ScaleOut Software

Unlocking the Power of Digital Twins for Streaming Analytics and Simulation of Large Systems

August 1, 2023

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

Agenda

ScaleOut Software

- A new vision for digital twins: real-time analytics and simulation at scale
- Some examples
- Why not “traditional” streaming analytics?
- Why digital twins?
- Target use cases
- Development process
- Enabling technology: in-memory computing
- Aggregate analytics
- Demo

About ScaleOut Software

ScaleOut Software

- Develops and markets software for **in-memory computing**:
  - Scales application performance and
  - Provides real-time analytical insights & simulation using digital twins
  - With proprietary in-memory data storage and computing technology
- Deep domain expertise:
  - Dr. William Bain, Founder & CEO. Bell Labs, Intel, Microsoft
  - Over 18 years in the market
  - Consistent track record of innovation and technology leadership
  - Introduced a digital twin hosting platform in 2018
- Flexible business model to meet diverse needs:
  - Fully supported software releases; on-premise or in the cloud
  - Dedicated to ease-of-use to minimize training and lower TCO
  - Choice of licensing models: perpetual, subscription, cloud-hosted

Microsoft Partner
Silver Application Development

amazon webservices Partner Network
TECHNOLOGY PARTNER

Windows Azure

cloudera CERTIFIED

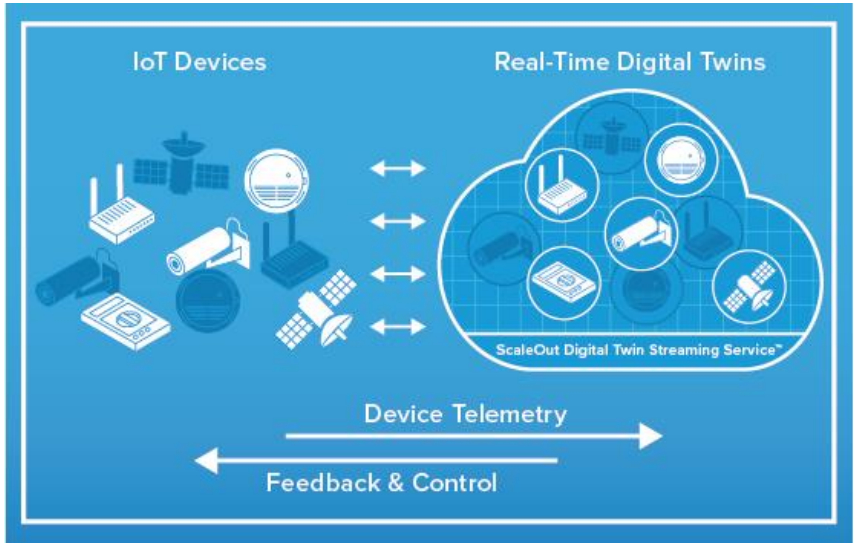
Gartner 2015 CoolVendor

ScaleOut Digital Twin Streaming Service™

ScaleOut Software

Uses a scalable in-memory compute engine to host digital twins for real-time monitoring and simulation.

- Build & deploy real-time and simulation digital twin models.
- Incorporate C#/Java code, business rules, and machine learning
- Create & visualize real-time aggregate analytics and continuous queries.
- Access an Azure-hosted cloud service or run on-premises.
- Use an intuitive web-based UI.
- Connect to data sources using Azure IoT Hub, AWS, Kafka, and REST.



ScaleOut Digital Twin Streaming Service™

A New Vision for Digital Twins

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 large collections of digital twins to track systems with many data sources:
  - Vehicle fleets
  - Logistics systems
  - Large infrastructures
  - Ecommerce shoppers

Designing a Jet Engine

Monitoring an Industrial Robot

Tracking the US Railway System

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

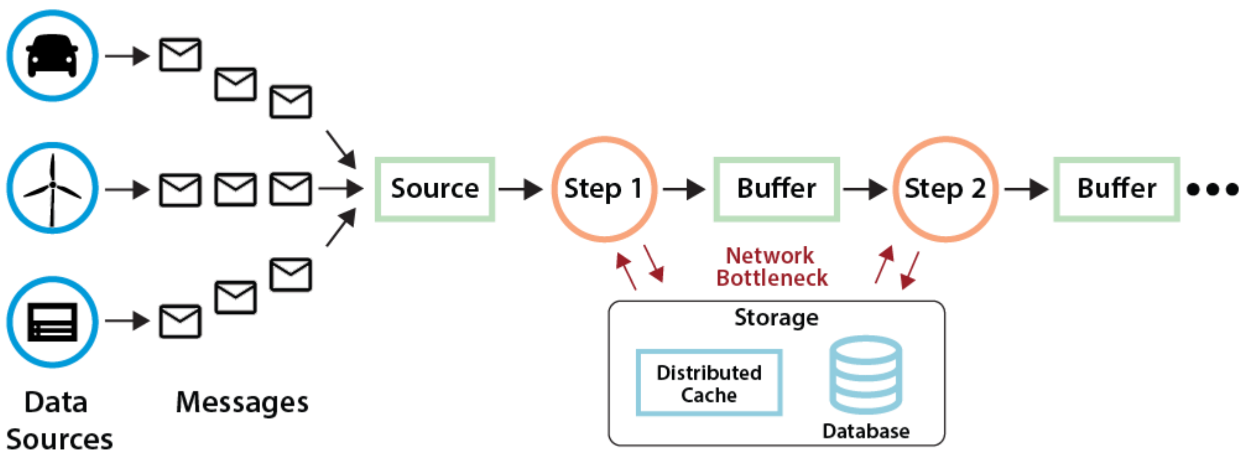
How track the safe distribution and delivery of millions of time-critical items?

- 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?

