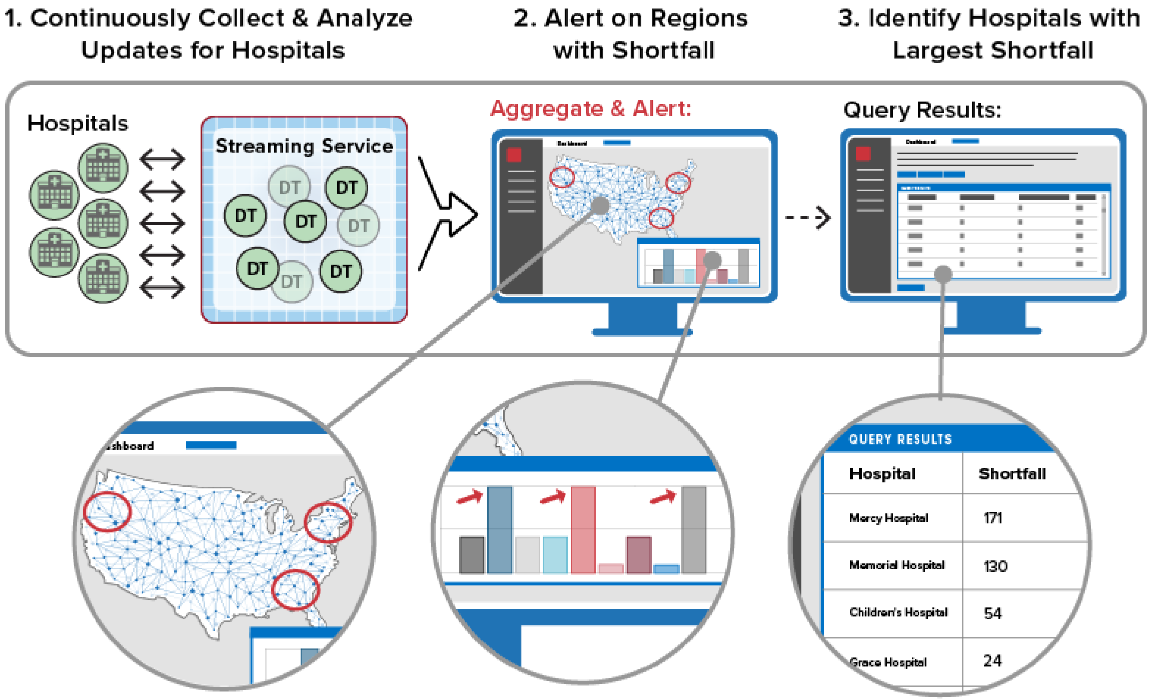


Example of Aggregate Analytics & Query

ScaleOut Software

Maximizes situational awareness for a logistics application:

- Integrated analytics engine combines key digital twin data in seconds.
    - Example: Determine largest shortfall in hospital supplies by region.
- Streaming service lets users visualize results.
    - Example: Show shortfall by region as a bar chart to alert on problem areas as they occur.
- Users query digital twin data to identify issues and take action.
    - Example: Query digital twins to find specific hospitals with largest shortfall in affected regions.



Using Digital Twins for Simulation

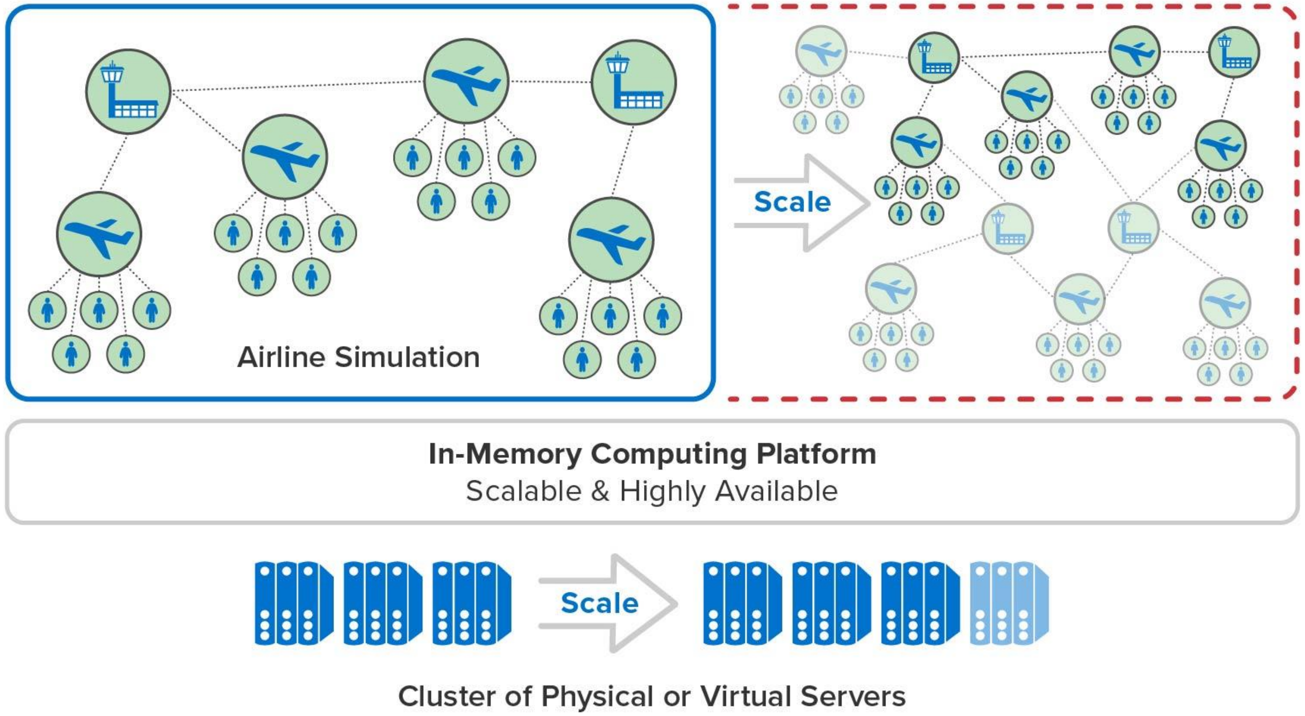
ScaleOut Software

Digital twins:

- Enable the construction of large-scale simulations (1000s to millions of interacting entities).
- Simplify simulation development.

Key benefits:

- **Improved decision making** for managing complex systems.
- Ability to **make predictions** for *live* systems.
- **Testing and validation** for streaming analytics.

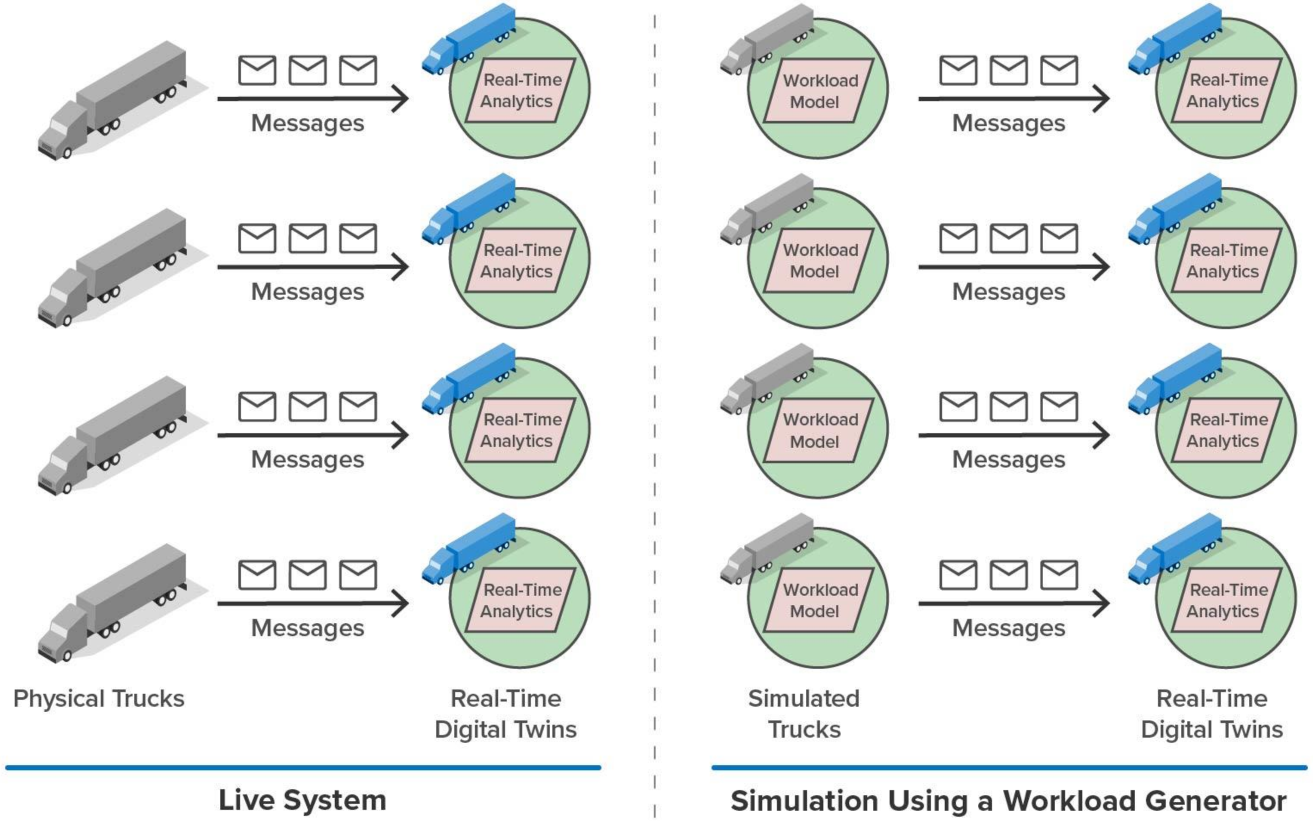


Airline Simulation

Using Simulation to Validate Streaming Analytics

ScaleOut Software

- Digital twins can model live data sources:
  - Generate telemetry in simulation and send it to real-time digital twins.
  - Real-time digital twins can implement streaming analytics.
- Simulation validates streaming analytics in a variety of scenarios.
- Real-time digital twins can then be deployed to analyze live telemetry.



