Fast Data and Architecting the Digital Enterprise
Fast Data drivers, components, requirements, and results

BIG AND FAST DATA
Big data and Fast Data are the twin pillars of a leading-edge business strategy using applications and data. Although the twins perform different functions and address different challenges, both are key for companies wanting to transition to digital business.

DATA AT REST
Today companies are collecting information from as many internal and external sources as possible:

• Web, mobile, and social channels
• Interactions with customers, business partners, and prospects
• Embedded devices
• Operational and logistics data
• B2B and partner networks
• ERP and other front- and back-end systems

This data collection results in what is known as big data or “data at rest,” large, rapidly growing volumes that are stored, then analyzed to identify trends and patterns that can improve decision-making. Big data is generally stored as static data on large distributed file systems, such as the Hadoop File System.

1 Digital business refers to the ability to create new business models by blurring the digital and physical worlds. It relies on the convergence of people, business, and things to disrupt existing business models — even those born of the Internet and e-business eras. What makes digital business different from e-business is the integration of things, connected and intelligent, with people and business.
DATA IN MOTION
Fast Data is “data in motion,” data in the process of being collected or moved between applications as part of a transaction or business process flow.

Fast Data is real-time data not yet stored as big data. It offers an opportunity for immediate response based on insights derived from deep analytics of incoming data streams. Fast Data processing sits in front of the big data fire hose, sifting through the massive amounts of incoming information to identify actionable business opportunities or threats.

The Time Value of Data
Fast Data can be generated at incredible speeds and volumes, at sample rates of thousands to tens of thousands of times (or more) per second, providing click-stream data, financial ticker data, or sensor and instrumentation data.

Fast Data is time-critical, existing within a small window of opportunity where it identifies and correlates with other data and kicks off an action or decision based on the events identified and the analytics applied.

Real-time insights derived from Fast Data allow you to:
• Proactively manage threats. Monitor suspicious activities to spot fraud patterns or operational data to find equipment issues before problems occur.
• Identify opportunities. Act on opportunities that exist only if you have a first-mover advantage. Or, identify new opportunities using predictive analytics and advanced data algorithms.
• Personalize customer engagement. Create a personal interaction on-the-fly using context from your real-time data stream (the customer is near your store) and data at rest (their buying preferences and purchase history).
• Make accurate predictions. Generate better-quality predictive analytics to affect future outcomes by combining the relevance and freshness of data in motion with deep analytics on data at rest.

Real-world Examples
Organizations in many industries are seeing tremendous benefits from Fast Data applications. According to Harvard Business Review, companies in the top third of their industry using data-driven decision-making are on average 5% more productive and 6% more profitable than their competitors.

A few real-world examples of the time value of Fast Data:
• Railways. Networks of sensors providing updates on train speed and rail traffic can be used to make adjustments to avoid unnecessary stops and save fuel, resulting in enormous savings and operational efficiency.
• Logistics. Real-time monitoring of traffic patterns, weather, construction delays, and other conditions along shipping routes can be used to reroute vehicles to optimize fuel consumption and delivery times and notify customers of changes or delays.
• Trading. Equity trading relies on Fast Data to evaluate potential real-time trades to prevent unbalanced portfolios and risk exposure.
• Retail. Brands can increase loyalty and purchases per customer with targeted coupons or offers awarded at checkout based on rapid correlation of the current purchase, purchasing history, and available inventory.

Oil & Gas. Drilling operators can detect drill bit damage before it occurs by analyzing historical data to identify failure patterns, and using business rules to detect these patterns in real-time operations data.

Banking. Banks are devising great customer experiences with the ability for timely and relevant conversations with customers across channels. Clerks, call-center staff, and bank officers can see what other interactions the customer has had recently, and continue the conversation.

Data Explosion, Analytics Advantages
Today, there are roughly 12 billion web connected devices globally, a number expected to grow to 50 billion devices over the next five years.

In 2013, IDC\(^3\) estimated that 4.4 zettabytes of data existed in the world (4.4 trillion megabytes). The amount of global data has been doubling every two years and it is estimated that by 2020 there will be 44 zettabytes of data, with a significant amount of it being generated by the Internet of Things (IoT), much of which will flow through applications as Fast Data.

Other interesting findings by IDC:

- In 2013, only 22% of the data in the digital universe was a good candidate for analysis (for example, useful if it were tagged). Less than 5% of that was actually analyzed. By 2020, the useful percentage could grow to more than 35%, mostly because of the growth of data from embedded systems and sensors, which is easily annotated with metadata or tags.
- Of the 22% of data identified as candidates for analysis, only 5% of it was especially valuable. That percentage is expected to double by 2020, as enterprises take advantage of new analytics technologies and new data sources.
- In 2013, less than 20% of the data in the digital universe was used in the cloud; either stored, even temporarily, or processed in some way. By 2020, the cloud factor is expected to double to 40%.