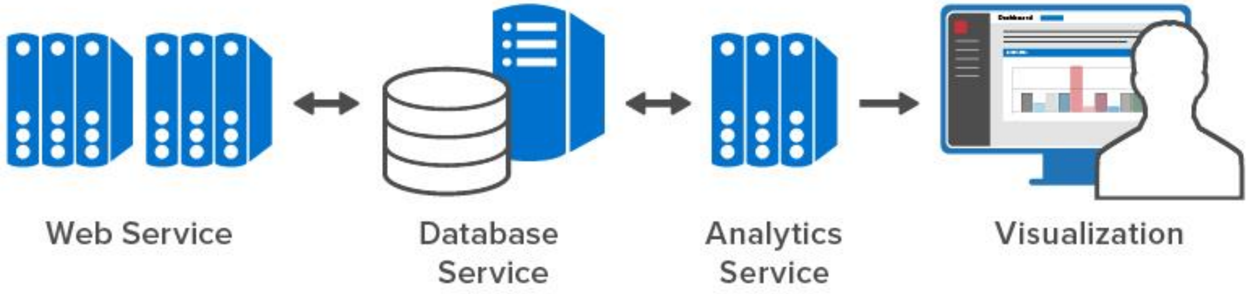
Why Do We Need **Digital Twins?**

Challenge: simultaneously track and analyze the dynamic state of 1000s of data sources

- Traditional stream-processing pipelines (e.g., CEP, Flink) cannot handle this:
  - Push all messages through a pipeline of processing steps.
  - Lack a mechanism for storing dynamic state and tracking each data source.
  - Cannot respond to individual data sources.



- Typical work-arounds (ad hoc network of services plus offline analytics) are ineffective:
  - Complex to design and implement, requiring multiple skills
  - Introduces scaling bottlenecks and availability challenges.
  - Offline analytics delay results.



Example with Human in the Loop

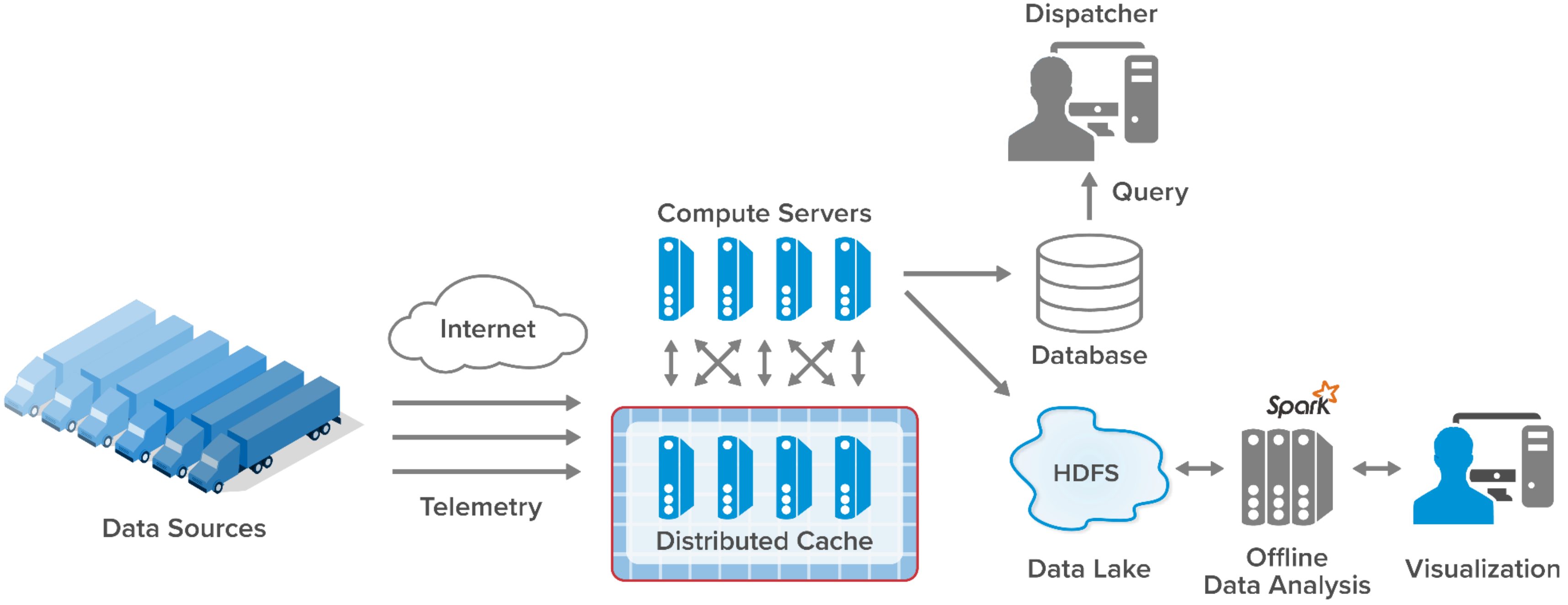
ScaleOut Software

Typical telematics systems do not:

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

As a result, they cannot:

