



The Internet of Analytics

Discovering actionable insights from
high-velocity streams of real-time IoT data

Sami Akbay, Founder and EVP, WebAction
In-Memory Computing Summit | San Francisco, CA | June 2015

Enabling Internet of Things

Cheap and Efficient Data Capture

Affordable sensors,
RFID, antennas,
aggregators,
cameras

Smaller footprint

Low energy consumption

Continuous Connectivity from Everywhere

Wired networks / wireless networks

Reliable, high bandwidth connectivity

Ubiquitous access virtually from anywhere

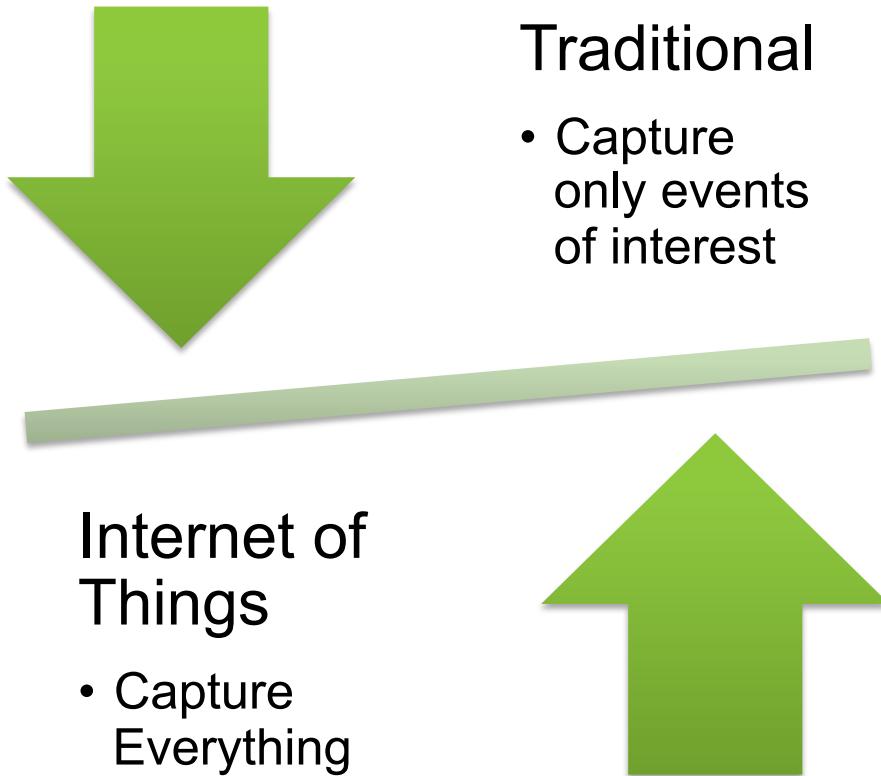
Abundant Compute Power and Storage

Faster chips

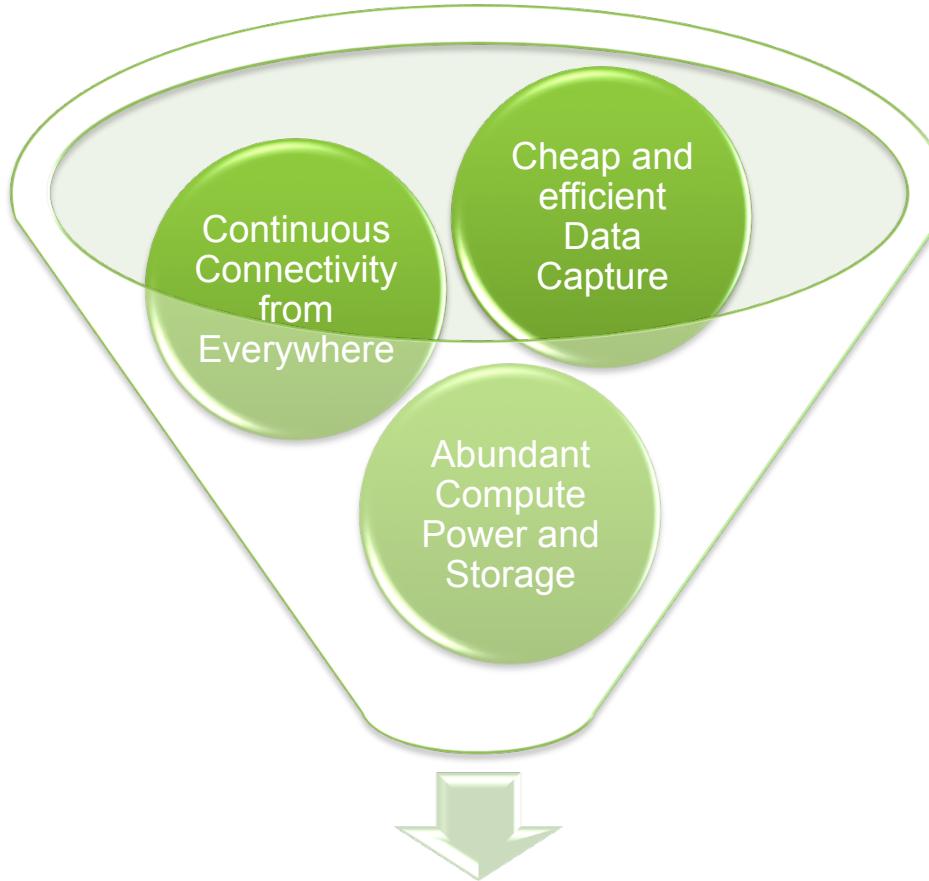
Cheaper Memory

Hadoop / big data stores

IoT Data Is Not Useful Until Analyzed



- ERP, CRM, Billing ...
 - Human generated/captured
 - Stored in Databases
 - Data inherently valuable
 - Data already useful for operations before analytics
-
- Sensor, log, location, ...
 - Machine generated/captured
 - Stored in big data frameworks
 - Most of the data has little inherent value
 - **Value of data unknown until after it is analyzed**