What these statistics indicate is that, while data generation will continue at an explosive rate, it will become increasingly difficult to find the information that really matters. High performance analytics that sift through real-time streaming data will be an absolute necessity for companies seeking to gain a competitive edge.

The Evolving Nature of Applications
Application design is also rapidly evolving. For a long time, applications were constructed as vertically integrated stacks of software, fully self-contained and deployed as independent silos for solving specific business problems. In these environments, it was not uncommon to see similar business functionality implemented repeatedly by different teams.

With the adoption of SOA architectures, these vertical silos began to break down as more business functionality was developed as a set of common services, shared horizontally between applications (for example, new account creation, credit authorization services, e-commerce shopping carts).

This trend accelerated with the broad adoption of public APIs that allow access to data outside of the data center, and by the incorporation of microservice architectures, which make it even easier to share fine-grained business functionality between applications.

---

The boundary of a data center and the edge of the enterprise is less well-defined today and a lot more permeable. Basically, the vertically integrated application silos of yesterday have been turned on their side and are now becoming horizontally integrated ecosystems of internal and external software services, APIs, and their creators. This is a very different approach to building an application that requires a different way of thinking.

The Cult of the Customer
One more relatively recent change has been the rapid uptake of mobile devices by consumers, and the corresponding growth in mobile applications, creating a lot more touch points and ways to interact. In a mobile environment—and now in all environments—consumer expectations are high that responses will be instantaneous. Numerous studies have shown that a consumer will navigate away from a web page or application that does not meet those expectations, likely going to a competitor.

Fast Data can improve the customer experience you provide in a number of ways:

- **Fast Response.** A high speed in-memory architecture ensures rapid responses for mobile and web applications and low latency IoT networks.
- **Customized, Higher Value Interactions.** The ability to use Fast Data to identify and create targeted, timely, high value offers for customers can increase customer satisfaction and brand loyalty.
- **Closer Relationships Based on Knowledge.** Fast Data—data in motion, real time analytics, correlation of historical data—can identify more ways to interact with customers and enable more relevant interactions at the appropriate times.

ANATOMY OF A FAST DATA APPLICATION
Big and Fast Data, an omni-channel customer experience, evolving applications: Interacting with Fast Data is a fundamentally different process than interacting with big data at rest, requiring systems that are architected differently. With the correct combination of functionality, an architecture can be built that delivers extraordinary value from both data at rest (big) and data in motion (fast).

**Fast Data as the Foundation for Digital Business**

Fast Data components include integration of real-time streaming data, machine analytics (including event processing software), human intelligence (and analytics software), an in-memory architecture, Master Data Management applied to big data stores, and automated actions throughout the process.
INTEGRATION IS THE FOUNDATION
It all starts with integration. A Fast Data application is only as useful as the breadth of data sources it can connect to. Every business relies on a different set of data sources and technologies. Some organizations require everything to be under local control in on-premises data centers where oftentimes it is simply not cost-effective to modify mainframe or legacy systems. Other organizations adopt cloud strategies or use outsourced services extensively and have little interest in operating their own IT infrastructure.

In any case, with the rapidly growing use of public APIs to consume or publish data, many data sources are not within the direct control of the IT organization, regardless of the strategy chosen. In most cases, a single approach is insufficient to meet the needs of the business, and most companies end up with a mix that results in a highly distributed hybrid IT architecture. Your approach to integration has to accommodate a need for flexibility and change.

TIBCO’s integration platform, ActiveMatrix BusinessWorks™, allows you to connect to virtually any application or data source, whether it’s in your data center or the cloud. Its connectivity capabilities span data sources, internal and external public APIs, and packaged, mobile, custom, and cloud applications. Out-of-the-box connectors are available for most popular SaaS and packaged applications, and an SDK supports developing connectors for custom legacy applications.

Integration is the core enabling technology that provides access to all of your Fast Data sources and creates the real-time data stream that all of the other Fast Data components interact with.

IDENTIFICATION OF BUSINESS MOMENTS
The more data sources you connect to, the greater the volume of data you will need to sift through to find those events (patterns) that have business value. Finding high value data to act upon is a lot like finding the proverbial needle in a haystack. It requires a combination of powerful machine analytics that filter the incoming data stream looking for specific events, coupled with rich visual dashboards that help people develop insights from the data. When used together, these two analytics components let you rapidly discover high value data and create “business moments” you can act on.

Capturing the evidence needed to detect the occurrence of certain types of business moments as they happen requires real-time processing of large streams of fast-moving data gathered from multiple digital channels including cloud services, APIs, social networks, mobile apps, and “things.”