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

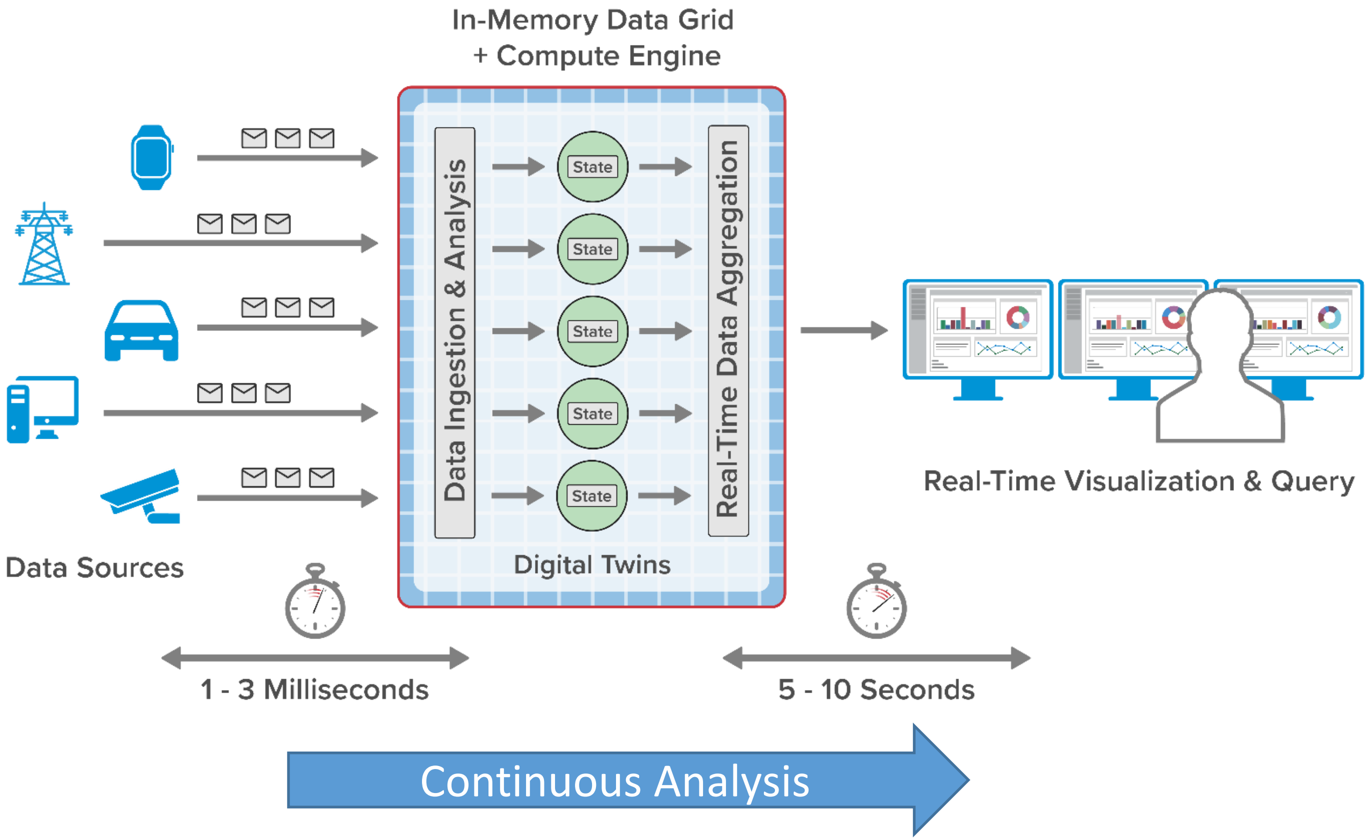


Typical Telematics Architecture for Streaming Analytics

Benefits of Using Digital Twins

ScaleOut Software

- **Deep introspection**: Track and update information about *each* data source.
- **Fast responses**: Continuously analyze incoming telemetry.
- **Situational awareness**: Continuously aggregate & visualize derived state.
- **Transparently scalable**: Seamlessly scale using in-memory computing.
- **Easy to use**: Use simple, object-oriented APIs.

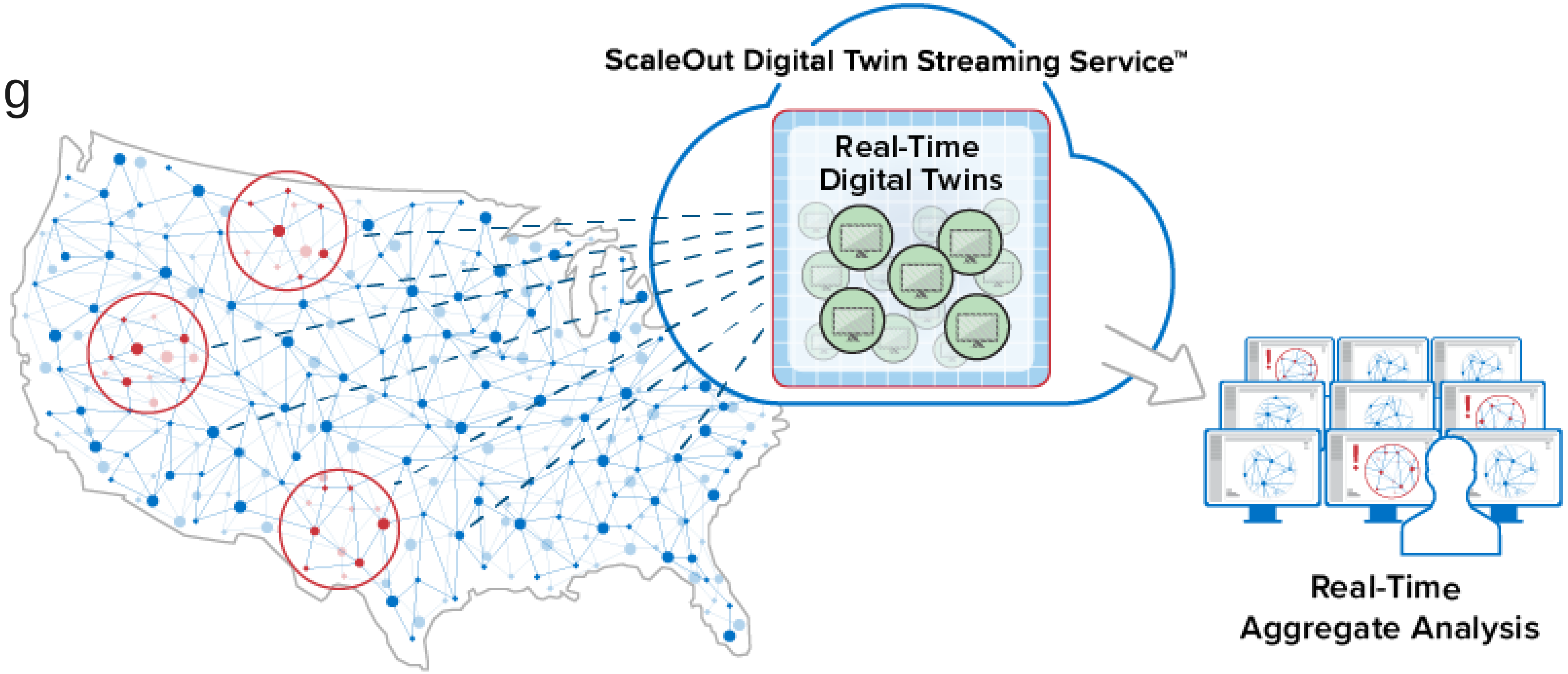


Software Architecture for Streaming Analytics Using Digital Twins

Many Target Use Cases

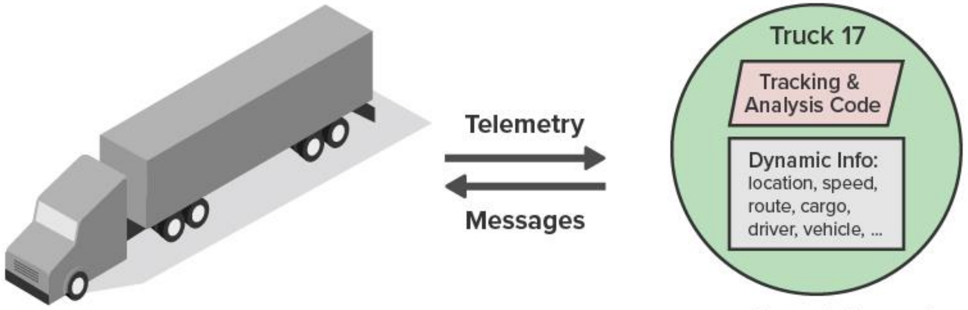
ScaleOut Software

- Applications that track **thousands of data sources which** require **fast response times, aggregate analysis,** and **situational awareness**
- General category: **real-time intelligent monitoring**
- Examples:
  - Security/safety monitoring
  - Telematics, logistics
  - Disaster recovery
  - Health tracking
  - Ecommerce recommendations
  - Fraud detection
  - IoT / smart cities
  - Transportation safety



