

The Build vs. Buy Decision

Exploring a Third Alternative

This whitepaper explores the classic IT conundrum of build vs buy, specifically in terms of acquiring a real-time streaming analytics platform.

Introduction

Enterprises are realizing the power that resides in real-time information, and they want it. Now.

As a result, streaming analytics platforms are fast becoming a must-have technology for enterprises seeking to build real-time applications. But, where to begin? Open Source? Build it yourself? Or buy it? And if buying is the answer, then from whom?

The question of build vs buy is not a new one; in fact, it's a classic IT conundrum. Unmediated, the conversation can devolve into nothing more than an esoteric, philosophical debate. However, there are many reasons the debate continues, and we'd like to offer a fresh look at that as it applies to the decision around building vs buying a Real-Time Streaming Analytics (RTSA) platform.

The first fact that we must admit is this: There is something incredibly enticing about having complete control over the creation of your own analytics platform.

But then again, few things compare to having a team of experts do it for you.

And therein lies the foundation of the problem. Which is better?

This whitepaper will outline the trade-offs involved.

There are obviously attractive qualities to both options which is why it is such an ongoing debate. Many companies that plan to adopt a solution will consider the in-house option first because they like the idea that such a solution will be easier to adopt, (especially when Open Source solutions abound). The sense of freedom, control, and lack of dependency on an external provider may also be alluring. Will this option ultimately be less costly? Will it be more scalable? This paper considers the pros and cons of doing-it-yourself against the option of buying a pre-built solution. Some of the considerations include issues of time, cost, human resources and the organizations' goals and strategies. This paper also looks at the four categories of buying options to help you determine which road to take if that is the direction you decide to go.

The build vs. buy decision is an important part of the development process. Either choice can be valid depending on the situation. If you're in a build versus buy quandary, this paper tackles these questions as they apply to developing an RTSA platform and attempts to offer some insights for your consideration.

The Market for Streaming Analytics Platforms is Dynamic

The market for streaming analytics platforms is gaining momentum far beyond its origins in industrial operations and financial services. You don't need an outside consulting analyst to tell you that streaming analytics is red hot right now. Hadoop may have been what opened our eyes to what is possible with big data, but the exhilaration around real-time streaming is all about what is possible when dealing with "fast data." Fast data (or data in motion) is about compressing and accelerating time to insight and ultimately time to action when dealing with the many streaming data sources that are available to enterprises today.

Today, streaming analytics platforms are available for every industry from a variety of enterprise software companies, pure-play vendors, and open source platforms.

Categories

There are four major categories of sourcing options for organizations wanting to address real-time streaming use cases.

BIG ENTERPRISE SOFTWARE VENDORS

- **Pros:** Proven, industrial-strength solutions that include enterprise integration and tooling.
- **Cons:** Potential cost and complexity of implementation; not specialized in RTSA.
- **Summary:** Enterprise software platforms generally cost more than solutions based on open source; may require the purchasing of other prerequisite software from the vendor and may not be as easy to integrate with other software platforms. The plus side is they typically deliver feature richness and have a certain level of longevity in the marketplace.

OPEN SOURCE STREAMING PROJECTS

- **Pros:** Free to use and an increasing number of projects to choose from. The open source community is a potentially powerful source of innovation.
- **Cons:** Lacks enterprise development/management tools and relies on the community for support.
- **Summary:** Open source offerings are getting a lot of attention today, especially Apache Storm which is in use by well-known companies with high volumes of streaming data, such as Twitter, Spotify, and Groupon. It is, however, a very technical platform that lacks the higher order tools and streaming operators that are provided by highly-rated vendor platforms. This means you really must be prepared to "Do-It-Yourself."

EMERGING STREAMING ANALYTIX PROVIDERS

- **Pros:** Purpose-built for today's diverse streaming analytics apps. These companies often leverage open source and augment it with enterprise class tooling and support.
- **Cons:** Still early in the market cycle to predict the extent of adoption of solution.
- **Summary:** A handful of streaming analytics providers are out there who are driven by the availability of innovative open source technology.

CLOUD EXCLUSIVE STREAMING ANALYTICS SERVICES

- **Pros:** Pay-per-use model and integration with other cloud services.
- **Cons:** Provides only limited analytical operators and lacks an on-premise solution.
- **Summary:** There are some cloud exclusive streaming analytics services, such as those offered by Amazon and Google, but these platforms are designed exclusively for application developers, can be difficult to code and may lack many of the higher level tools that enterprise-class vendors provide.

Decision to Build

The following list outlines a few trade-offs involved with a “Do It Yourself” strategy.