The Internet of Analytics

IoT Data is Fast, Big, and *Different*

IoT generates large **volumes** of data

- Expensive to store in traditional data stores
- Much of it is not useful

Data comes in spikes or **high-velocity** continuous streams

- Requires adequate connectivity
- Uses significant network resources

Data arrives in a **variety** of different formats

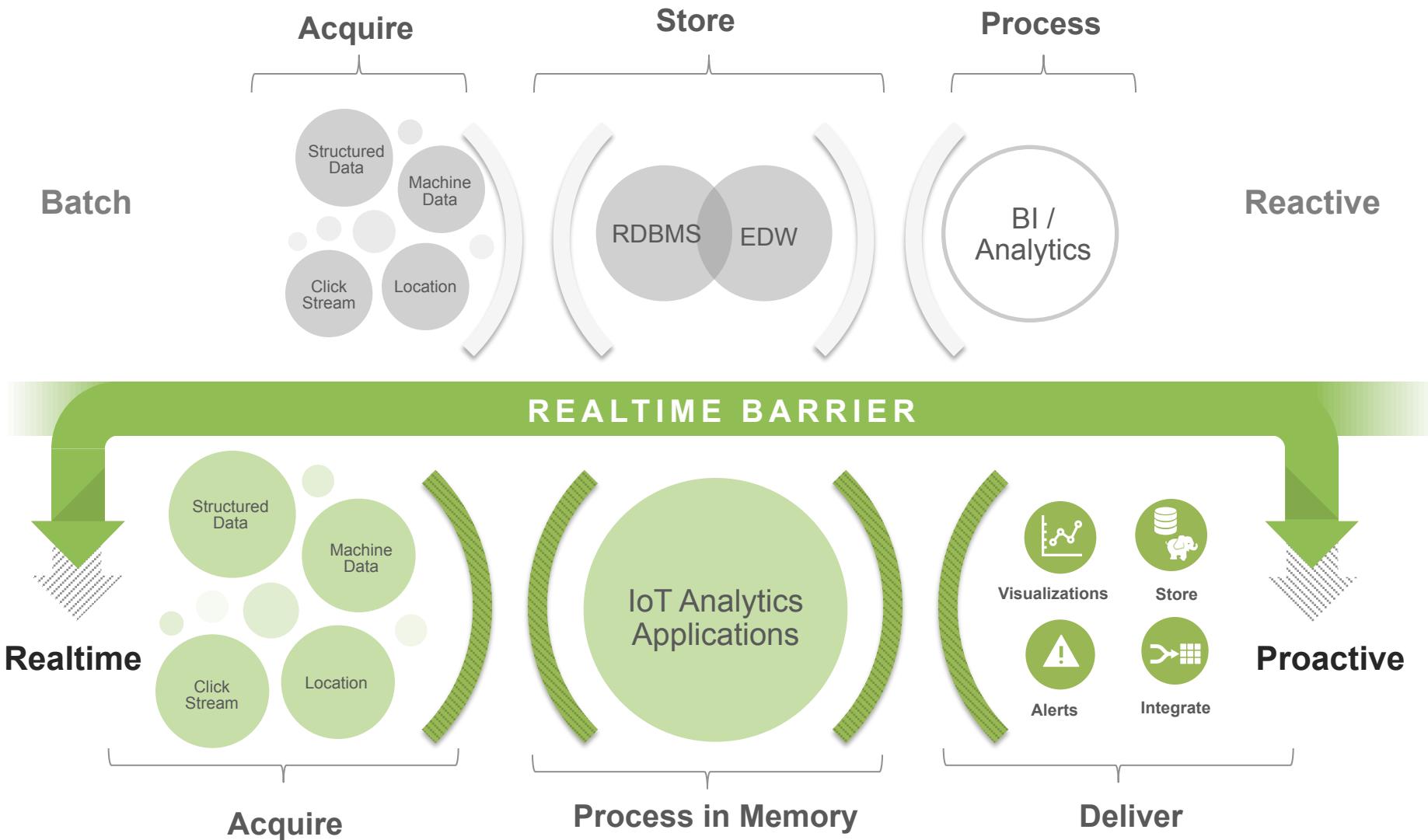
- Data transformation is required
- Data models need to facilitate analytics

Data contains **perishable insights**

- Requires platforms that can perform sophisticated Stream analytics

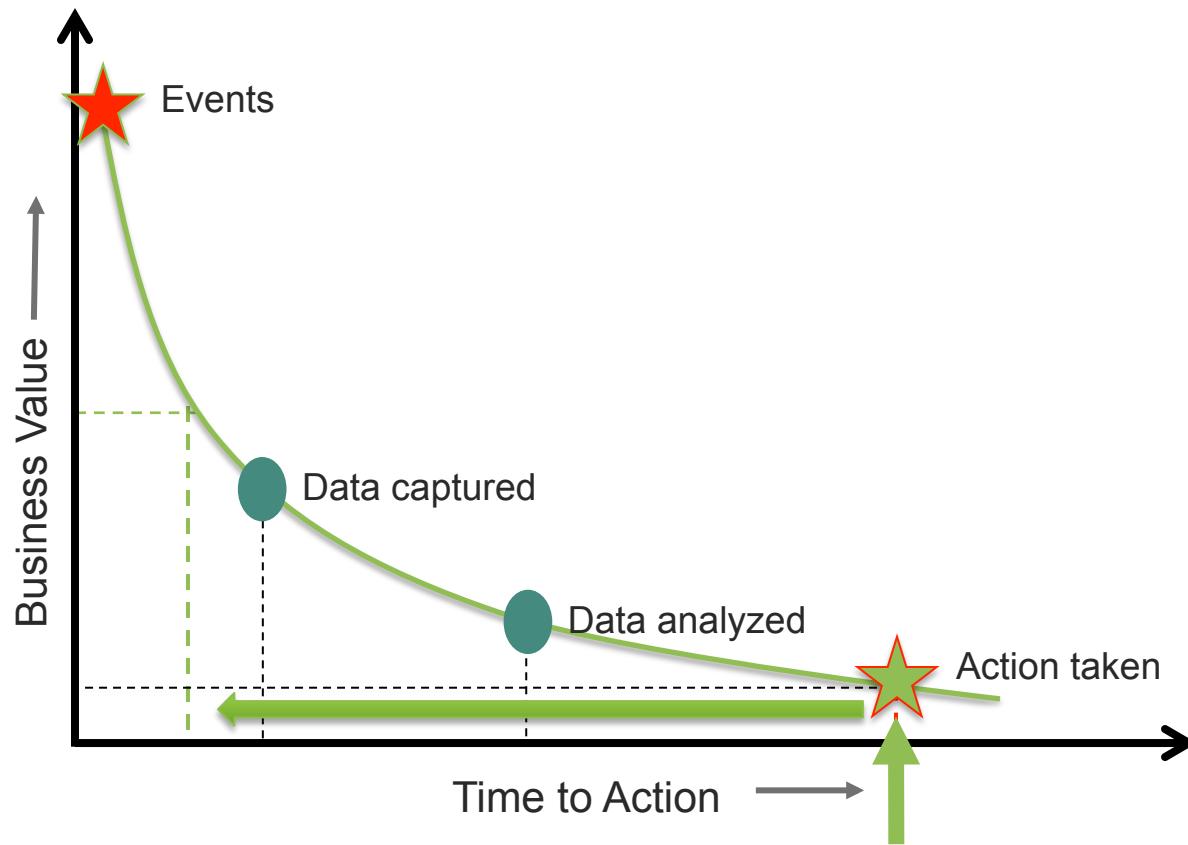
Source: *Internet Of Things Applications Hunger for Hadoop and Real-Time Analytics in the Cloud*
by Mike Gaultieri and Rowan Curran , Forrester

Paradigm Shift in Computing – 3Vs



Minimize Time-to-Value – Perishable Value

Reduce the Latency to Capture, Analyze, and Ultimately Take Action to Increase Value



The Perfect Storm

Perfect Storm: An event where a rare combination of circumstances aggravate a situation drastically



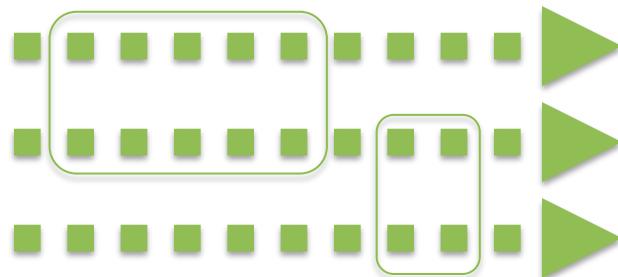
- Oil Rig Drill Sensor:
 - Temperature up 10°C → continue drilling
 - Temperature up 10°C + Viscosity down → stop drilling
- Hospital Bed:
 - Blood O₂ below 93% → Patient went to the bathroom
 - Blood O₂ below 93% and Pulse @ 150BPM → send a doctor

**You need realtime correlation of multiple data streams
to handle the perfect storm**

Need to Combine Context with Realtime

Actionable insights come from combining current events with context

$$\text{Event} + \text{Context} = \text{Realtime Action}$$



Real-time Event Stream
e.g. Real-time Sensor Events



Historical Context
+
Reference Data

e.g. Shopper profile,
Store ID, Inventory,
Profitability



e.g. Present Next-Best-Action,
update current price, modify sourcing

Why Realtime?