Example: Fleet Telematics

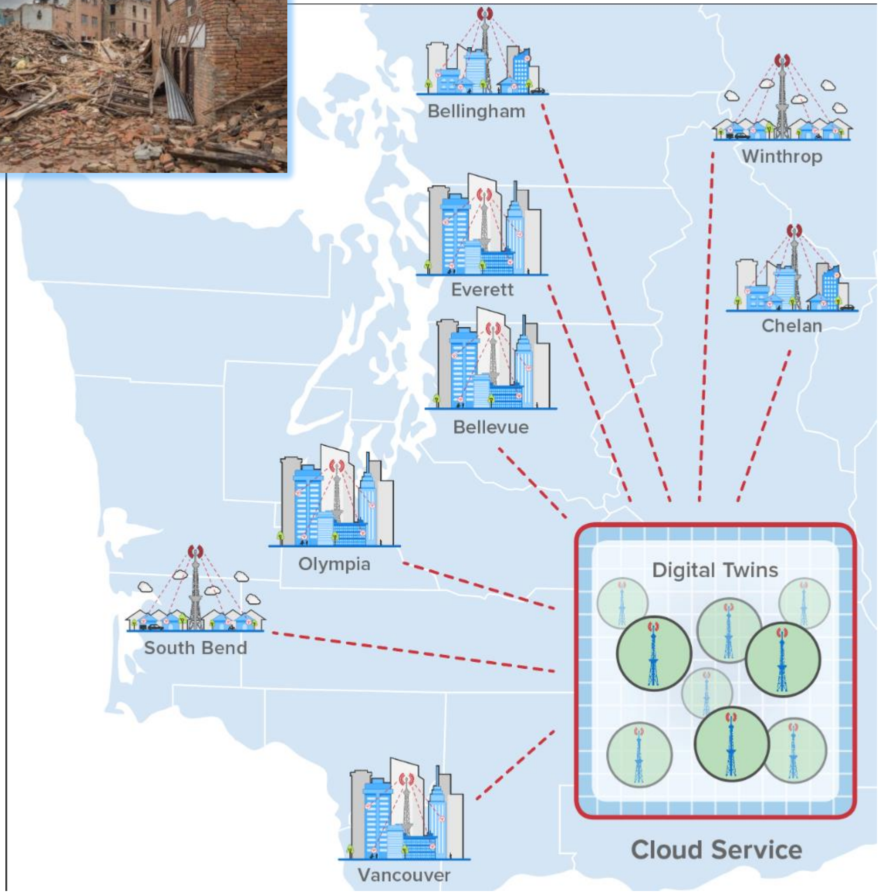
- Real-time tracking for a car/truck fleet (typically, thousands of vehicles)
- Telemetry includes location, speed, mechanical & cargo parameters.
- Digital twins add route, cargo, info on driver, service history & issues, weather, etc.
- Using incoming telemetry, digital twins can:
  - Alert driver 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

ScaleOut Software

- 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.



Also Use Digital Twins for Simulation

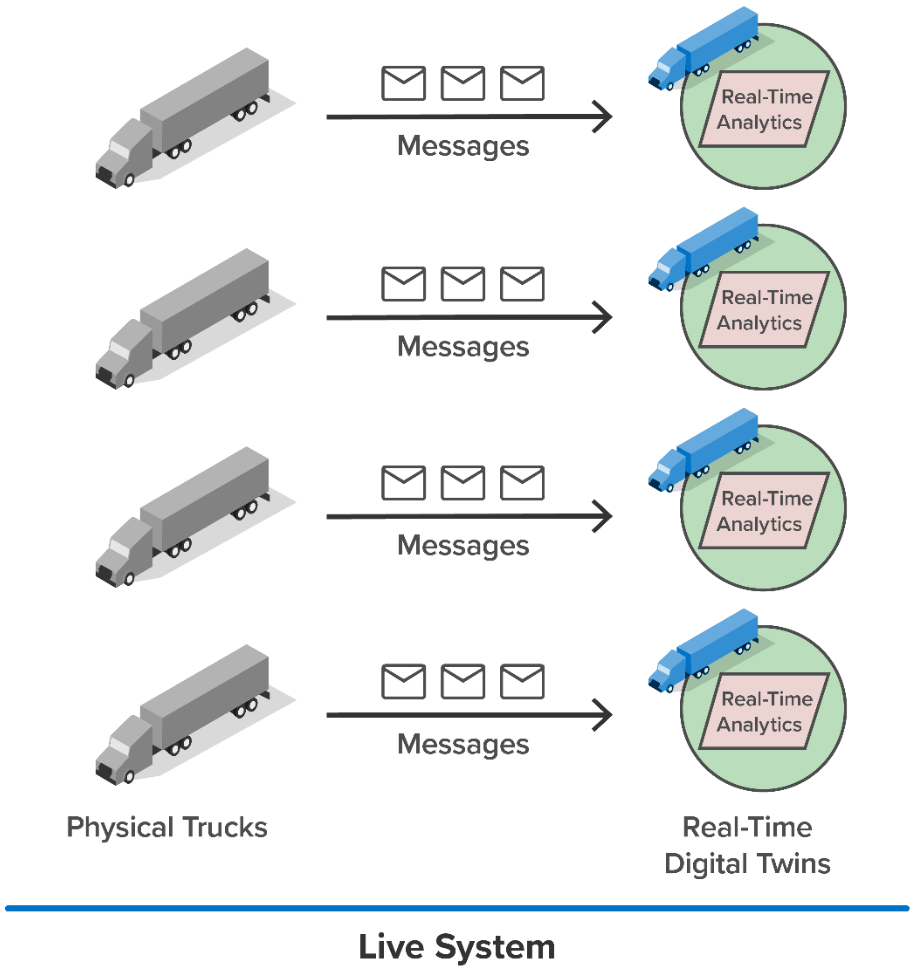
ScaleOut Software

Digital twins simplify the construction of large-scale simulations (1000s to millions of interacting entities).

One use case: a **workload generator** for testing streaming analytics.