Telematics Simulation

Many Use Cases

ScaleOut Software

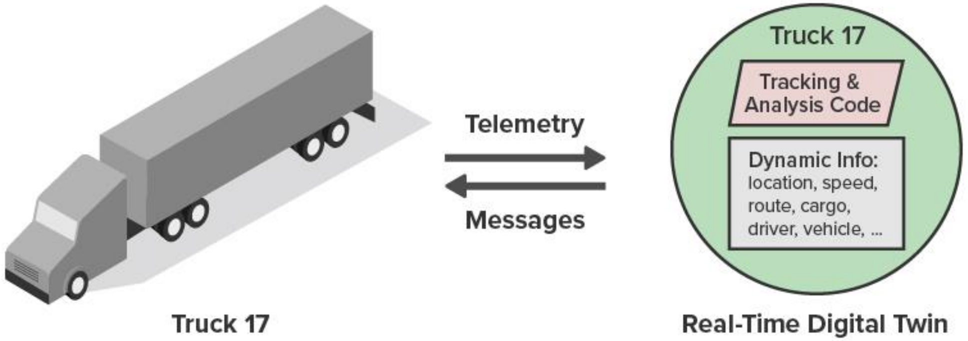
Here are some use cases for digital twins tracking and modeling complex live systems:

- Telematics – vehicle fleets
- Transportation safety – railroads
- Transportation systems – help manage passengers and assets for an airline
- Logistics –tracking shipments
- Disaster recovery –tracking wildfires and alerting personnel
- Security – tracking physical and cyber security perimeters
- Crime prevention – cloned license plate and convoy detection
- Fraud detection – credit card fraud
- Health informatics – tracking with wearable devices
- Smart cities and building management

Example: Fleet Telematics

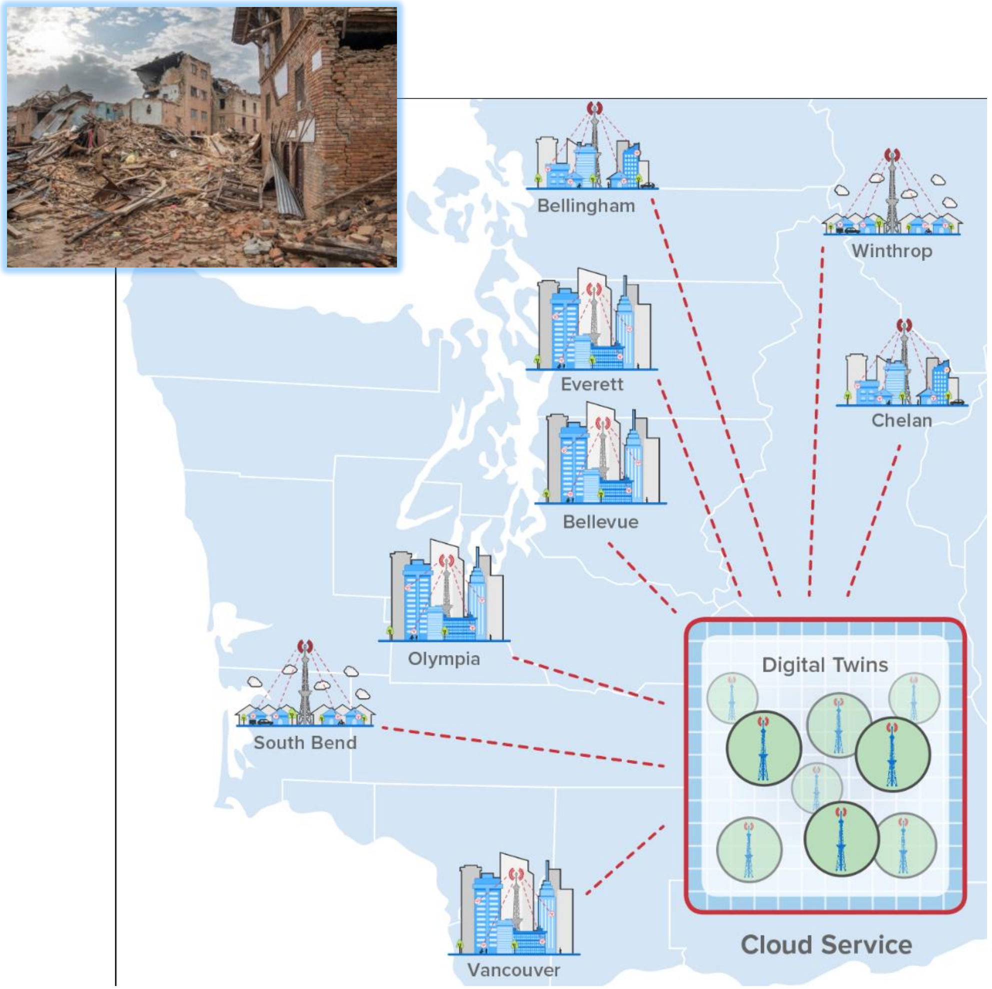
ScaleOut Software

- Real-time tracking for a car/truck fleet (typically, thousands of vehicles)
- Telemetry includes location, speed, mechanical & cargo parameters.
- Digital twins add route, cargo, info on driver, service history & issues, weather, etc.
- Using incoming telemetry, digital twins can:
  - Alert driver to upcoming hazardous road conditions or weather delays.
  - Assist lost driver or alert if driving too long or unsafely.
  - Track emerging mechanical issues with vehicle or risk to cargo.
  - Maintain status which can be aggregated for all trucks to enhance dispatcher's situational awareness of the fleet.



Example: Disaster Recovery

- Goal: help find buried survivors after an earthquake using their cell phone data.
- How?
  - 5G cell towers can track direction and signal strength for each subscriber.
  - This information can help locate survivors.
- There are about 350K 5G cell sites in the U.S.
- Digital twins can maintain current status of all cell towers.
  - Can track fast-changing updates to call status for each cell tower.
  - Aggregate analytics can immediately pinpoint areas of greatest need.



Example: Tracking the Freight Rail System

ScaleOut Software

- Each year in the US, thousands of freight trains carry 1.6 billion tons of freight across 140,000 miles of track:
    - Approx. 300 trains per week
    - Approx 500K carloads per week
- In 2022, there were more than 1,100 train derailments, causing over 100 million dollars in damage.
- 6,000 hot boxes around the US monitor the temperature of wheel bearings, which can cause derailments if they get too hot.
- Hot boxes just alert operators by radio when high temperature is detected; they do not track trends.
- Digital twins can solve this problem:
    - Track and analyze temperature trends for all wheel bearings.
    - Integrate service history and other relevant data to assess danger and create timely alerts.

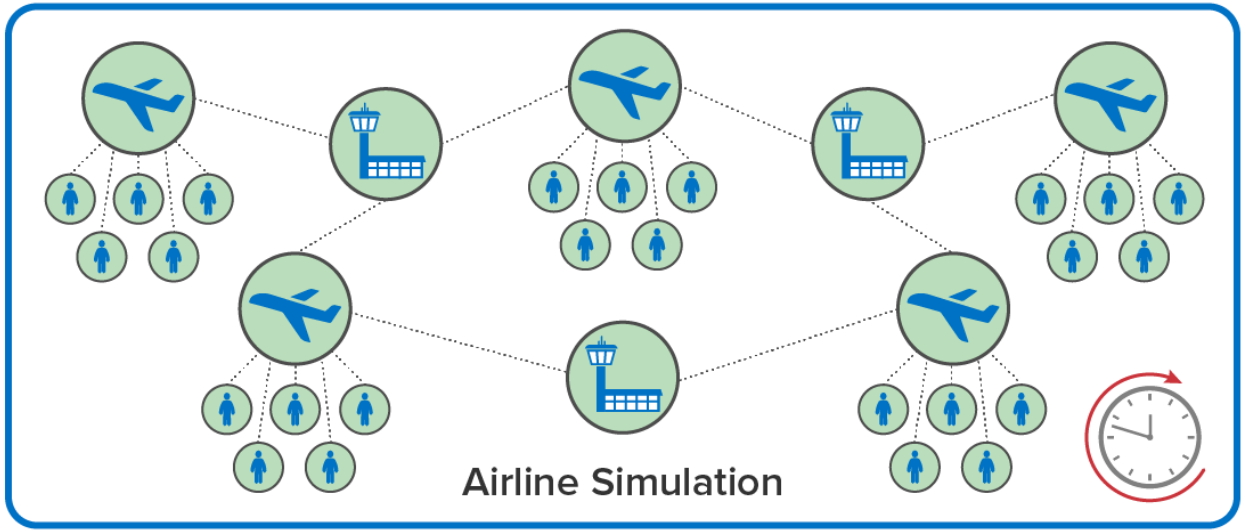
July 28, 2023

Example: Managing Airline Operations

ScaleOut Software

Build **system simulations** with interacting digital twins exchanging messages for performance evaluation & prediction.

