IoT and Big Data

Bosch Software Innovations and MongoDB

Table of Contents

Introduction

Bosch Software Innovations – the Bosch Group's software and systems house, and MongoDB – the fastest growing database company in the market, are collaborating to provide a platform for the development of highly scalable applications for the Internet of Things (IoT).

From advanced telematics solutions to smart power tools and appliances to micro-electromechanical sensors, The Bosch Group is at the forefront of the IoT, driving its IoT activities through the Bosch Software Innovations (Bosch SI) unit. MongoDB is the leading database for operational big data applications, leveraged by thousands of customers for highly scalable real-time applications.

This whitepaper outlines the joint IoT application platform and how customers can benefit from Big Data technology in realizing the Internet of Things vision.

Internet of Things — The Opportunity

Regardless of who you talk to, everybody agrees that the IoT is growing exponentially: IHS forecasts that the IoT market will grow from an installed base of 15.4 billion devices in 2015 to 30.7 billion devices in 2020 and 75.4 billion in 2025. GE predicts investment in the Industrial IoT is expected to top $60 trillion during the next 15 years.

The excitement for IoT is supported by:

- Ever shrinking sensor size and price, while offering dramatic gains in functionality and performance
- Ubiquitous connectivity (almost every Telco carrier is now offering dedicated network solutions for IoT)
- Rich IoT application platforms (such as the one discussed in this whitepaper)
- An increasing amount of IoT applications in many use-cases and industries.

Some of these IoT scenarios are very end-user oriented, e.g. driven by crowd funding and the Maker Movement. However, many IoT examples are more focused on industrial applications including Fleet Management, Telematics, Smart Metering and Smart Grids, TeleHealth, and so on. This whitepaper will focus on three of the most innovative industry domains to illustrate the transformational power of the IoT, and how it can leverage Big Data:

- Retail and Logistics
- Manufacturing
- Mobility.

Use Case 1: Retail and Logistics

Retail and logistics is one key area where IoT is expected to have a huge impact as an enabling technology. RFID (Radio Frequency Identification) has been used successfully in logistics to track containers, pallets and crates for some time now, primarily in closed loop systems and mostly with high-value goods. The massive investments in IoT technologies are promising to help reduce costs for RFID and similar technologies, eventually making the tracking of goods on an item-level a feasible business case. For retailers, this has many advantages, including inventory accuracy, reduction of administrative overhead, automated customer check-out processes and a reliable anti-theft system.

Other emerging technologies include beacons, used as indoor positioning systems, which can interact directly with modern smart phones, e.g. using Bluetooth Low Energy (BLE). A network of in-store beacons can identify the location of a customer in a store and send them push notifications. For example, a user might create a shopping list on their smart phone and share it with the store app. Upon entering the store, the store app will display a map to the customer, which highlights all the products on his shopping list. Every time the customer gets close to a position where a group of products from their shopping list is located, the app will notify them and make a recommendation for a particular brand. At the check-out point, the system could identify all the products in the shopping cart automatically via RFID, create and confirm an invoice, and use the smart phone to process the payment. The store's inventory system is automatically updated when the checkout process is complete.

An IoT application supporting this kind of scenario would greatly benefit from using IoT application middleware, as explained later in this paper. The use of a NoSQL data repository would be of great benefit for storing all kinds of structured, semi-structured and unstructured customer related data, including shopping history and movements through the store. Advanced data analytics algorithms could be used to analyze the customer's movements and past shopping decisions. This enables the IoT application to generate shopping recommendations that can be pushed to the customer's smart phone while in the store, or to notify them of special offers – for example if the system detects that the customer is returning to an area in proximity to the store.

Use Case 2: Manufacturing

"Industry 4.0", "Smart Factory" and "Industrial Internet"– these are some of the terms used to describe the social and technological revolution that promises to change the current industrial landscape. There are many examples discussed and explored in this area, from leveraging IoT supply chain optimization to the modularization of production lines with the help of intelligent products.

One interesting example that we explore here is related to the increasing use of hand-held tools in manufacturing, e.g. for the assembly of automobiles, airplanes, trains and ships. In recent years, these tools have become more powerful (e.g. torque) and are now equipped with long lasting batteries, enabling workers to use them without the limitations of power cables or a fixed connection to an air compressor. This greatly enhances flexibility, but also poses certain challenges from a manufacturing process point of view, which can be addressed by leveraging IoT capabilities.

One of the key IoT concepts is the development of intelligent, connected "edge" devices. One example for such an IoT device is the Bosch Rexroth Nexo, a powerful nut runner which is equipped with an on-board computer and wireless connectivity. The on-board computer supports many aspects of the tightening process, from configuration (e.g. which torque to use) to creating a protocol of the work completed (e.g. which torque was actually measured). In addition, the Nexo features a laser scanner for component identification.