Using TIBCO’s event processing solutions, events you are interested in can be defined, which are then applied in real time to the incoming data streams. When an event is spotted, a predefined action can be automatically executed to react. For example:

• When a credit card transaction in New York comes in at 2:00 p.m., and a second transaction in London comes in 30 minutes later, it will be denied, flagged as suspicious, and routed to the fraud department for follow-up; there’s no way the same person could be in two locations 3,500 miles apart with only 30 minutes time difference.

• Monitoring inventory across multiple stores provides alerts when levels are exceeded, kicking off actions to rebalance, perhaps by shipping to stores needing inventory or by reducing prices/running promotions to sell the excess quickly.
HUMAN INTELLIGENCE
When you know what you’re looking for, event processing allows you to process very large data streams in real time, then take automated actions to act upon the data as the event is triggered. Actions can be very simple—forwarding data to an analytics dashboard for further analysis or issuing a push notification to a person or application—or more complex, such as invoking a long-running business process.

However, when you don’t know what you’re looking for, and because data changes continuously, you have to find the patterns and trends that indicate new value or opportunities. The way to uncover these patterns is to bring human intelligence into the loop, using analytics tools and dashboards that correlate events to help bring new insights. Products like TIBCO Spotfire analytics provide powerful dashboards and visualizations combining diverse data types from many different sources for analysis and predictive analytics.

The key to making all this work is getting the right data to the right person at the right time. Fast Data applications are fundamentally architected to do just that.

ADDING CONTEXT TO INCREASE VALUE
Two of the cornerstones of TIBCO’s Fast Data platform are Master Data Management (MDM) and Business Process Management (BPM). MDM allows you to normalize similar data elements across applications and data stores. BPM enables you to bring together different resources (people, processes, and systems) from across your organization to take actions on your data.

In most cases, the more relevant the context you can add to a piece of data, the more valuable it becomes. Greater context generates deeper insights, better business decisions, and more potential touchpoints with the customer. The fact that a member of your loyalty program was recently browsing your website and left a product sitting in her shopping cart becomes a business opportunity when the GPS in her phone shows she’s near your store. Now you can instantly send an SMS offer inviting her into the store to purchase the item at an attractive price.

Creating context requires combining data from multiple sources and applications. In the example above, the customer is probably represented differently in different databases; perhaps as Jane_Doe in one, Jane Doe in another, and Doe, Jane in yet another. Unless you can reconcile these differences, creating correlation and context is impossible. This is the problem that TIBCO MDM addresses.

Some of this context can be added automatically through event processing or discovered in streaming analytics or analytics dashboards. The complete context will come from combining real-time streaming data with the historical data in your Hadoop file system or other databases.

Once an opportunity is discovered, you need the ability for the right people to take action, which is where BPM comes into play. TIBCO ActiveMatrix BPM provides processes flexible enough to react to the right business events in real time. Rather than follow a prescriptive, predesigned path, your business processes are designed to change on-the-fly. Both people and systems can compose the best possible process for the situation and information at hand. This ability allows you to take instant actions on your insights and produce contextual and customer-centric processes to differentiate your business.
IN-MEMORY ARCHITECTURE TO INCREASE SPEED

Conventional technology and applications were never designed for real-time analysis of high-volume streaming data generated by sensors, web applications, mobile devices, market data providers, social media, and other sources. Hard-disk data access is simply too slow.

Fast Data applications are architected end-to-end using in-memory computing, an architecture where applications assume that all data required for processing is located in the main memory of computing environments. An in-memory architecture that incorporates in-memory databases can also act as a bridge between the real-time environments of streaming analytics and the slower disk-based data in your big-data stores. Queries and analytics can be continuously run against big-data sources, with the results cached to an in-memory database, providing the ability to perform high-speed correlation of historical and real-time data.

TIBCO incorporates in-memory computing across all major product components of the Fast Data platform—integration, event processing, analytics, BPM, and MDM—whether on-premises or in the cloud. No other vendor delivers this level of support for in-memory computing.

SUMMARY

Data coming from your applications, systems, processes, customers, partners—and now the Internet of Things—fuels your business. The Fast Data Platform integrates this data to make the right information available at the right time, powering services and APIs, as well as automating processes.

Once you have the right data at the right moment, you need to provide it to the right person. The Fast Data platform puts information in the right hands, at the point of decision-making. With complete understanding of a situation’s context, decision-makers have insight. Whatever the volume or velocity, data scientists, business users, and operational teams have access to relevant information for making informed decisions in the moment. The platform applies these decisions immediately, adapting on-the-fly or triggering a preprogrammed action.

Your competition might have much of the same data as you—but with the ability to make informed decisions and quickly take action while the outcome is still in play, you can change what happens next . . . while your competition is still crunching numbers. That’s Fast Data.