Key benefits:

- Allows testing and validation prior to deployment.
- Simplifies application design.
- Enables seamless scaling to model large systems.



Also Use Digital Twins for Simulation

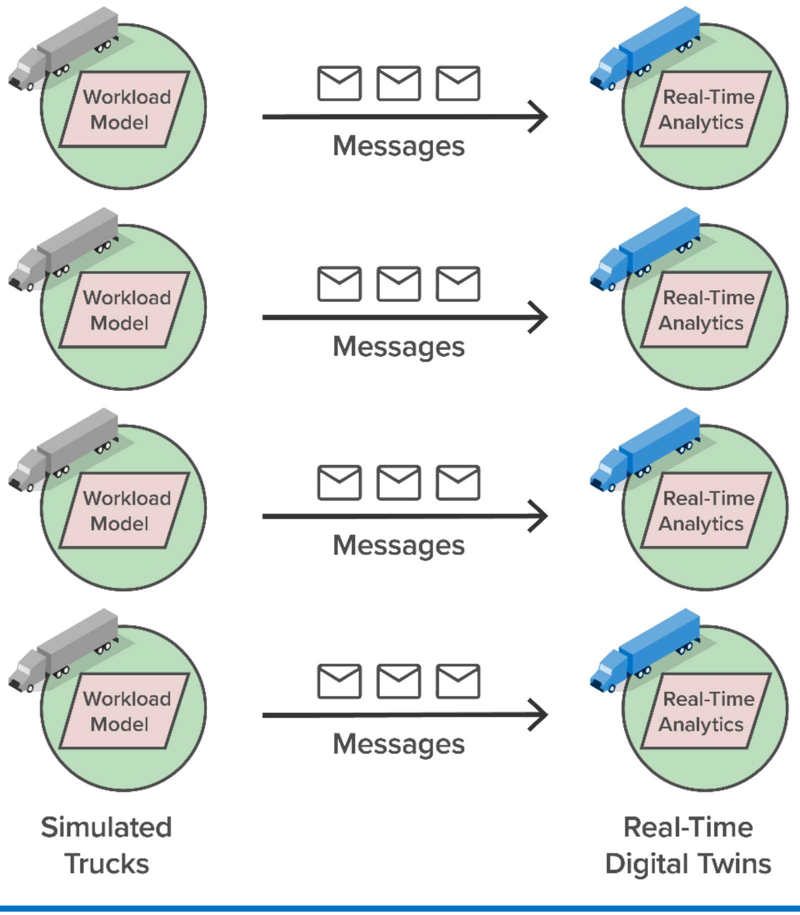
ScaleOut Software

Digital twins simplify the construction of large-scale simulations (1000s to millions of interacting entities).

One use case: a **workload generator** for testing streaming analytics.

Key benefits:

- Allows testing and validation prior to deployment.
- Simplifies application design.
- Enables seamless scaling to model large systems.



Simulation Using a Workload Generator

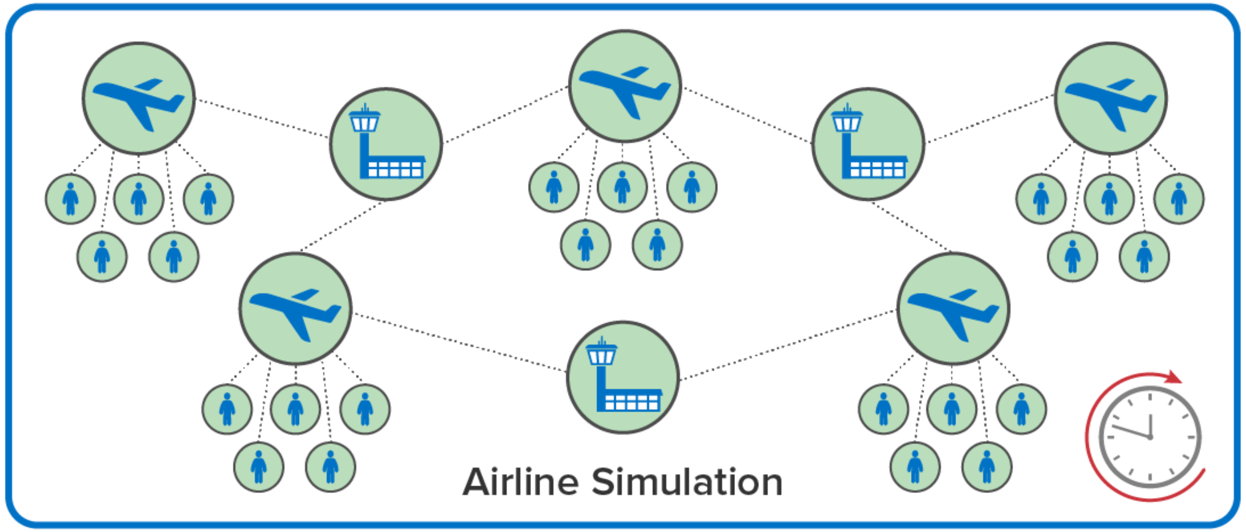
Another Simulation Use Case

ScaleOut Software

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

Example: an airline system simulation

- 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.