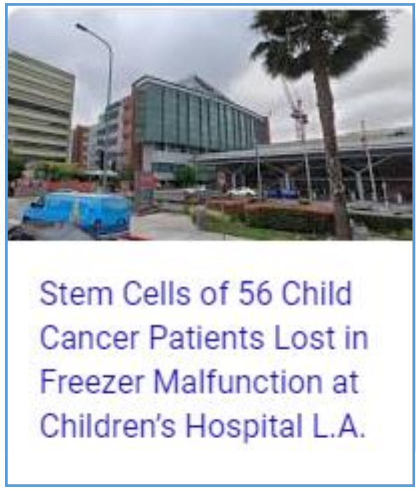
- Use digital twins to model physical entities:
  - Airplanes, passengers
  - Airports, gates, etc.
- Model and measure complex interactions.
- Evaluate management decisions faster than real time.
- Enable improved flying experience.



Example: Medical Refrigerators

- Cloud-based streaming service monitors 10,000+ medical refrigerators:
- Refrigerators hold highly important tissue samples, embryos, etc.
- Service receives periodic telemetry:
  - Temperature
  - Power consumption
  - Door position, etc.
- Must predict failure before it occurs:
  - Notify user to migrate contents to another refrigerator.
  - Avoid false positives.
  - Identify widespread power outages (aggregate analytics).

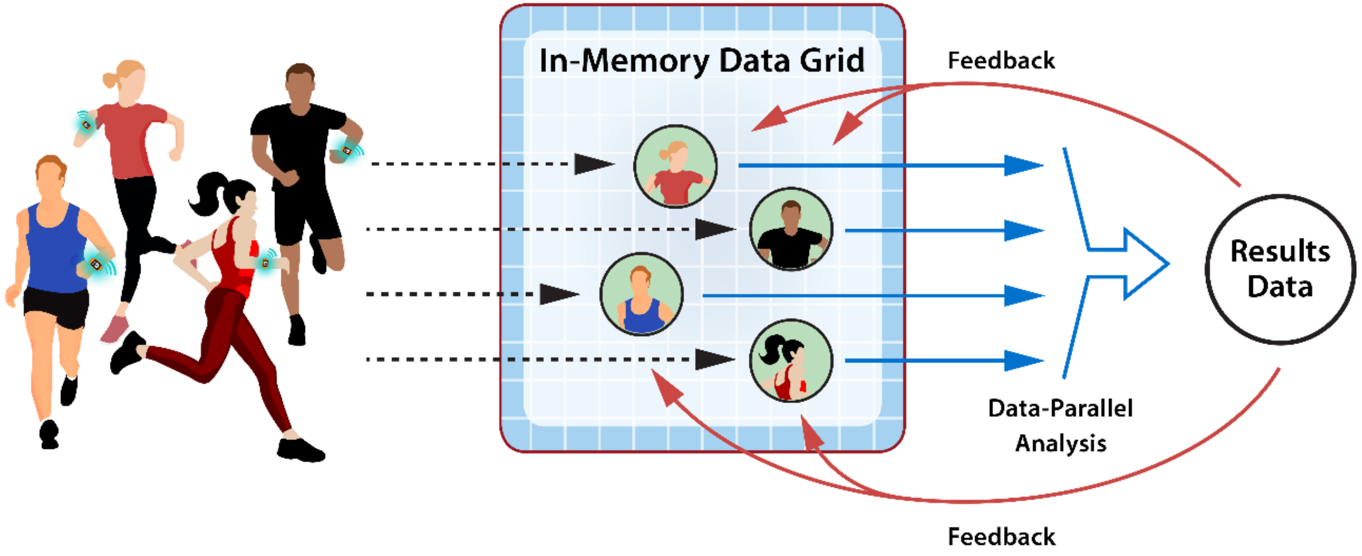
Sep 26, 2019



Example: Personalized Health Tracking

Real-time digital twins analyze telemetry from health-tracking devices to help ensure wearer's safety and give feedback:

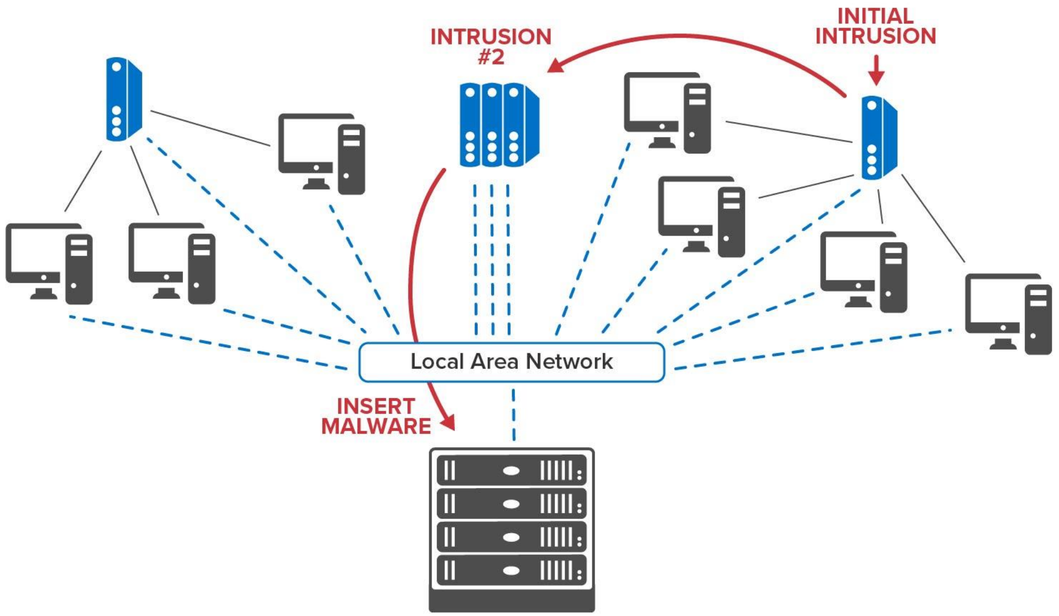
- RTDTs receive periodic messages with key metrics (heart rate, blood oxygen, etc.).
- RTDTs track person's health history, medications, limitations, recent medical events.
- Data-parallel analysis combines dynamic results from large populations for feedback.



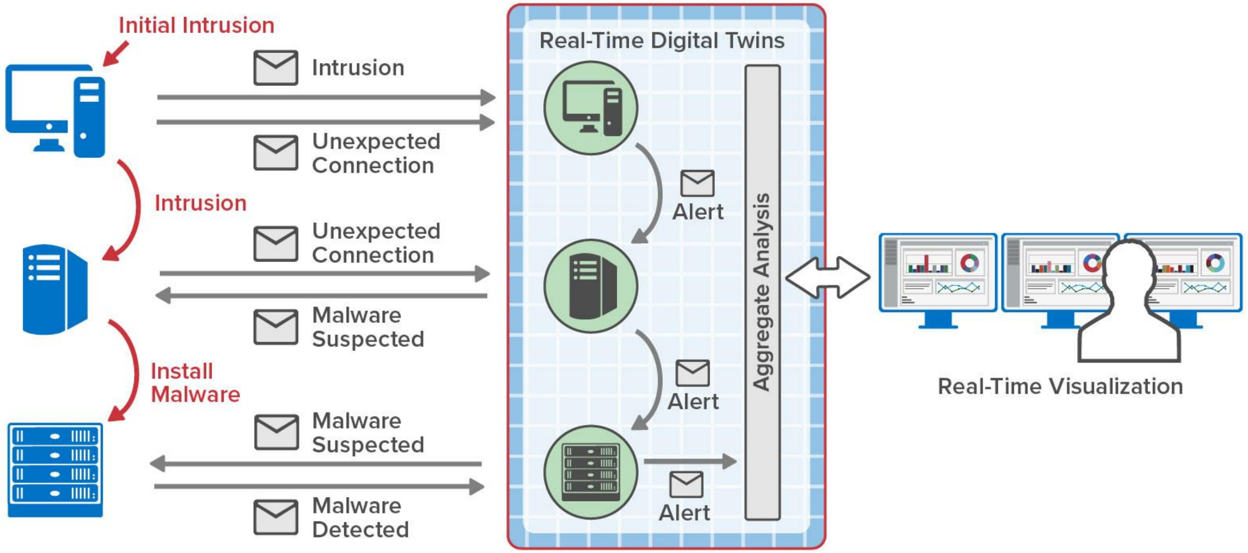
Example: Cybersecurity

ScaleOut Software

- Real-time digital twins can be used to:
  - Track the evolving state of entry points and assets (e.g., desktops, routers, servers).
  - Display the dynamic state of assets, not just events in SIEM consoles.
  - Track (with ML) and possibly stop the progress of an intruder within an infrastructure.