Because actionable insights come from combining current IoT events with context

Event	+	Context	=	Realtime Action
In the last 30 minutes a store has sold \$8,000		This store typically sells \$3,000 on Tuesdays in June		Alert the store manager to require ID at checkout
In the last hour, 2 visits by shopper X in Store Zone 3 for 16 minutes		Zone 3 has mobile phones. Shopper X due for device refresh.		Offer promotion package for new device with 2 year contract renewal
A mobile subscriber drops 3 calls in 2 hours		A subscriber will drop 8 calls in a week before becoming a churn risk		When a 611 call is made, alert the agent NOT TO offer a service discount

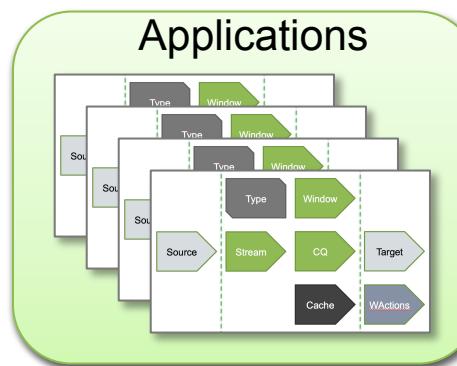
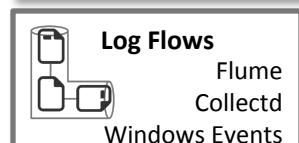
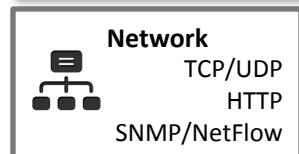
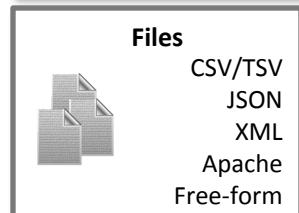
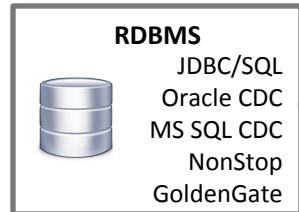
IoT Platform Requirements

- Support Data in-Motion and Data at-Rest
 - Process events and groups of events (data windows) as Streams
 - Correlate multiple Streams in Realtime before disk storage
 - Leverage analyzed context from historic data sources
 - Store aggregate data, analyzed data, and raw payload on various storage frameworks
- Implement an Easy-to-Use Development environment
 - Allow users to quickly discover and analyze data
 - Convert analysis patterns into IoT Analytics Applications
 - Provide an easy-to-use development / deployment interface
- Address industrial and operational needs
 - Offer linear scalability
 - Run on commodity infrastructure / virtualized environments
 - Provide redundancy, failover, recovery

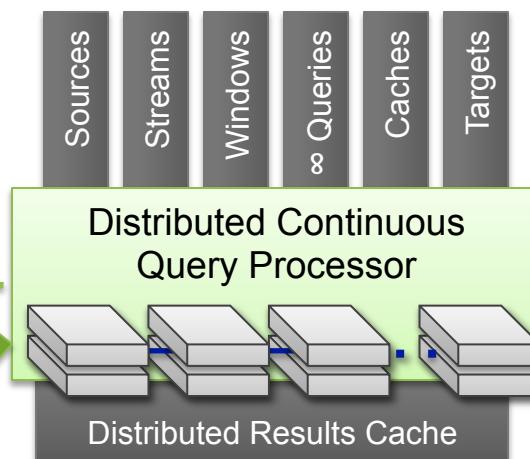


IoT Stream Analytics Ecosystem

Sources

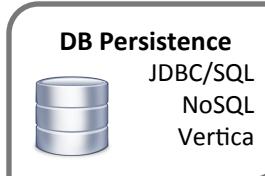


Business-Level Logic
With Tungsten QL (extended SQL)



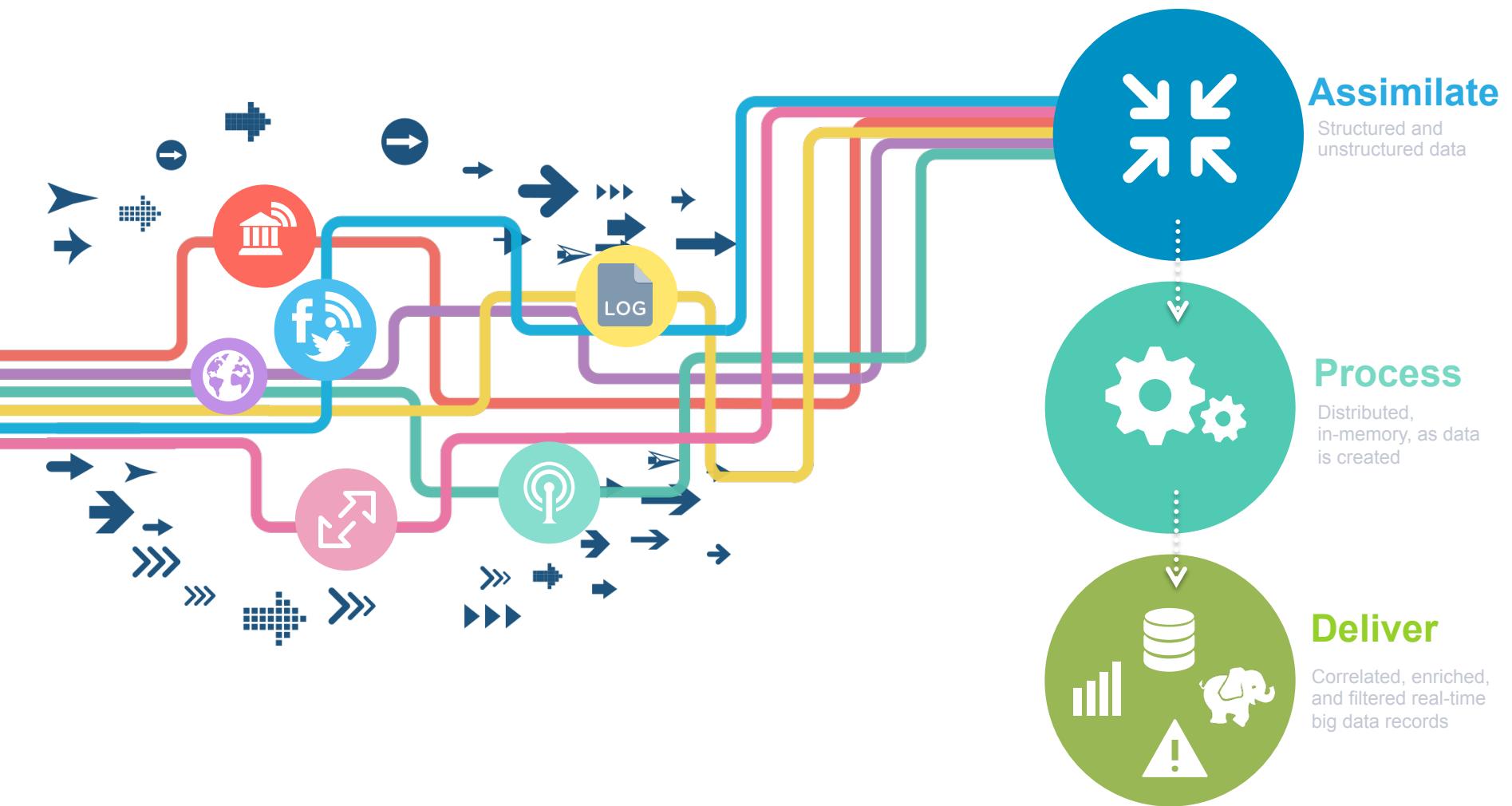
External
Context

Delivery



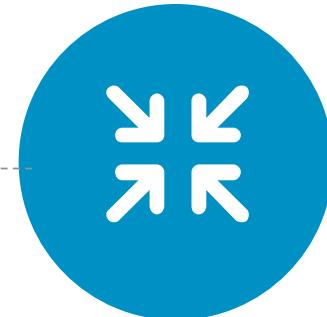
WebAction Solution

How Does WebAction Work?