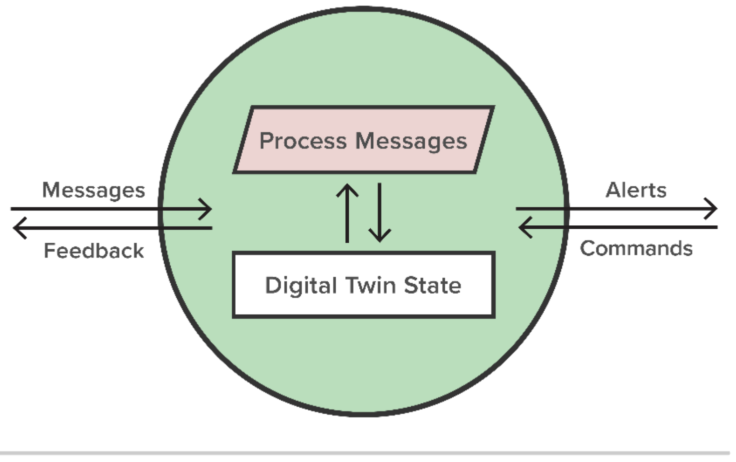


Creating and Hosting Digital Twins

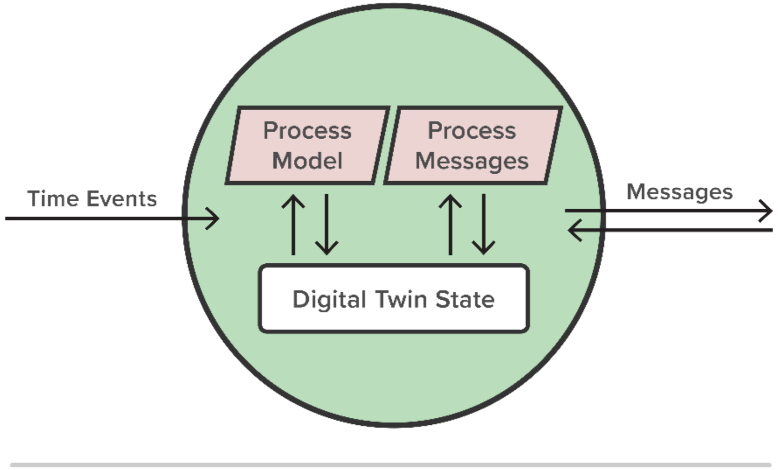
ScaleOut Software

Goals:

- Use a simple, flexible software architecture for implementing digital twin models.
- Leverage the inherent object-oriented nature of digital twins:
  - State information for each instance of a model
  - Common analytics for all instances (code, business rules, and machine learning)
- Let the platform handle the rest:
  - Create and manage digital twin instances at scale.
  - Ensure fast access to digital twin state.
  - Enable real-time aggregate analytics (e.g., map-reduce and query) for digital twin state.



