Solve the Big Data challenge in real-time
SQLstream transforms streaming Big Data into real-time business value. High data volumes are processed with ease, and answers delivered with ultra low latency, even when the data arrival rate is high.

Built on a standards-based platform for massively scalable real-time data management, SQLstream is the only platform that offers streaming integration with location-based and real-time analytics from the live data as it streams past.

SQLstream integrates seamlessly with Hadoop and other popular high performance data warehouses. Log file, clickstream and sensor data are collected in real-time, analyzed 'on the wire' and then streamed continuously to your Big Data stores for further analysis.

Why SQLstream for Streaming Big Data?
SQLstream is the only mature, 100% SQL-compliant solution for real-time Big Data analytics and integration. SQLstream is built on patented technology for high volume, high velocity and low latency distributed stream processing. Input data is cleansed, transformed and analyzed in real-time, enabling businesses to respond to each and every record. With SQLstream, you will be able to:

- Manage the three Vs of Big Data in real-time - Volume, Velocity and Variety
- Dramatically reduce the time to deliver Big Data solutions
- Perform complex location-based and time-based analysis, such as monitoring rate of change and varying averages
- Discard uninteresting and redundant data before the store-compute loop of map/reduce

Streaming for Hadoop
Apache Hadoop offers highly scalable, resilient data processing over commodity hardware. It is ideally suited to the scaling of large scale data problems over hundreds or thousands of servers. However, Hadoop does not support, nor was it designed to support, real-time streaming data processing and analytics.

<table>
<thead>
<tr>
<th>Hadoop Map Reduce</th>
<th>SQLstream Streaming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unlimited scalability</td>
<td>Unlimited scalability</td>
</tr>
<tr>
<td>Batch processing</td>
<td>Stream processing</td>
</tr>
<tr>
<td>Historical queries</td>
<td>Continuous queries</td>
</tr>
<tr>
<td>Shared ‘fat’ store of tuples</td>
<td>Shared ‘fat’ stream of tuples</td>
</tr>
</tbody>
</table>

Both Hadoop and SQLstream offer unlimited scalability across thousands of clusters. Where Hadoop is ideally suited to distributed batch processing, SQLstream complements this with real-time streaming data analytics, plus the real-time conditioning, integration and formatting of streaming data into and from Hadoop and Map/Reduce.

SQLstream also adds real-time geospatial analytics, offering insights into spatial patterns in the arriving data, as well as standard time-sensitive analysis.

Hadoop Ready
The SQLstream adapter for Hadoop integrates SQLstream into your Big Data storage environment, enabling SQLstream to stream data directly into and from Hadoop. SQLstream enables real-time streaming data management to be deployed painlessly as a core component of your Big Data strategy.
Relational Streaming for Big Data scalability

Relational Streaming is a paradigm for processing streaming Big Data tuples using standard SQL queries. SQL offers familiar and powerful relational expressions that lend themselves readily to automatic optimization and distributed parallel processing of streaming data. Relational Streaming queries execute continuously over arriving data. SQL is a declarative, mathematically tractable language that provides an ideal model for massively parallel execution. Streams can also be optimized according to operational requirements, such as the following:

- Computational Efficiency - minimum CPU consumption
- Latency - lowest end-to-end time for results
- Throughput - maximum number of records per second
- Reliability - maximum availability with lowest MTBF

Hadoop and SQLstream - Two sides of the same coin

Hadoop and SQLstream have similar characteristics and architectures for scalable, distributed processing of data. However, the diagram below illustrates that there are also significant differences. Hadoop excels at batch-based processing, SQLstream at streaming data processing. Where Hadoop executes historical queries on stored data, SQLstream executes continuous queries on arriving data, and pushes real-time analytics, data and results to downstream systems.

SQL is also declarative, where Hadoop and similar approaches are based on low-level procedural code. SQLstream extends Big Data storage platforms into the world of real-time, streaming data processing. Essentially SQLstream is the Big Data pipeline between real-time log file, sensor and clickstream data and your Hadoop environment, as well as simultaneously providing real-time analytics from the data as it streams past.

What next?

SQLstream offers the leading streaming data management platform as well as a range of Big Data and real-time consultancy services. Whatever your Big Data problem, we’re here to help, and would be happy to discuss your requirements.

Call us, or email inquiries@sqlstream.com.

About SQLstream

SQLstream is the leading standards-based streaming Big Data platform, forging real-time competitive advantage from streaming service, system and sensor data. SQLstream’s standards-based, distributed and scalable architecture uses industry standard SQL for the rapid analysis of high volume, real-time data streams. Standards mean lower costs, proven performance and seamless integration. With SQLstream, our customers are turbo-charging their Big Data environments for real-time, and responding with confidence to business exceptions based on accurate, up to the second information. SQLstream is headquartered in San Francisco, California and is on the web at www.sqlstream.com.