Example of a Cyber “Kill Chain”

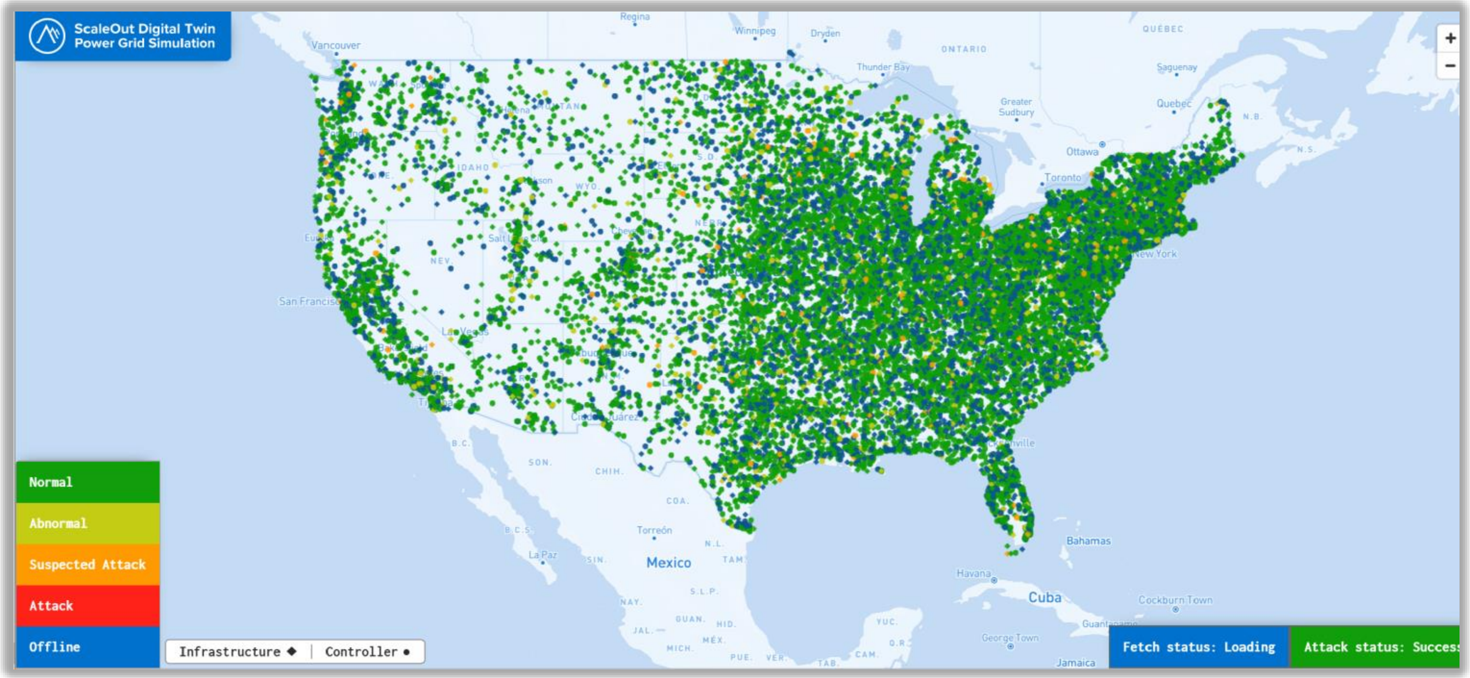


Use of RTDTs to Interrupt Intruder’s Progress

Demo: Simulated Power Grid

ScaleOut Software

- **Goal**: **Demonstrate value of enhanced situational awareness**
- **Overview**:
  - Simulates 20K nodes within a power grid spanning the US and subject to outages or attack.
  - Each node streams telemetry with its status to its real-time digital twin in the cloud service.
  - Real-time twin evaluates telemetry and updates derived state: *alert level*:
    - Assesses threat level for that node (range 0-20)
    - Updated by examining telemetry with knowledge of node's function and reporting history (e.g., false positives)

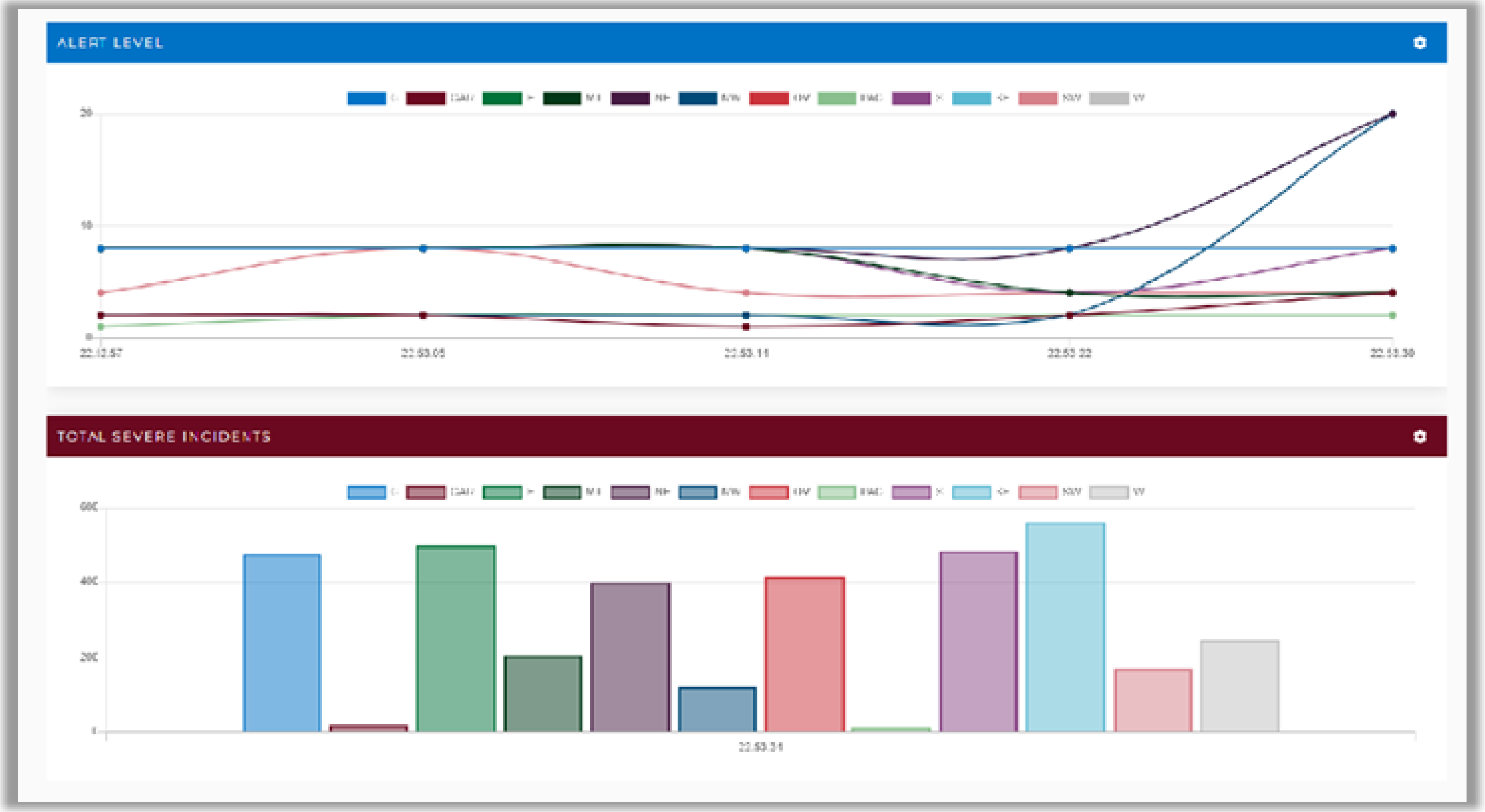


Demo: Aggregate Analytics

ScaleOut Software

Aggregate analysis by the cloud service quickly pinpoints emerging threats for managers:

- Evaluates alert level in digital twin state for all nodes and aggregates results.
- Dashboard widgets display maximum alert level by region.
- Runs every 5-10 seconds as a background MapReduce job.

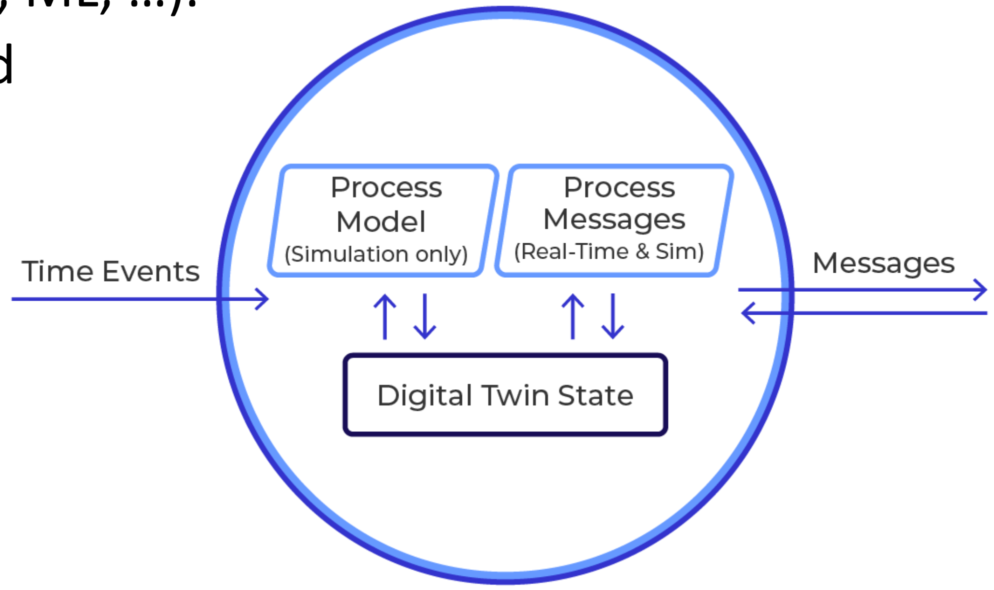


Anatomy of a Digital Twin

ScaleOut Software

A **digital twin model** It consists of a:provides integrated support for both streaming analytics and simulation. For each device type, it consists of:

- A **message processing** and optional **simulation processing method**:
  - Receives and analyzes live or simulated messages and simulation time events.
  - Encapsulates analysis algorithm (Java, C#, rules, ML, …).
  - Updates state information, generates outbound messages and alerts.
- A **state object**, which holds dynamic data:
  - Context/metadata for analyzing events
  - Also: time-ordered event lists, cached parameters
  - One **instance** per data source or device



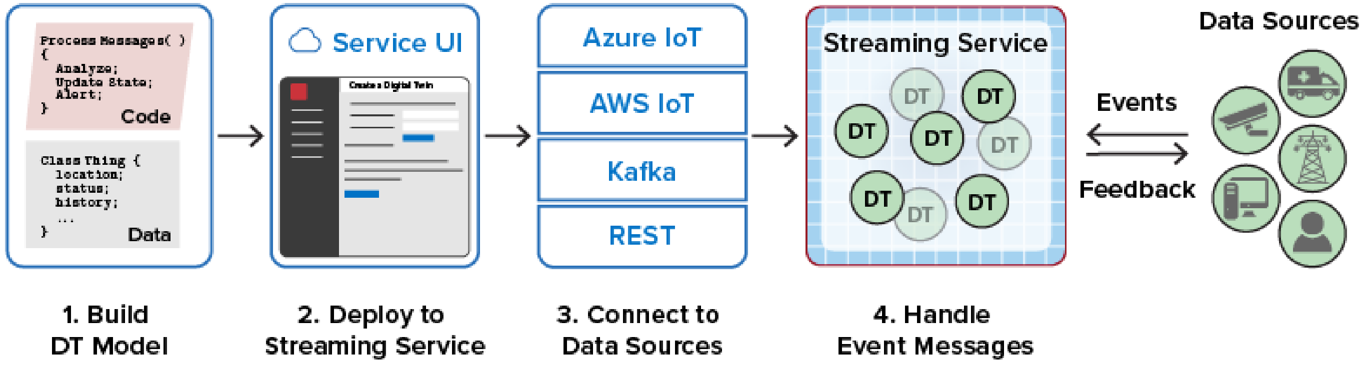
ScaleOut Digital Twin Model

Simplified Development + High Performance

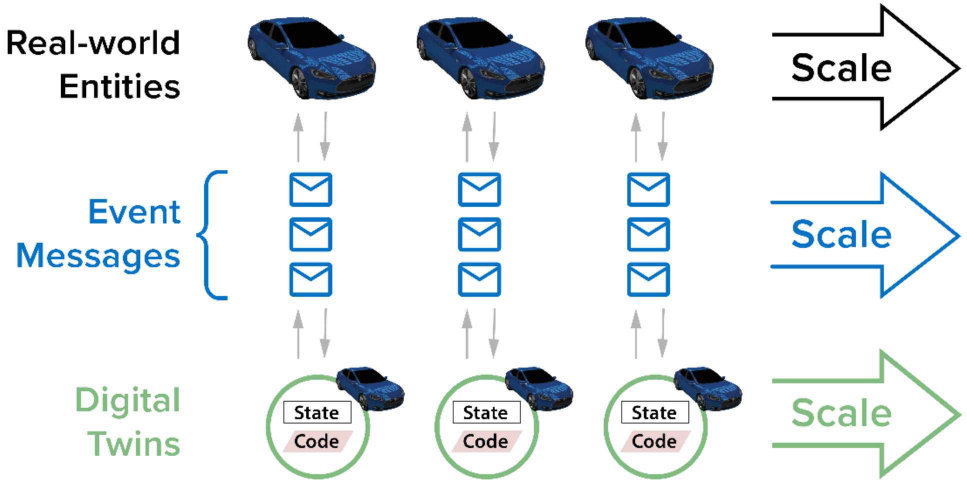
ScaleOut Software

Enables easy-to-build, "agile" applications.

- Digital twin model minimizes code development:



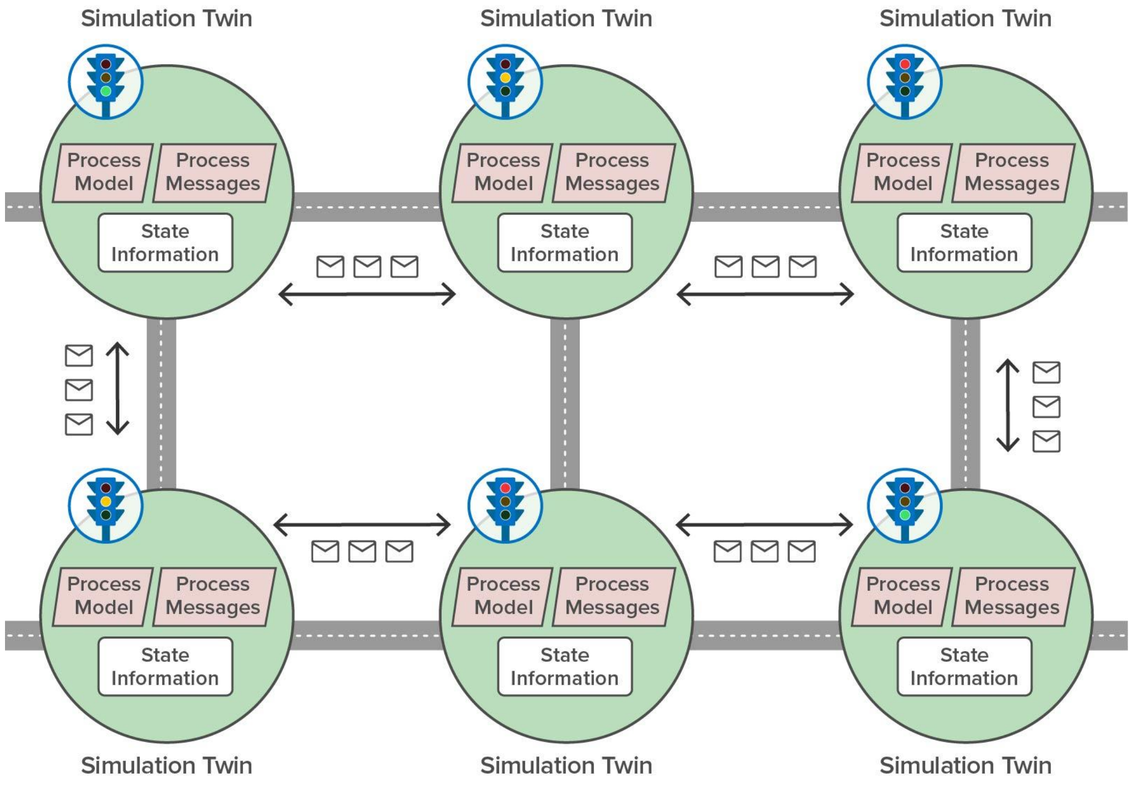
- Streaming service handles scaling and ensures high availability:



Implementing Digital Twin Simulations

ScaleOut Software

- User's digital twin model implements:
  - State information to be maintained by each instance
  - A ProcessModel method that runs at each time step
  - A ProcessMessages method that handles incoming messages
- ScaleOut's in-memory platform:
  - Distributes digital twin instances across a server cluster.
  - Runs all instances at each time step.
  - Delivers and processes messages sent between instances.
  - Provides data visualization and query.

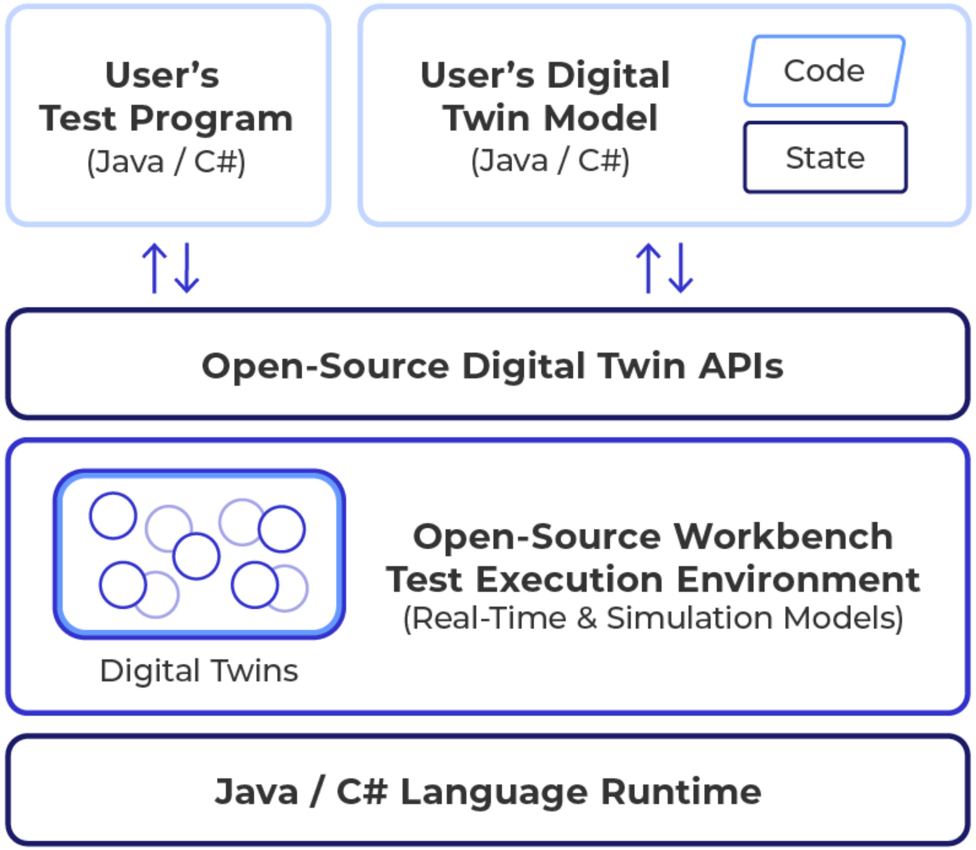


A Traffic-Flow Simulation

Development Workbench

ScaleOut Software

- Debugging and test within the cloud service can be challenging.
  - Cannot set breakpoints and examine state.
  - Difficult to verify all code paths.
- Workbench environment allows development and test on developer's workstation:
  - Uses same base classes as production deployment.
  - Replaces `ProcessingContext` with mock version.
  - Deploys model with same APIs as used on-premises.
  - Allows sending/receiving messages from instances.
  - Allows instance state to be examined.