DECISION TO BUILD

- Extended customization
- Ownership of product and differentiation
- Personnel risk
- Timelines and opportunity cost

Extended customization

In-house development means more control over the process. This gives you greater personalization and a platform built to your exact specifications.

Ownership of product and differentiation

Most organizations opting for ‘build’ are interested in the end-to-end ownership of the product without the need for vendor dependencies.

Personnel risk

Requires a development team that is familiar with the rigors of product development and the need to ensure that the platform is scalable, extensible and supportable. If you do have such a team, there is the risk of over-dependency on one or a few individuals who designed and developed the solution.

Timelines and opportunity cost

Application development cannot start until the development platform is ready. This could mean delays in delivering the application's business benefits. Time spent building the platform could have gone directly into building the application — potentially delivering improved time-to-value and higher ROI.

The bottom line

Ultimately the decision here may be a function of to what extent the company wants to be in the platform development and maintenance business as a strategy of being totally self-sufficient vs implementing a pre-built solution that may carry some vendor dependencies but allows them to get started in quickly building solutions for the business.

Decision to Buy

The following list outlines some of the tradeoffs when buying a vendor provided solution.



DECISION TO BUY

- Rapid integration and simplified approach
- Training and support services
- Inadequacies, cost management and inefficiencies

Rapid integration and a simplified approach

Commercial software platforms developed by industry-recognized providers come with an implied guarantee that the system has been tried and tested and is relatively error-free. The expectation is that commercial platforms are standardized and stable. Implementation should follow a repeatable, well defined methodology. Application development can begin right away since you are not waiting for the platform development to be completed.

Training and support services

Commercial software vendors frequently deliver a broad range of support and maintenance capabilities, including training, and professional services.

This offers reduced costs for managing and maintaining a skilled internal staff to enhance or repair the platform.

Inadequacies, cost management and inefficiencies

Product may include unneeded features or unused functionality.

The platform may possess more generic capabilities to meet the needs of a broad user base.

Out of pocket costs for vendor software services and support may exceed expectations.

The buyer could experience vendor lock-in and over-dependence on the vendor's road map for future enhancements.

The bottom line

A vendor provided solution potentially removes the risk involved with being in the product development business and presents the opportunity to focus on delivering business solutions sooner by leveraging a pre-built platform and vendor experts who have "been there before."

The tradeoffs may be some loss of self-sufficiency and the ability to tailor the platform to what might be perceived as unique requirements. Dependence on the vendor for enhancements and even "vendor lock-in" may also become inherent risks.

The Third Alternative: “The Best of Both Worlds”

Nested between the two options of buying a 100% vendor provided solution and a 100% open source solution is a third alternative. Solutions such as our own StreamAnalytix product provide a fully supported enterprise class solution that is based on open source technology.

By combining the strengths of each, we are able to offer the following features and benefits:

- Full enterprise class support and professional services
- Pre-built enterprise class functionality
- Visual developer Interface for ease of use
- Rapid Application Development productivity that delivers quick time to benefit
- A pre-built and pre-integrated development platform that offers you the ability to tailor applications to enterprise requirements, including pre-defined operators, workflow, alerting mechanisms, a real-time dashboard, and a powerful CEP engine
- A steady stream of cutting edge innovation from the worldwide open source community
- Reduced costs compared to proprietary enterprise solutions
- Vendor managed product roadmap of future enhancements



A STATE-OF-THE-ART PLATFORM

In our next section we introduce a, state-of-the-art RTSA platform that is quickly finding early user acceptance and critical appreciation from industry experts and analysts.

Among modern RTSA platforms, StreamAnalytix stands out as a robust and hyper-efficient data management tool capable of multiple processing tasks that delivers fast, effective and real-time results while also providing strong customization capacities.

The StreamAnalytix Solution



StreamAnalytix is a horizontal product for comprehensive data-ingestion across industry verticals. It is developed on an enterprise-grade scale with a best-of-breed open source stack including Apache Kafka and Apache Storm while also incorporating the popular Hadoop and NoSQL platforms into its structure.

StreamAnalytix (www.streamanalytix.com) is a state-of-the-art RTSA platform combining enterprise class with open source technology. Built on an engineering framework that articulates the essential requirements for modern businesses, StreamAnalytix delivers:

- Reliability
- Large-scale efficiency
- Robust intuitiveness
- A future-proof and context-efficient abstraction architecture

StreamAnalytix is a horizontal product for comprehensive data-ingestion across industry verticals. It is an enterprise-grade solution developed on a best-of-breed open source stack, including Apache Kafka and Apache Storm. It also includes the popular Hadoop and NoSQL platforms in its structure.