Real-Time Digital Twin

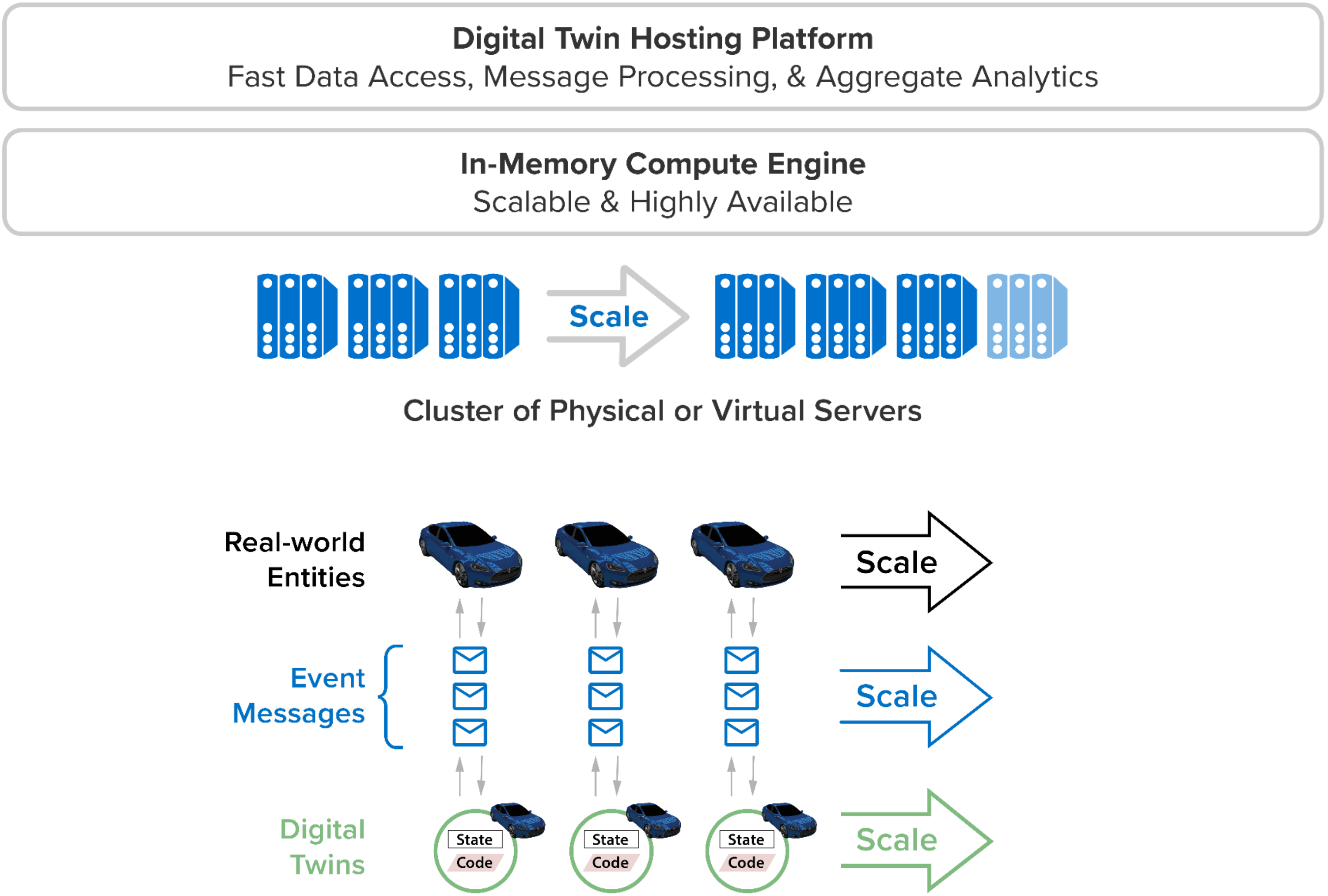


Simulation Digital Twin

Benefits of In-Memory Computing

ScaleOut Software

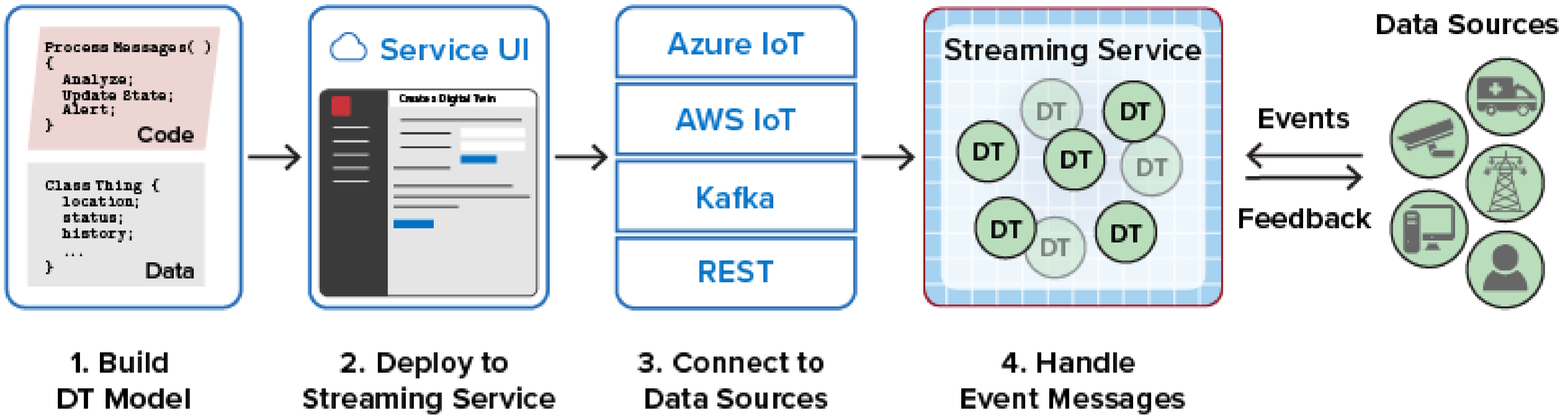
- What is “in-memory computing”?
  - A scalable platform for hosting in-memory objects with integrated aggregate analytics
  - Transparent message processing, load-balancing, scaling, and high availability
- Scales to host large populations of digital twins for both stream processing and simulation



Digital Twin Development Process

ScaleOut Software

- Application developers create one or more digital twin models and deploy them to the hosting platform using the service's UI.
- For real-time analytics, connect to data sources using popular message hubs or REST.
- For simulation, spawn initial digital twin instances and start simulation.
- Use aggregate analytics to query and visualize state of digital twins.



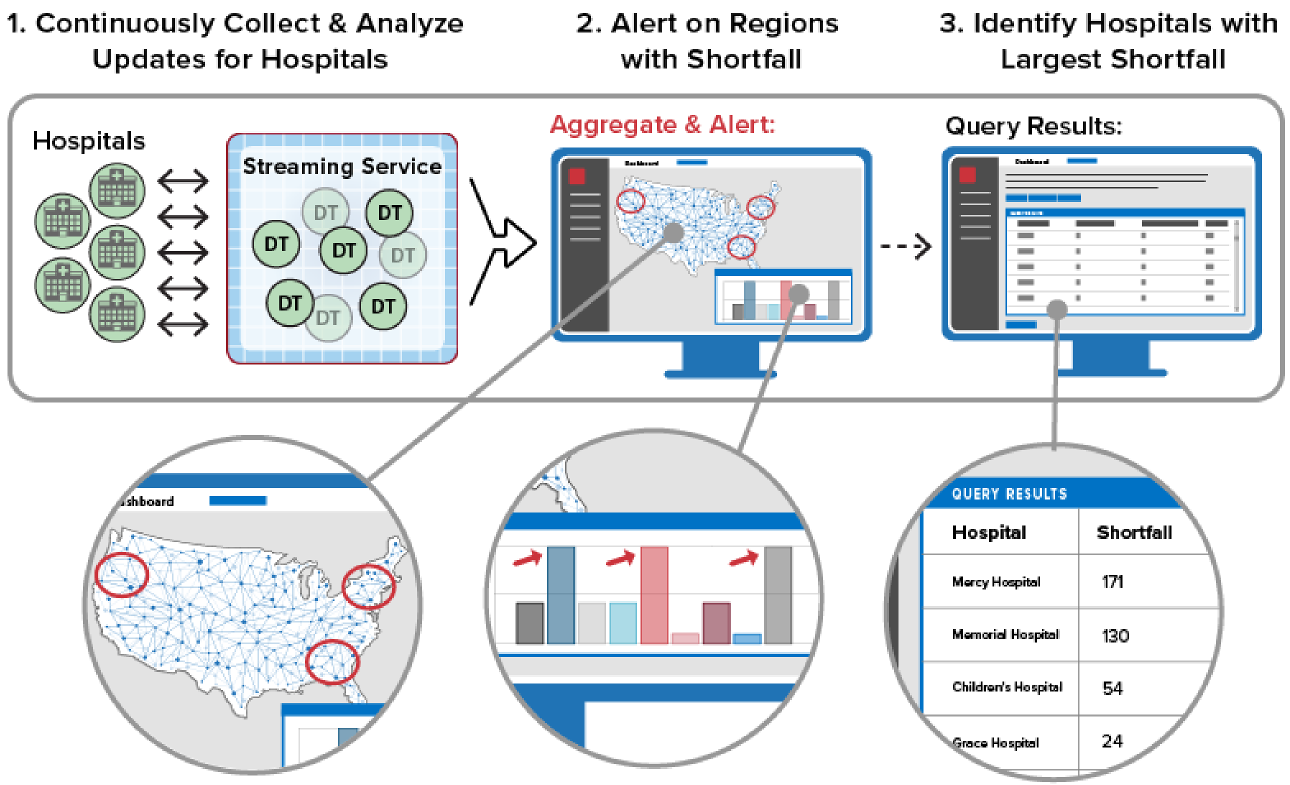
Using Aggregate Analytics & Query

ScaleOut Software

Aggregate analytics maximize situational awareness.

Example: 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.