Assimilate

- Data from transactional sources is acquired via redo or transaction logs
- Structured and non-Structured data
- No Production Impact
- No Application changes

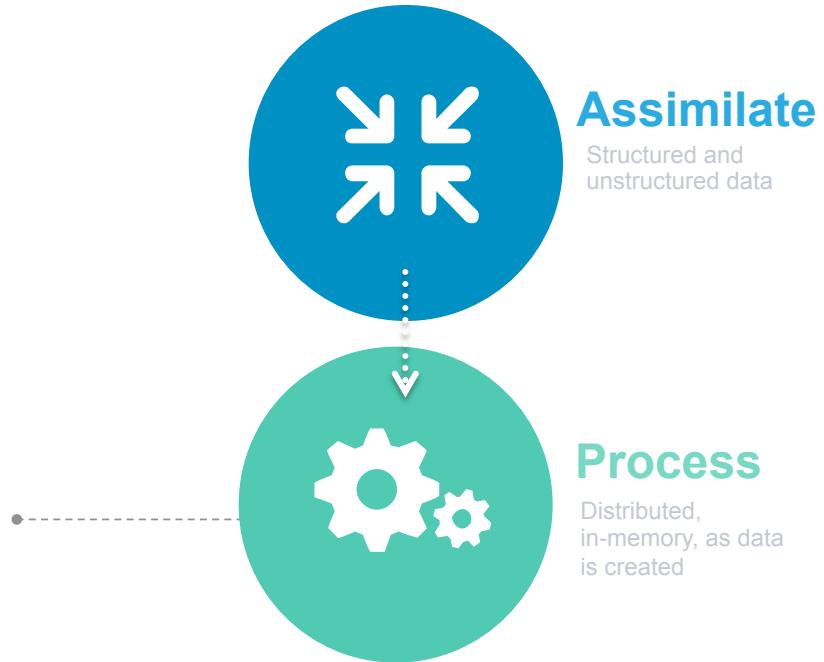


Assimilate
Structured and unstructured data

TYPE	EXAMPLE	COMPLEXITY
Common File Format	 CSV, JSON, XML	SIMPLE
Social Feeds	 Facebook, Twitter	SIMPLE TO MEDIUM
System/ IT Data	 Syslogs, weblogs, event logs	MEDIUM
Device Data	 SmartMeter, Medical Device, RFID, Netflow, iBeacon, CDR	MEDIUM
Industry Data	 SWIFT, HL7, FIX, ASN	HIGH
Real-Time Transaction Data	 Oracle, DB2, SQLServer, MySQL, HP NonStop	VERY HIGH

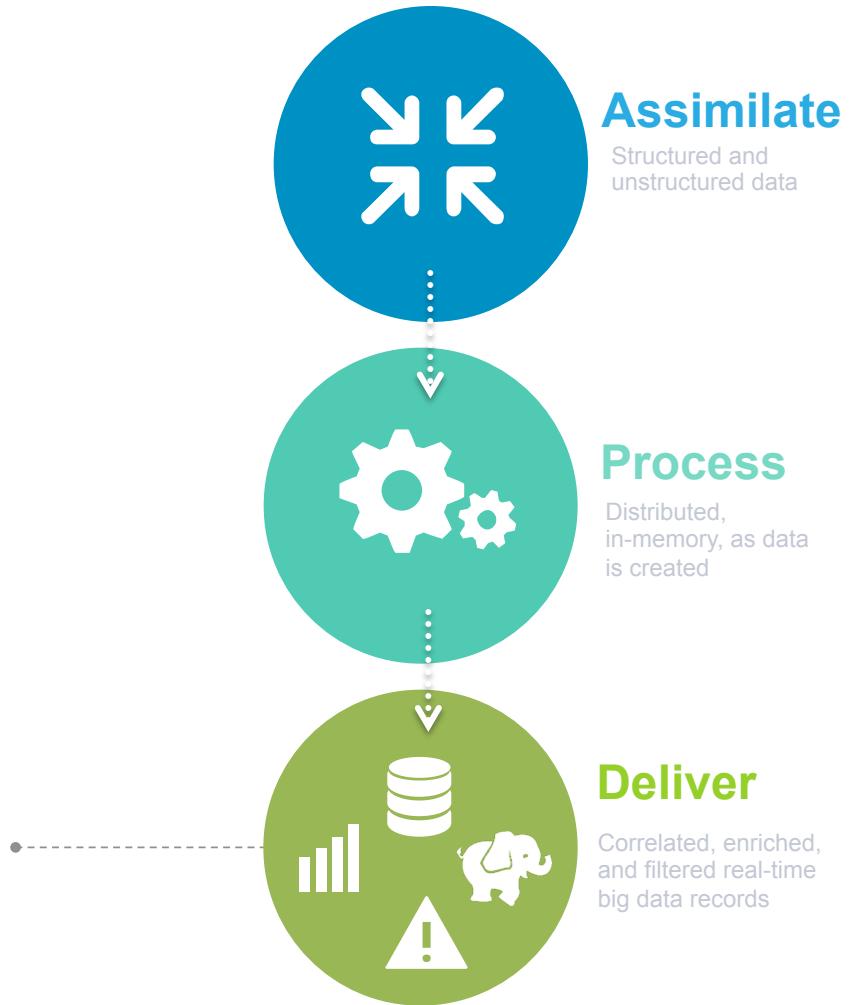
Process

- Enrich live Big Data with historical data sources
- Process Big Data faster using partitioned streams, caches, and additional nodes
- Execute SQL-like queries of in-memory Big Data
- Alert in real-time based on predictive analytic model results