Additionally, it supports custom extensions and multiple streaming engines and delivers unmatched analytical streams.

StreamAnalytix is powered by Storm and use Spark as its underlying engine.

StreamAnalytix is a powerful RTSA tool designed to help organizations seamlessly manage data.

Architecture

The following is an overview of the product and its architecture:

- **Open source technology:** StreamAnalytix is built on Apache Storm (an open source distributed real-time computation system) and is therefore able to leverage the numerous upgrades, improvements and flow of innovation that are foundational to the global open source movement.
- **Versatility and comprehensiveness:** StreamAnalytix is a 'horizontal' product for comprehensive high-speed data-ingestion across industry verticals. The platform caters to a variety of use cases spanning fields such as:
 - E-commerce
 - Online advertising
 - Security
 - Fraud
 - Insurance claim validation
 - Sensor data analytics
 - Call center analytics
 - Log analytics
 - IoT
 - Machine generated m2m data
 - Credit-line-management.












Its IDE development environment offers a palette of applications based on customer requirements. Multiple components can be dragged and dropped into a smart dash-board to create a customized work-sphere.

- **Abstraction layer driving simplicity:** The platform's architecture incorporates an abstraction layer beneath the application definition interface. This innovative setup enables automatic selection of the ideal streaming engine while also allowing concurrent use of several engines.
- **Compatibility:** Built on Apache Storm, Kafka and Hadoop, the StreamAnalytix platform is seamlessly compatible with all Hadoop distributions and vendors. This enables easy ingestion, processing, analysis, storage and visualization of streaming data from any input data source, proactively boosting split-second decision making.
- **'Low latency' capability and flexible scalability:** The platform's ability to ingest high-speed streaming data with very low, sub-second latencies makes it ideal for use cases which warrant split-second response, such as flight-alerts or critical control of risk factors prevalent in complex manufacturing environments.

- **Intricate robust analytics:** StreamAnalytix offers a wide collection of built-in data-processing operators. These operators enable high-speed data ingestion and processing in terms of complex correlations, multiple aggregation functions, statistical models and window aggregates.
- **Detailed data visualization:** StreamAnalytix provides comprehensive support for 360-degree data visualization. This means the system delivers incoming data streams instantaneously in the form of appropriate charts and dashboards.

KEY FEATURES

StreamAnalytix offers several differentiating value propositions and beneficial outcomes. Here are some of the key areas of innovation and superior tech-deployment within the product:

	Visual Application Development and Monitoring
	Real-time Dashboards
	Multi-tenancy Support
	High-speed Data Ingestion
	Elastic Scaling
	Flexible Data Parsing
	Rule-based Alerts
	Pluggable Workflow Management
	Real-time Index and Search
	High Fault Tolerance and Data Integrity
	High Performance Optimization

Summary

The key to getting real-time applications that your business needs now is to take advantage of new innovation using an abstracted streaming architecture that One of the greatest opportunities of a prepackaged solution is the ability to begin creating your application today and gleaning the results of it immediately, instead of taking time, money and resources to build the platform first.

The right RTSA platform should be geared to deliver the following essential elements, representing a collaboration from both sides of the divide:

will deliver today and not lock you in tomorrow.

Build	Buy
<ul style="list-style-type: none">● Open source● 'Do-it-yourself'● Low cost● Future-proof	<ul style="list-style-type: none">● Enterprise class● Reliable● Manageable● Professional

An RTSA platform holds the potential of being a definitive game-changer, drastically altering the way organizations interact with customers, data, and the tasks required to run their businesses. Any single vendor technology may not be sufficient to undertake this task

We recommend that organizations select open source options to build an RTSA platform using big data technologies. The success of such a platform depends entirely on the tools that are used and the services that support it. Therefore, organizations must use discretion in selecting the most appropriate tools from the available options.

GO GATHR

Data to outcomes, 10x faster.

- ✓ No-code/ low-code for data at scale, at rest or in motion
- ✓ Built-in ML to augment, automate and accelerate every step
- ✓ Drag and drop UI, 300+ connectors, 100+ pre-built apps
- ✓ Collaborative workspaces for Data, ML, Ops & Business users
- ✓ Open, extensible, cloud-native and interoperable



[Machine Learning](#) [Data Integration](#) [DevOps](#) [FinOps](#) [Business Process Automation](#) [More...](#)

[Schedule a demo →](#)

[Free 14-day trial →](#)