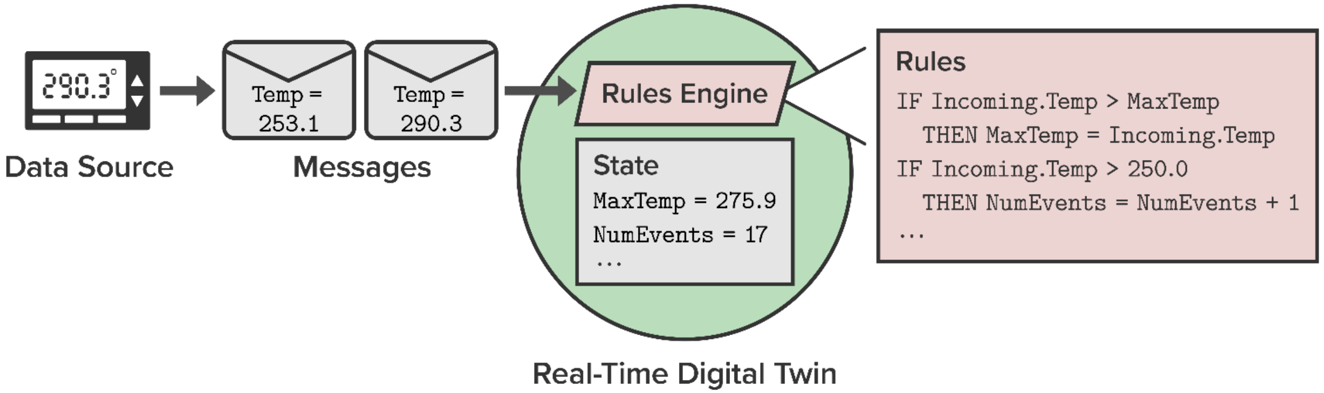


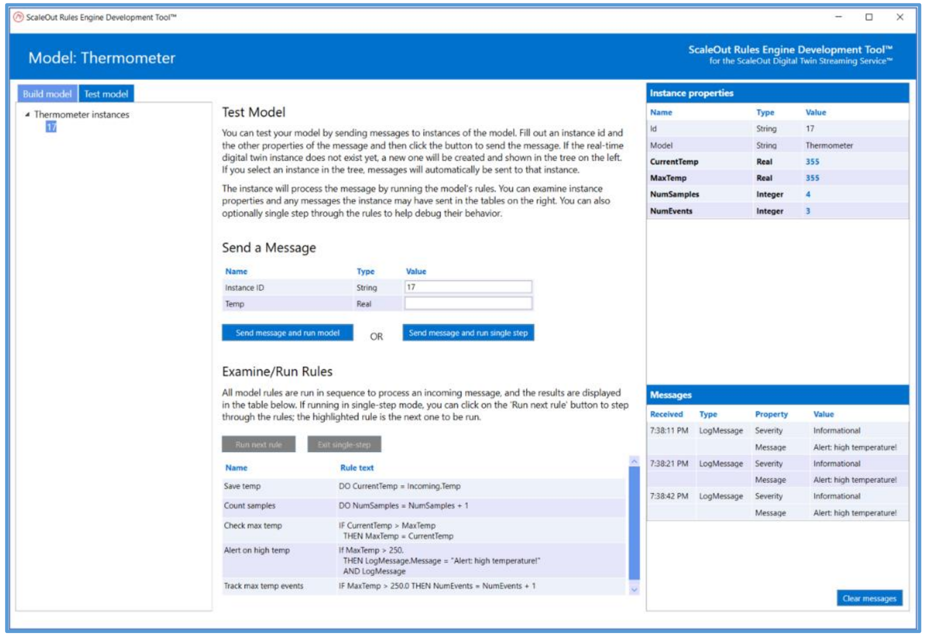
Optional Rules Engine Simplifies Design

ScaleOut Software

- Allows application code to be described as a sequence of rules:
  - Uses a simple "IF condition THEN action" rules language.
  - Avoids the need to use formal programming languages.

- Desktop-based **ScaleOut Model Development Tool** enables fast development and testing.



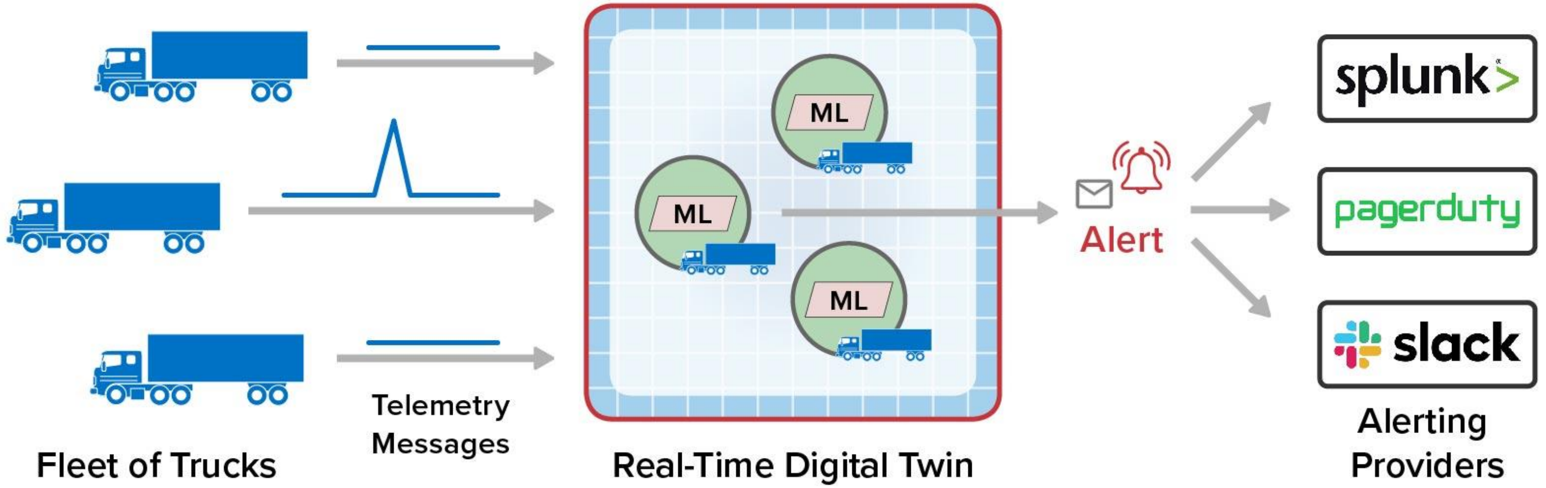


Optional Machine Learning Adds Power

ScaleOut Software

Digital twins can run machine learning algorithms from Microsoft ML.NET library to analyze telemetry from data sources:

- Performs spike, trend, and anomaly detection and send alerts with **no code required**.
- Runs separately for each data source in each real-time digital twin.
- Can be integrated with rules for additional analysis.