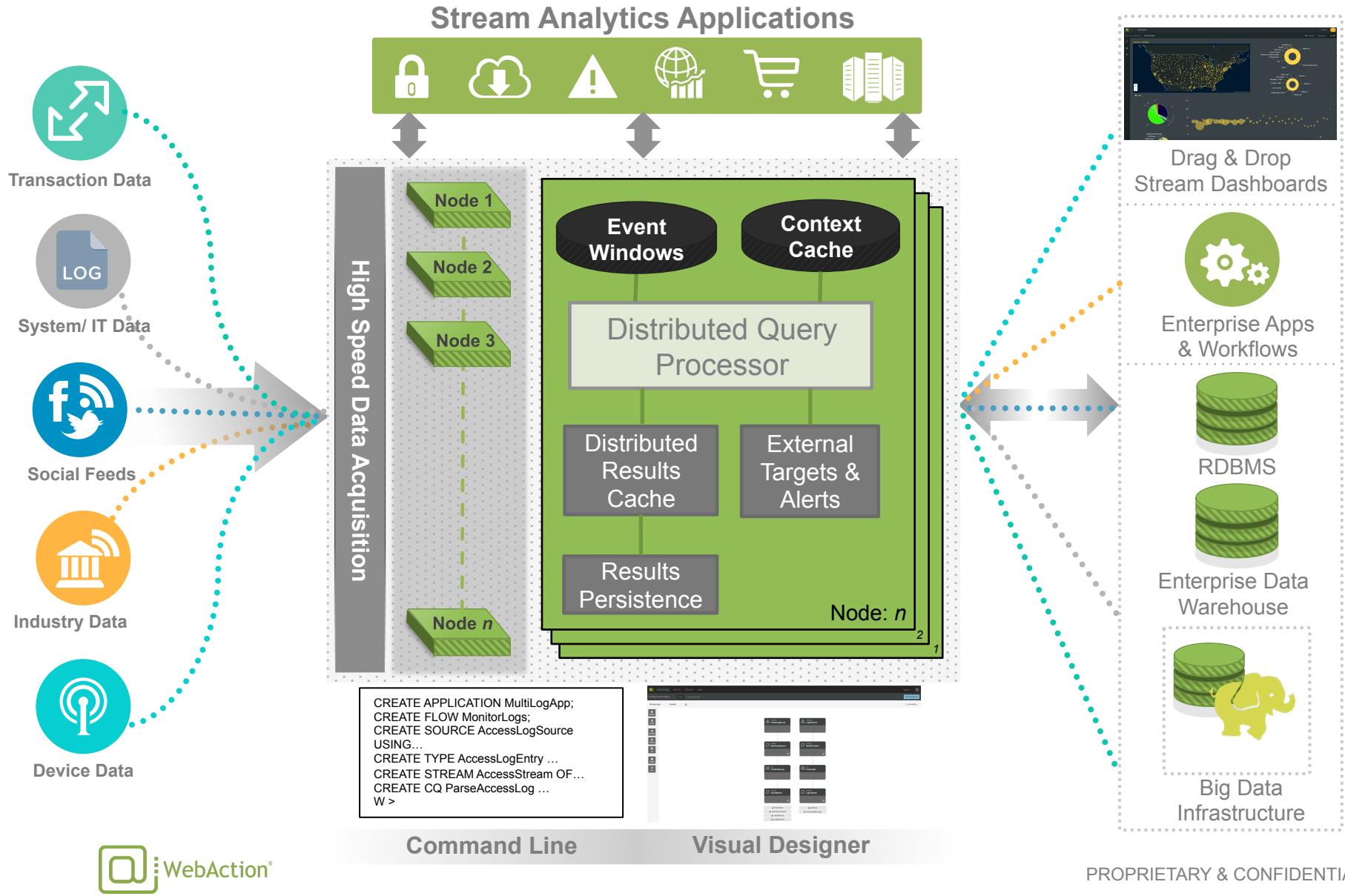


Deliver

- Continuous Big Data Records
- Realtime Drag & Drop Dashboards
- Predictive Alerts
- Business Trends
- Data Patterns
- Outliers



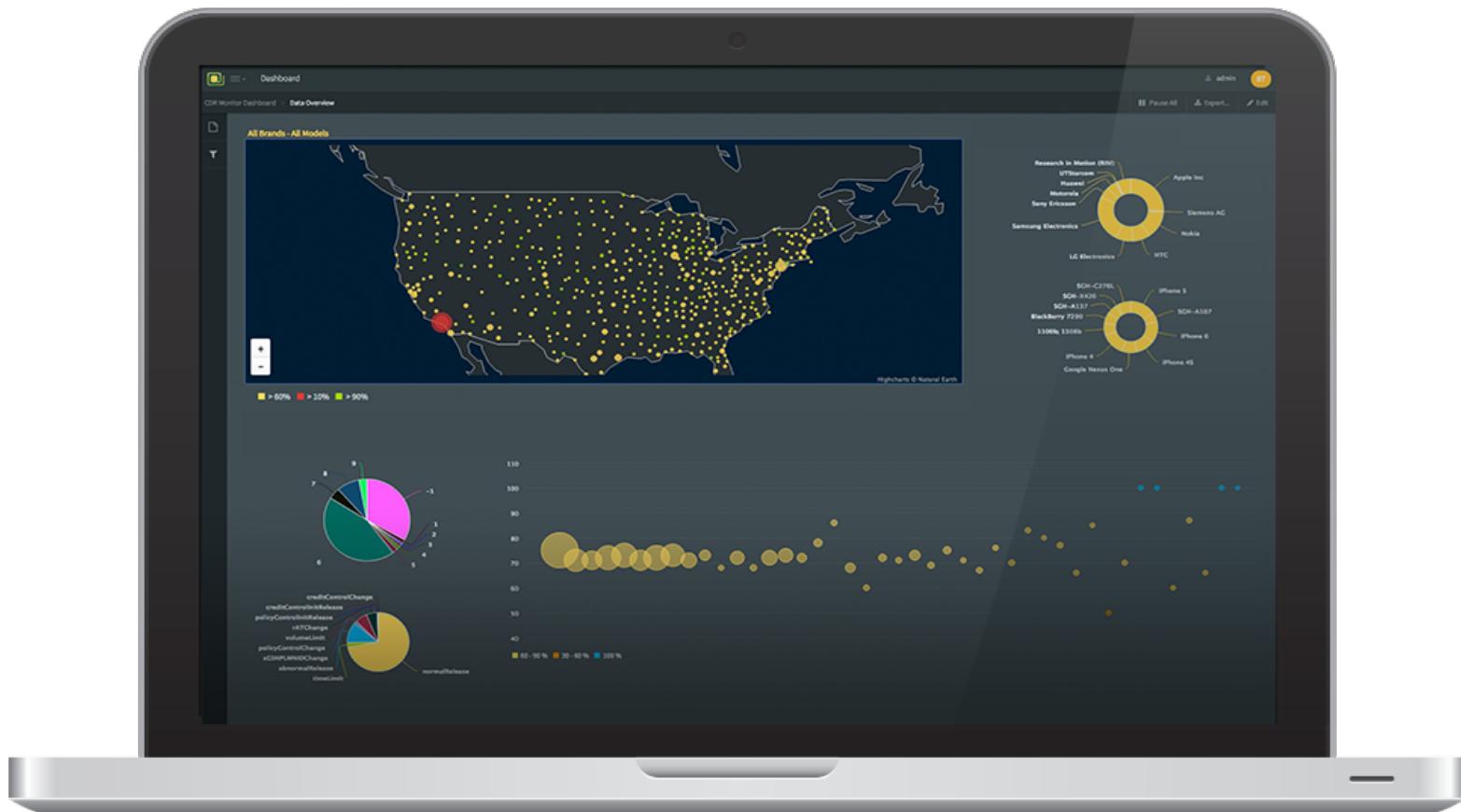
WebAction Platform Architecture



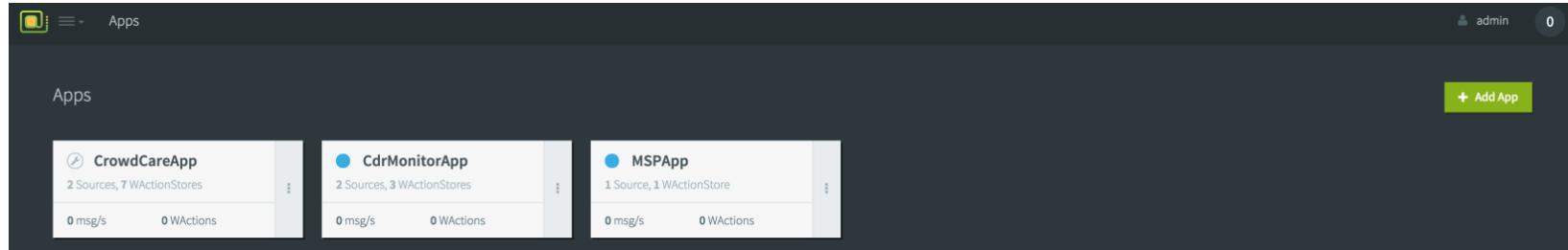
Core WebAction Components

Component Definition	
Source	Access external data and provides realtime continuous events into streams
Stream	Carries data between components and nodes
Window	Provides moving snapshot/collection of events for aggregates and models
Cache	External contextual data made available using distributed in-memory grid
(Continuous Query)	A Continuous Query emits big data records after processing realtime streaming events (can process data from streams, windows, caches, event tables, and stores)
WAction Store (big data records)	Resulting big data records from processing (aggregates, correlates, anomalies, predictions) - can be in-memory only or persisted to Elasticsearch / database
Target	Outputs realtime big data records to external systems
Application	A combination of the above components performing business logic
Dashboard	A drag and drop realtime view into stores, caches, and streams

