

Accurate Data and Insights — Faster

Agile, Continuous, Reliable Data Ingestion and Transformation Solution

Data Drives Business Success

Data is opportunity. And time is money. To succeed in today's data-driven economy, businesses must capture, store, and protect relevant data—and extract value from it in real time, despite volume and complexity. Success hinges on access to data and on strong data analytics. Is your business set up for success?

Traditional databases are too slow and inefficient in ingesting streaming data. As a result, data scientists can spend 50 to 80 percent¹ of their time collecting and sanitizing data in data stores before it can be used to derive meaningful business insights. Those delays have consequences; incomplete or delayed data can seriously impact decision-making, often at great cost to the organization.

Key data challenges that businesses using legacy systems face include:

Timing

Legacy systems can cause delays in both data ingestion and insights derivation. The ideal is to execute ingestion pipelines as soon as data is generated—to better meet mission-critical SLAs and fulfill customer expectations.

Scalability

Is your system able to process billions of events per second now and scale seamlessly to meet future needs? Most legacy systems aren't.

Data Quality

Quality is paramount to accurate insights. Systems should ingest accurate data in the correct format without any data loss, preparing you to make more informed business decisions.

Operational Visibility

It should be quick and easy to manage and monitor the health of data pipelines and applications.

Talent

Building effective datasets and data pipelines takes skill and expertise.

Businesses thrive with an enterprise-grade solution that can help them ingest relevant data from increasingly heterogeneous sources, transform it to fit their analytical needs, and load it into data stores—all in near real-time. Prepare your business to meet today's data challenges with the DataTorrent Ingestion Solution.

DataTorrent Ingestion and Transformation Solution

Better data, delivered faster, leads to better business decisions. With the DataTorrent Ingestion Solution, businesses can filter, aggregate, transform, and enrich high volume, multi-source moving data streams with ease and without any data corruption or data loss. Powered by the open-source Apache Apex engine, the DataTorrent Ingestion Solution is a fast, scalable, fault-tolerant, in-memory solution designed to meet today's data demands in real time. With it, you have a production-ready, code-free solution to better meet SLAs and customer needs.

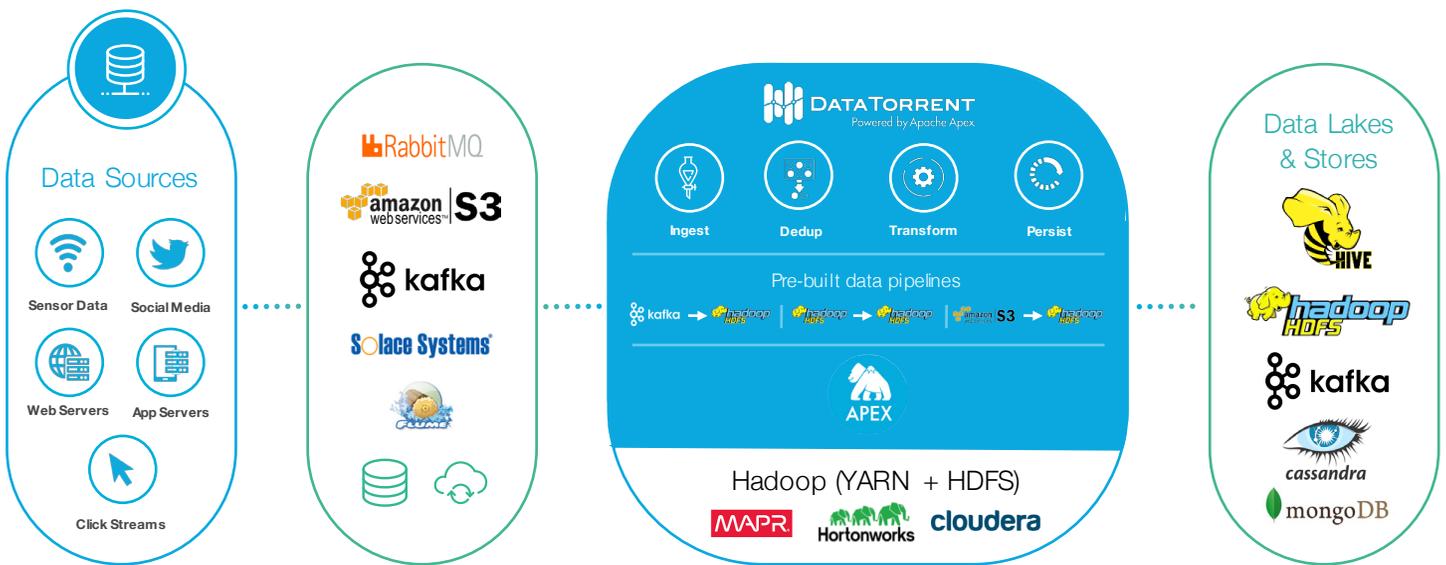
Rapid Results, Ease of Use

Our solution allows you to rapidly develop and configure data pipelines as well as to ingest data from a variety of data sources, including relational databases, sensors, log files, message queues, and more. And, the DataTorrent Ingestion Solution delivers analytics-ready data to different types of systems instantly with ease. Depending on your needs, DataTorrent offers the flexibility to deploy the ingestion solution either on-premise or in the cloud.

The DataTorrent Ingestion Solution includes these features:

- ✓ Comprehensive library of pre-built operators to ingest any type of data quickly and easily in and out of Hadoop, NoSQL, and relational databases without writing any code
- ✓ Pre-built data pipelines (HDFS to HDFS, Kafka to HDFS, Amazon S3 to HDFS) to accelerate the application development cycle
- ✓ Intuitive drag and drop user interface along with compositional APIs and declarative APIs
- ✓ Built-in transformations including dedupe, filter, sanitize, and enrich
- ✓ Dynamic topology modification to change data flow and add or remove operators with no downtime

¹“For Big-Data Scientists, ‘Janitor Work’ Is Key Hurdle to Insights,” The New York Times, Technology, August 2014



The DataTorrent Difference

Better Data, Better Business Decisions

The DataTorrent Ingestion Solution ensures that all business-critical data arrives on time and that you're able to use it to drive key business insights.



Get accurate, consistent data faster

- Pre-built data pipelines along with drag-and-drop application assembly enables you to ingest high-volume streaming data from a variety of sources within milliseconds
- Built-in transformations including dedupe, filter, sanitize, and enrich help you ingest and cleanse data while in motion, rather than in data stores



Ease of use

- Build powerful ingestion pipelines without writing code, so you can focus on meeting customer needs
- Ingest data from a broad spectrum of data sources, including message queues, sensors, web/app servers, file systems, and relational databases



Fast time to market

- Accelerate time to market from months to weeks with proven, production-ready solution and real customer-use cases
- Self-service, role-based accessibility for a broad set of developers, including Data Engineers and Java Developers, ensures fast application development



Mission-critical monitoring

- Web-based management and monitoring interface allows developers and operators to closely monitor data pipelines and application performance
- Ready-to-use visualization tool allows business users to create powerful interactive, customizable dashboards to monitor application output in real-time



Enterprise-ready

- Enterprise-grade scalability, fault tolerance, and security to help you meet growing data traffic demands, prevent any data loss, and protect data
- Easy integration with a variety of data stores, including HDFS, Hbase, Hive, Kafka, MongoDB, Cassandra, and relational databases

Imagine what you could achieve with real-time, powerful data-in-motion ingestion and transformation.

[Download](#) the Sandbox, Enterprise, Cloud, or Community version of the DataTorrent platform, and explore our difference today. Learn more about our products and solutions at www.datatorrent.com.