Example: Tracking the Freight Rail System

ScaleOut Software

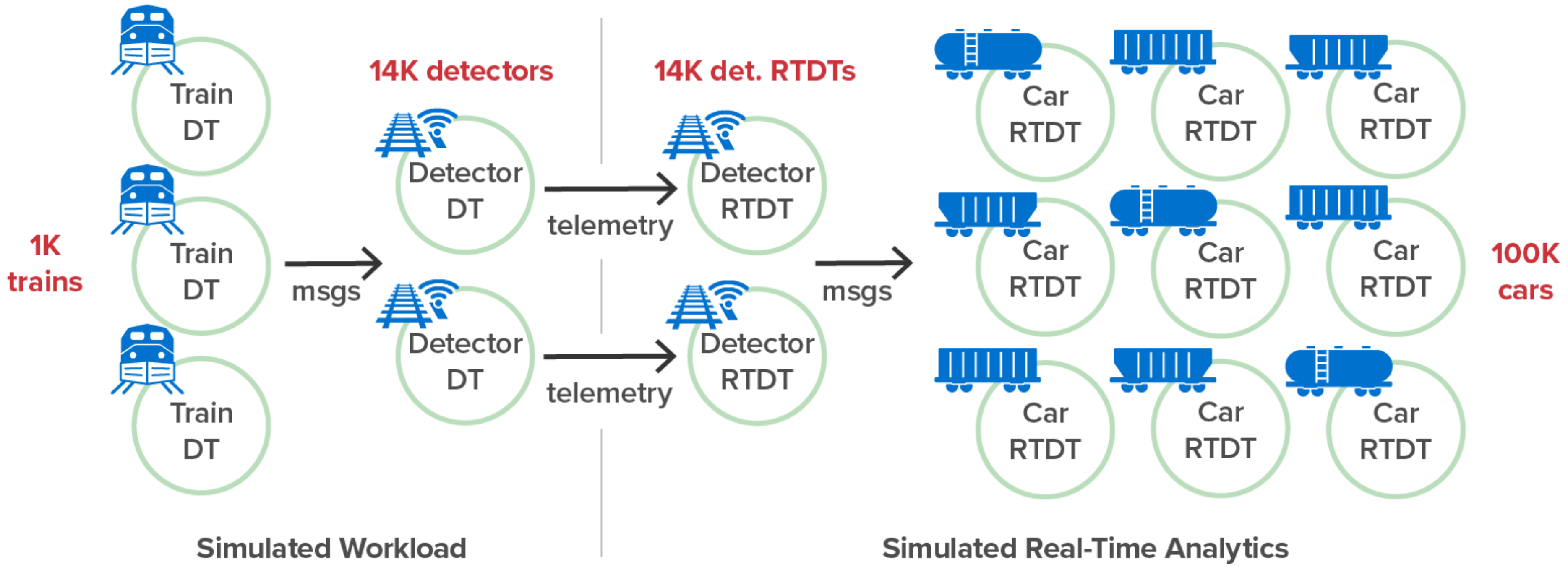
- Each year in the US, thousands of freight trains carry 1.6 billion tons of freight across 140,000 miles of track:
  - Approx. 300 trains per week
  - Approx 500K carloads per week
- In 2022, there were more than 1,100 train derailments, causing over 100 million dollars in damage.
- 6,000 hot boxes 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.

Create and Validate Digital Twin Analytics

ScaleOut Software

Goal: Implement and simulate telemetry tracking from track-side detectors and predict wheel bearing failures before an accident can occur.

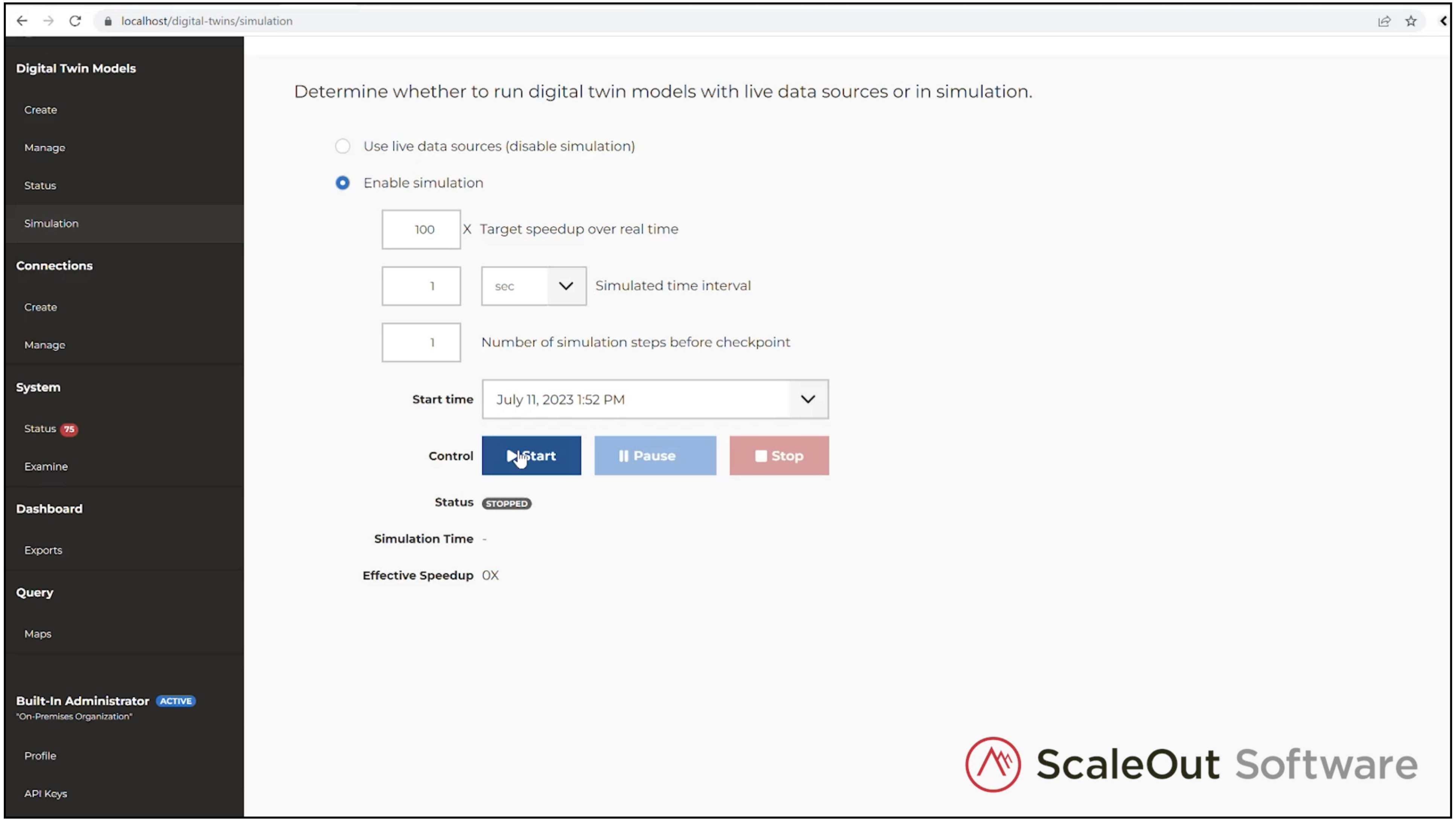
- Uses ~ 129K digital twins to both model the system and implement real-time analytics.
- Validates their ability to receive and analyze real-time telemetry from hot boxes.



Digital Twin Simulation

Demo of Train Simulation

ScaleOut Software



Key Takeaways

- Digital twins aren’t just for PLM.
- They offer a powerful software architecture for real-time streaming analytics and simulation 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

www.scaleoutsoftware.com