How to build Applications



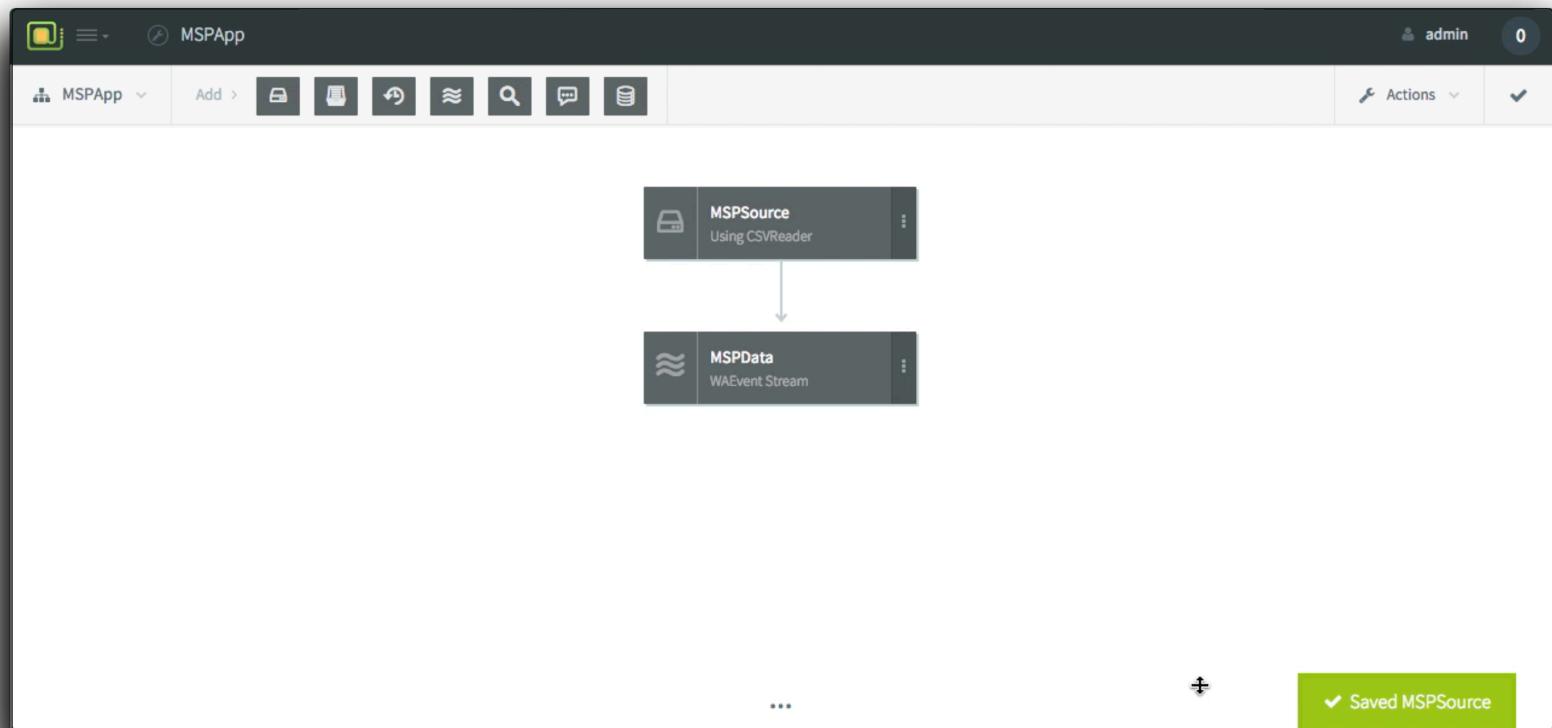
Application Developer UI



- Create Applications
- Add & Navigate through Flows
- Design Data Model
- Drag & Drop Components
- Configure Components
- Deploy Applications
- Start / Stop Applications
- View Alerts
- View Event Flow Rate

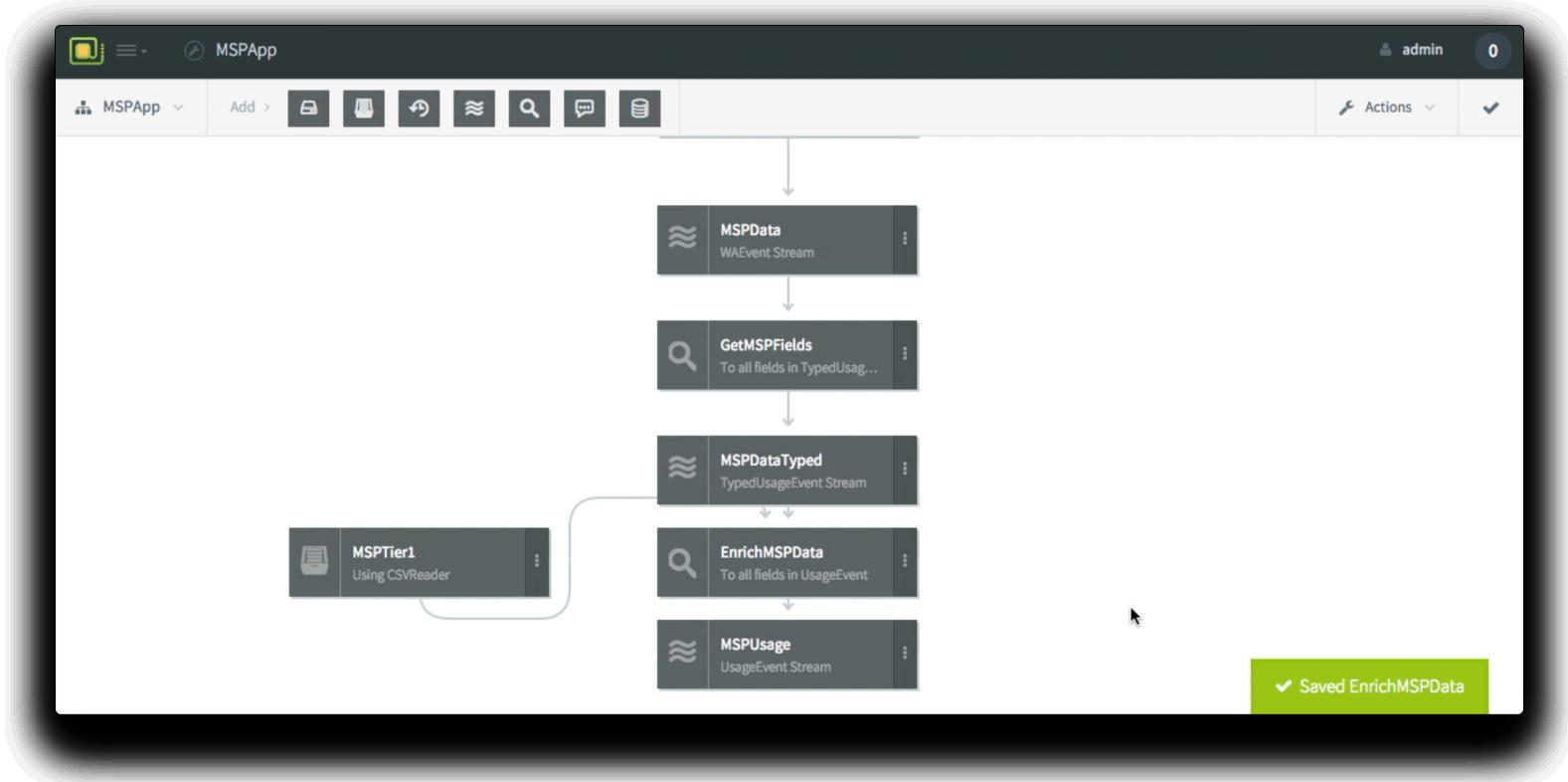
Building a Stream Analytics Application

- Add a new application
- Add a source (CDC, structured, semi-structured, etc.)
- Configure Source