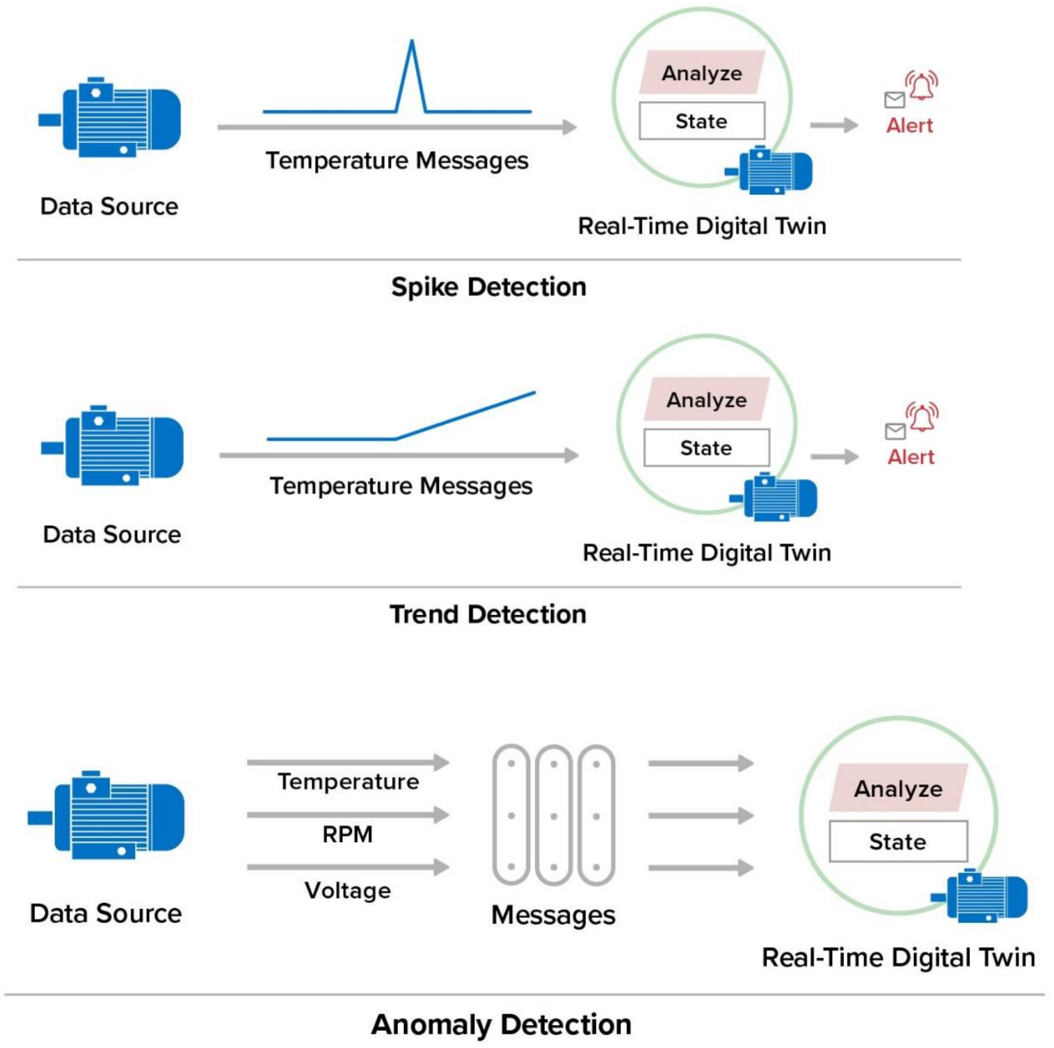


Machine Learning Capabilities

ScaleOut Software

Automatic spike, trend, and multi-parameter anomaly detection:

- Spike and anomaly detection use ML.NET.
- Trend detection uses ScaleOut linear regression algorithm.
- User configures with ScaleOut Model Development Tool.

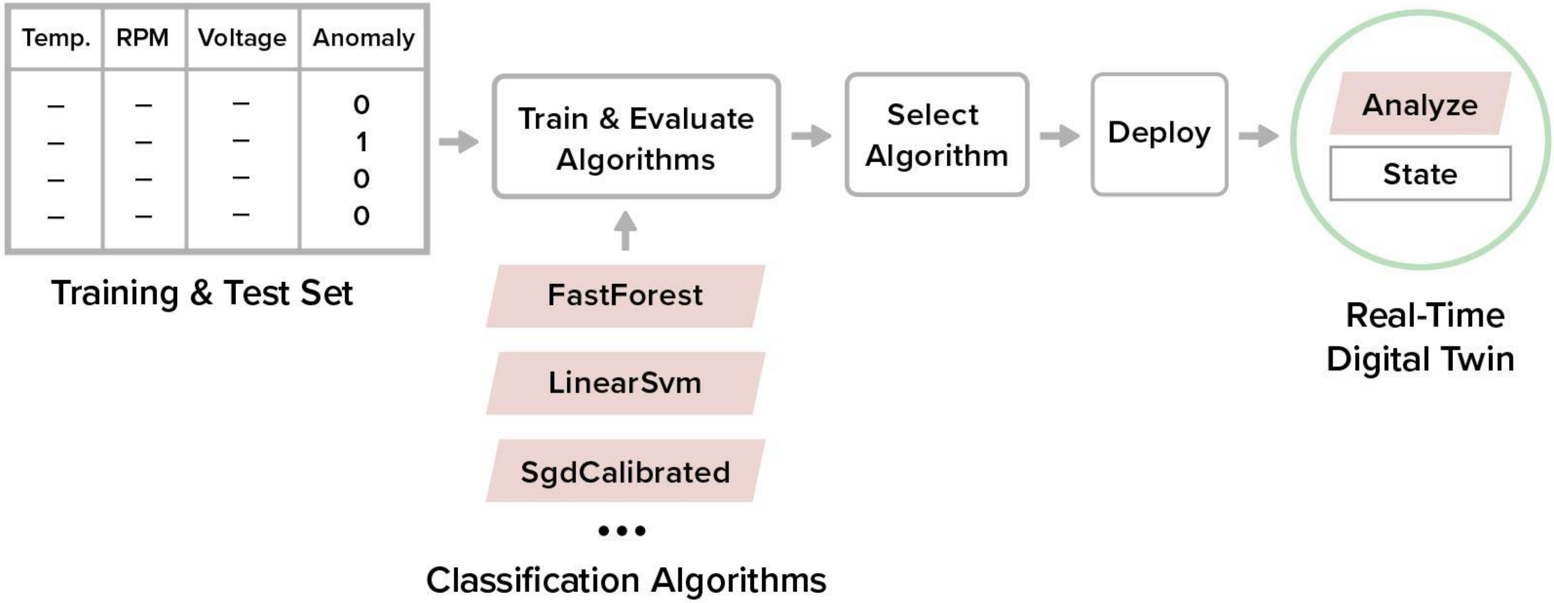


Training Anomaly Detection

ScaleOut Software

Anomaly detection uses binary classification algorithm trained with user's data set:

- User can train multiple algorithms from ML.NET.
- User selects one algorithm based on metrics from training.

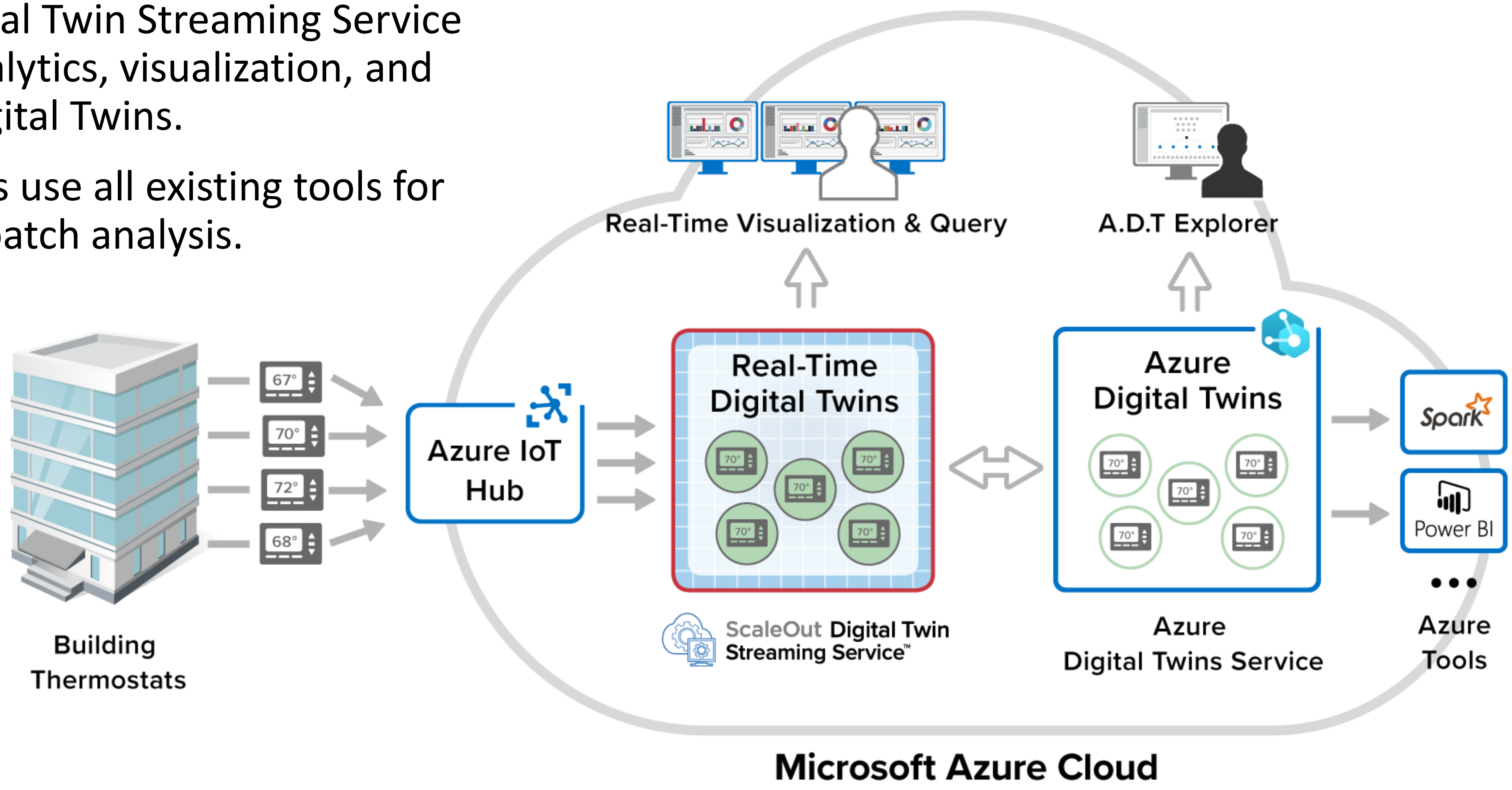


Workflow for Training Real-Time Digital Twin Model

Optional Azure Digital Twins Integration

ScaleOut Software

- The ScaleOut Digital Twin Streaming Service adds real-time analytics, visualization, and query to Azure Digital Twins.
- Azure Digital Twins use all existing tools for visualization and batch analysis.