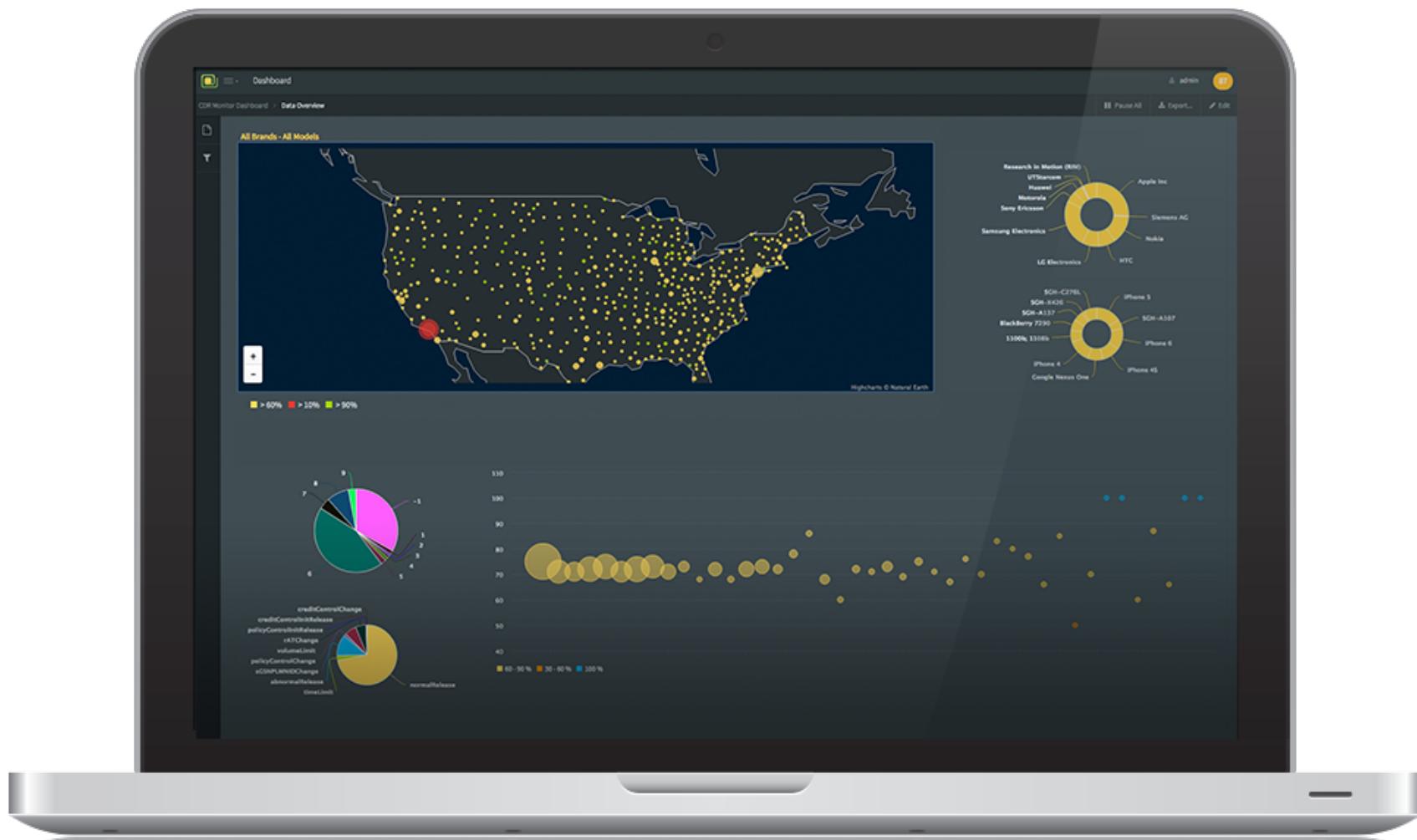


Building a Stream Analytics Application

- Add a typed stream
- Add a CQ to transform data types
- Add a Cache and CQ for context enrichment

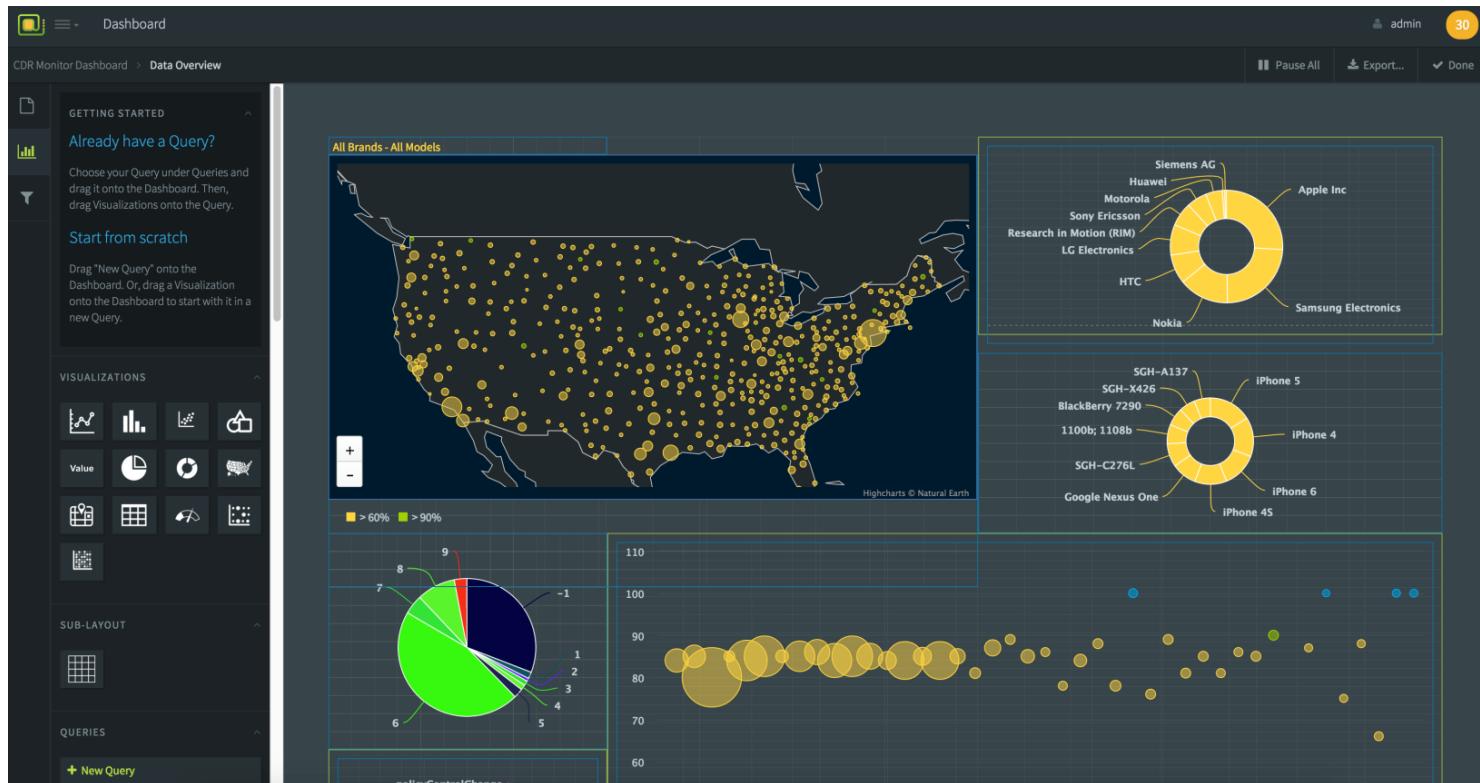


Streaming Visualizations



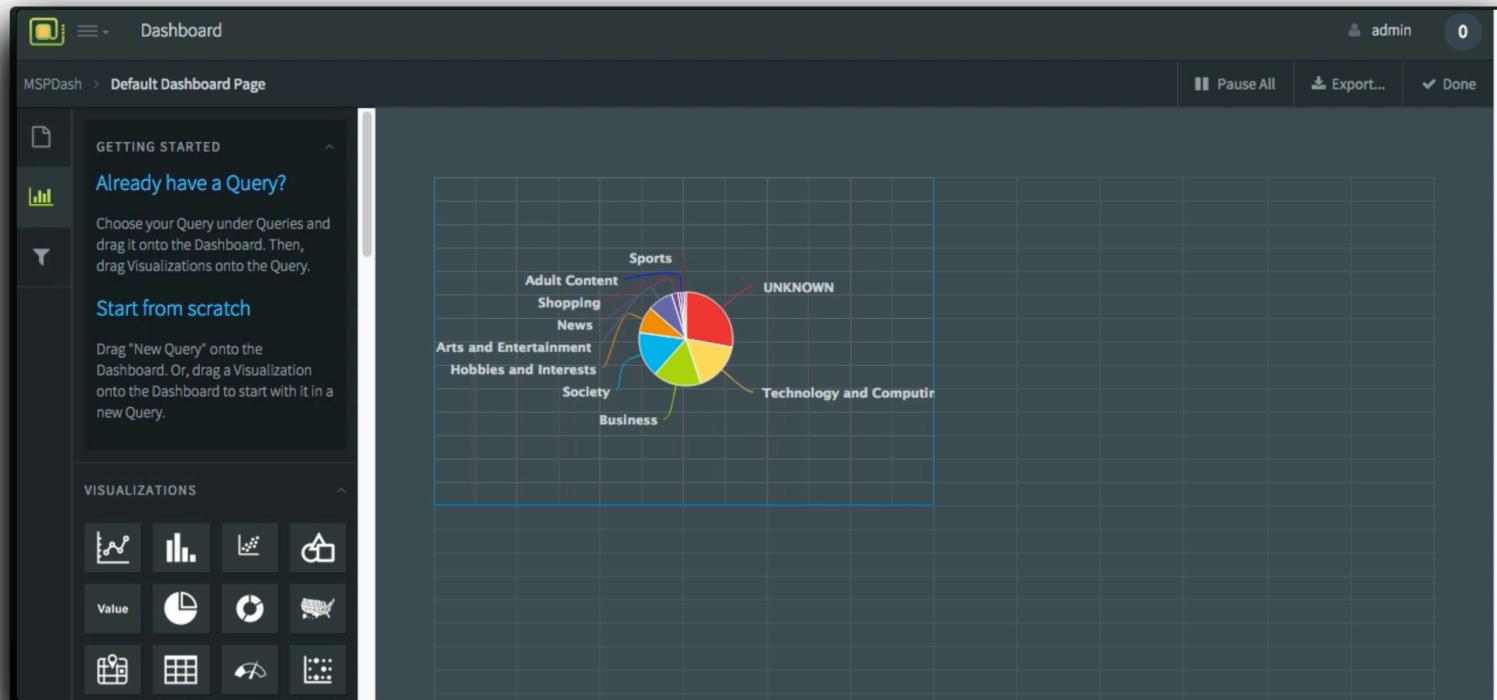
Realtime Dashboard Designer UI

- Design dashboard
- Create multiple pages / drilldowns
- Define data through queries
- Drag and drop visualizations: Values / Icons / Gauges – Maps – Line Charts – Scatter / Bubble Plots – Pie / Donut Charts – Bar Charts – Tables – Heat Maps



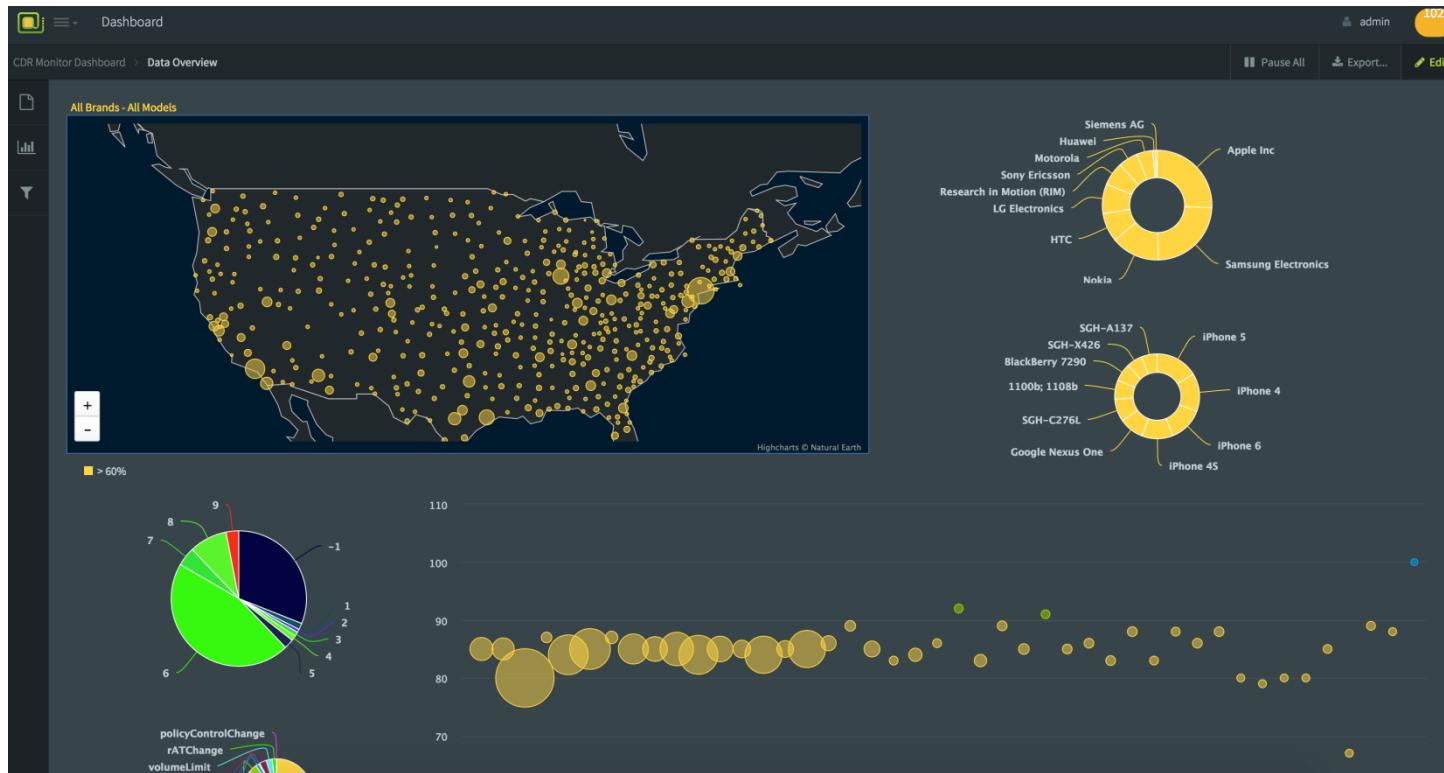
Building a Realtime Dashboard

- Add a new Dashboard
- Add a Visualization
- Configure Query and Visualization



Realtime Dashboard

- View Data Visualized
- Filter data in a page
- Drilldown to related and detail pages



Under the Hood

- Realtime log / database CDC reading in addition to push sources like TCP/JMS
- Bytecode generation for data types and query processing
- Scaling across multiple nodes with flexible deployment
- Auto failover of application components from one node to another
- Nodes can be added and removed while applications are running
- Recovery ensures no events are missed or processed twice
- Recovery takes window contents into account
- Role based security at the application through component level
- Integrated realtime dashboard visualizations using server push

Continue the Conversation...



WebAction Headquarters
info@webaction.com
+1 (650) 241-0680



[LinkedIn.com/company/WebAction-Inc-](https://www.linkedin.com/company/WebAction-Inc-)



facebook.com/WebActionSoftware



[Twitter.com/WebActionInc](https://twitter.com/WebActionInc)