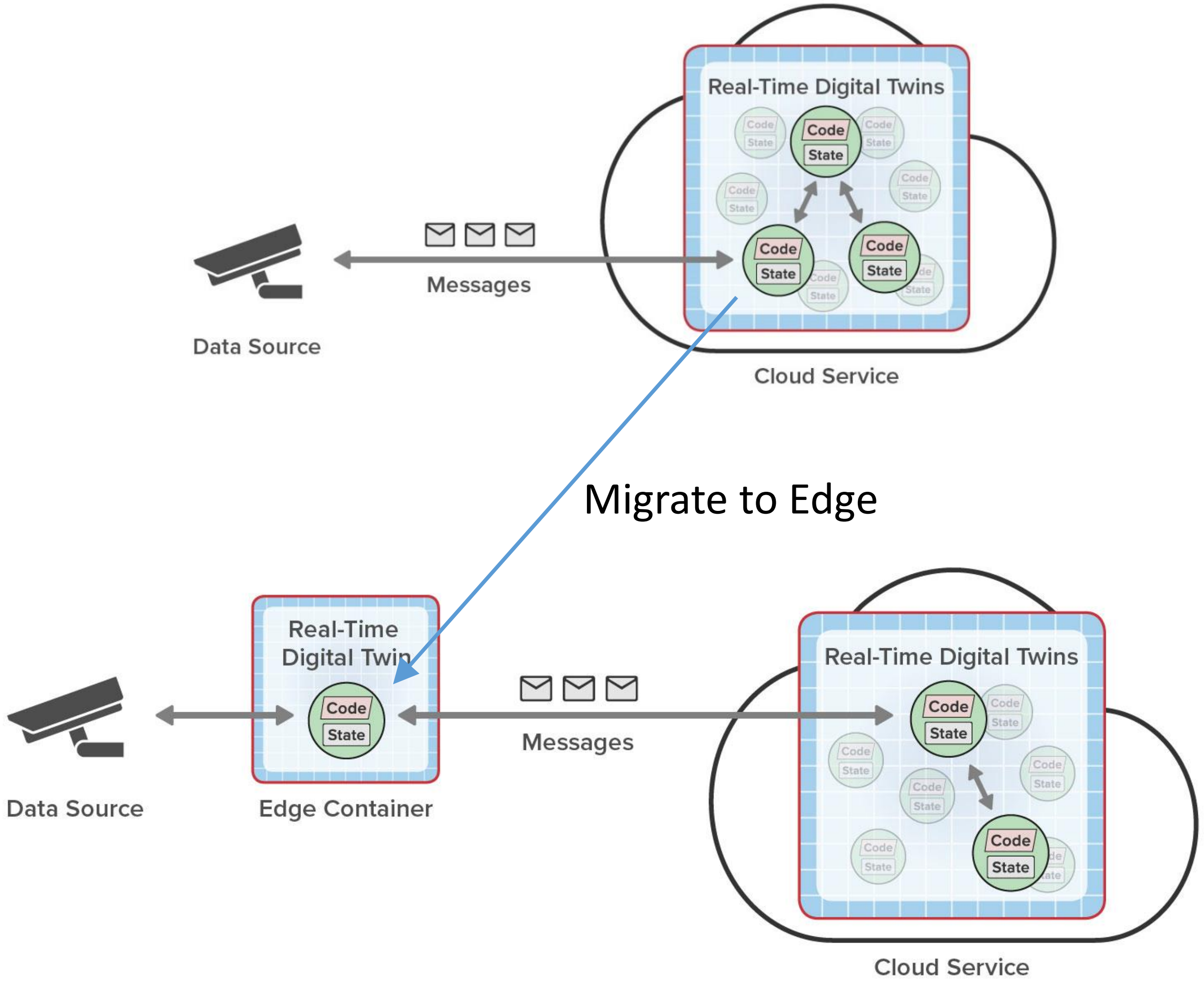


Thermostat Example Running in the Azure Digital Twins Ecosystem

Transparent Migration to the Edge

ScaleOut Software

- Real-time digital twins can naturally be organized into hierarchies.
- As edge computing becomes more capable, low-level real-time digital twins can migrate from a cloud service to edge containers:
  - Avoids programming changes and enables a uniform development model.
  - Allows division of labor between edge (device specific) and cloud (strategic).
  - Reduces message traffic.



Key Takeaways

- Digital twins aren't just for PLM.
- They offer a powerful software architecture for real-time monitoring and predictive modeling of large systems.
- Numerous applications in diverse verticals can benefit:
    - Transportation
    - Logistics
    - Disaster Recovery
    - Many more
- In-memory computing provides a key enabling technology:
    - Fast responses
    - Transparent scaling
    - Aggregate analytics
    - Real-time visualization

ScaleOut
Digital Twins

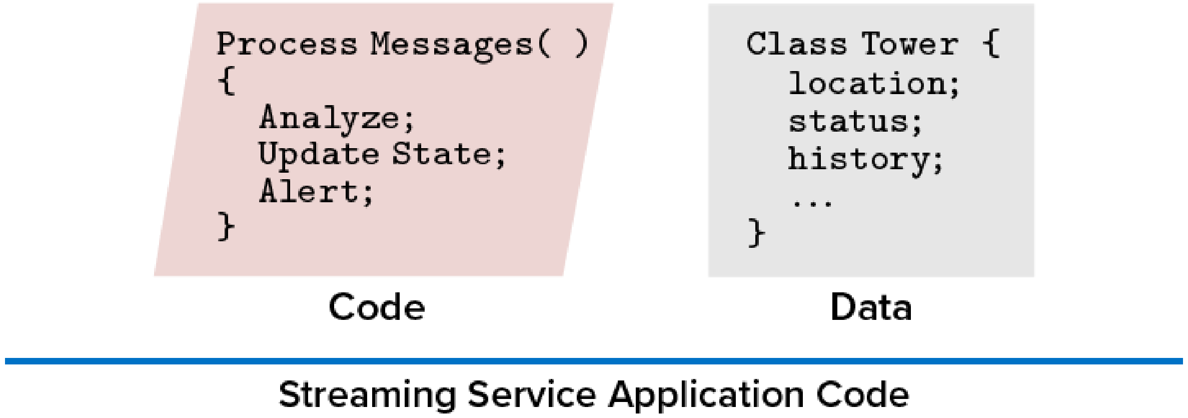
- Solves important unmet challenges for real-time monitoring and predictive modeling.
- Addresses key application scenarios in security, logistics, healthcare, disaster recovery, and more.
- Individually tracks 1000s of data sources in real time.
- Simulates 100s to millions of entities.
- Integrates real-time, aggregate analytics and query.
- Pinpoints operational issues fast with an agile solution.
- Dramatically simplifies application development.
- Automatically scales and delivers high availability.
- **Enables users to maximize situational awareness and decision-making with minimum effort.**

www.scaleoutsoftware.com

Example: Simulated Power Grid

Application code is dramatically simplified:

- Developer just implements one message processing method and state object for a single real-time digital twin.
  - Approximately 20 Java statements in this demonstration
- Platform automatically runs this code for all data sources.
  - Automatically correlates incoming messages by data source.
  - Automatically scales to handle large workloads.
- Developer avoids:
  - Selecting telemetry out of a single, combined event stream
  - Creating ad hoc storage for contextual data
  - Pushing the data into a data lake for offline analysis using Spark



Power Grid Demo Code

ScaleOut Software

- Definition of state object (Java):

```java
public class StatusTracker extends DigitalTwinBase {
    // State variables
    public String node_type;
    public String node_condition;
    public String region;
    public double longitude;
    public double latitude;
    // Derived state variables
    public int alert_level;
    public int minorIncidentCount;
    public int moderateIncidentCount;
    public int falseIncidentCount;
    public int severeIncidentCount;
    public int totalIncidents;
    public int totalResolvedIncidents;
    public boolean experiencingIncident;
    // Dynamic incident report list
    public List<IncidentReport> incidentList;}
```

Power Grid Demo Code

ScaleOut Software

- Sample message-processing code (Java):

```java
public ProcessingResult processMessages(ProcessingContext processingContext, StatusTracker digitalTwin,
                                        Iterable<StatusTrackerMessage> messages) throws Exception {
   // Iterate through the incoming messages:
   for(StatusTrackerMessage msg : messages) {
       // if the message indicates a moderate incident and this tracker has never had a severe
       // incident while the heuristic false incident ratio is greater than 50%, boost the alert level:
       if(msg.moderateIncident() &&
           digitalTwin.getSevereIncidentCount() == 0 &&
           digitalTwin.getModerateIncidentCount() > 0 &&
           ((double)(digitalTwin.getFalseIncidentCount()/
                     digitalTwin.getModerateIncidentCount()) >= 0.5)) {
               digitalTwin.setAlertLevel(Constants.MODERATE+3);
               digitalTwin.incrementModerateEventCount();
               digitalTwin.setStatusTrackerCondition(msg.getNodeCondition());
           }
       // ... [additional rules]
   }
   return ProcessingResult.UpdateDigitalTwin;}
```