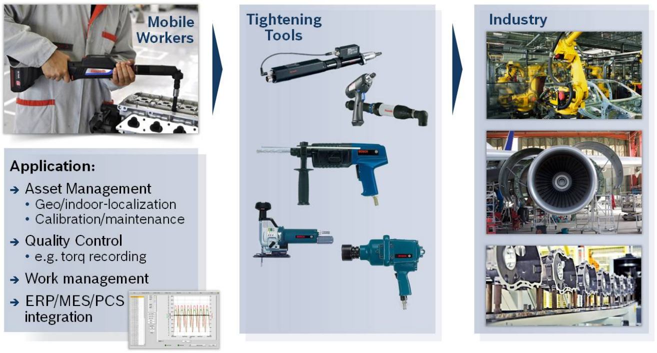

Figure 1: Manufacturing Use Case

By integrating such an intelligent edge device into the IoT, very powerful services can be developed that can help with supply chain optimization and modularizing the production line. For example, these intelligent tightening tools can now be managed by a central asset management application, which provides different services:

- Basic services could include features like helping to actually locate the equipment in a large production facility
- Geo-fencing concepts can be applied to help ensure only an approved tool with the right specification and configuration can be used on a specific product in a production cell.

The central asset management system can help with optimizing tool maintenance, for example by periodically reading calibration information from the remote tools via the factory WLAN. The asset management application can serve as the bridge between the power tools and the ERP (Enterprise Resource Planning) and MES (Manufacturing Execution System) systems that control the manufacturing process. For example, the asset management system can distribute work orders and configurations to the tools. In addition, the asset management application can document each tightening process by creating inspection lots (e.g. using torque recordings from the tools) and associate them with the BOM (Bill of Material) in the ERP system.

Such a production documentation system can benefit hugely from Big Data and NoSQL technologies that allow the aggregation of large volumes of heterogeneous, multi-structured data about the production process, including legacy data from many different systems, in addition to images and film recordings from different production modules. In an age where manufacturers can suffer huge costs from large product recalls, this can be a very powerful tool.

Use Case 3: Mobility

Rapid developments in mobility and automation are driving significant transformations across many industries – especially in the creation of new services and customer experiences. Telematics is a prime example of an industry harnessing the power of mobile connectivity and IoT.

While the engine data bus has long served to aggregate sensor events for engine diagnostics or geo-location, each new generation of vehicle is equipped with more sensors to extend services into fuel efficiency, autonomous operations, driver safety, theft prevention and more. The availability of these sensors, coupled with the integration of data to back-end enterprise systems via IoT application middleware is creating entirely new business models. For example, auto-makers and car rental companies have introduced new vehicle sharing offerings enabling customers to locate cars using their smartphones, rent them for a short time, and then park and return them anywhere within a defined zone. They may partner with local property owners to provide secure parking for the vehicles, and in the case of electric cars, with power companies for the location of charging points.

Consider leasing and car rental companies who are facing multiple challenges today. In a highly competitive market, margins can be improved only through cost savings. At the same time, customer requirements and expectations are increasing. Driven by rising fuel prices and general sustainability efforts, fleet managers expect additional solutions from car leasing companies to help them go green and reduce costs.

Another example is an increased interest in dynamic leasing contracts with flexible mileage and duration terms, offering better flexibility. In a project with a leading leasing provider, Bosch Software Innovations implemented a connected fleet solution addressing many of these challenges and therefore enabling the leasing provider to successfully compete in this market. Leveraging an on-board, built-in unit and remotely connecting this unit with a backend application allows the fleet operator to get real-time information about fleet performance, individual vehicle status, and so on. In the enterprise backoffice systems, this information is consolidated and fed into the relevant backend processes. Established approaches such as Business Process Management (BPM) and Business Rules Management (BRM) provide valuable tools and techniques to enable integration and automation – for example to schedule preventative maintenance and repair. Web-based access to vehicle information can be provided to the individual car lessees. Other mobility providers, such as gas station operators, are also integrated into the enterprise processes.

IoT Foundation: Asset and Event Management

As we could see from the three examples above, the IoT can enable a wide range of different applications. However, almost all of these IoT use cases share a common set of requirements. The goal of Bosch SI is to offer a generic application platform for the IoT which addresses these common requirements. Of course, flexible and scalable data management is one of the key requirements, and Bosch SI is leveraging MongoDB as the central data management technology.

Before looking at some of the benefits of this approach, the following section will provide a more general description of how the Bosch SI application platform is enabling IoT applications. This will help in understanding how these applications can benefit from the NoSQL and Big Data technologies provided by MongoDB.

If we look at the three use cases described above, we can see that there are two central entities in industrial IoT applications:

- Assets: Most industrial IoT applications are built around some kind of physical asset. The types of assets can vary greatly, from shopping carts to hand-held power tools to large trucks.
- Events: Managing events generated by the assets is critical to most industrial IoT applications. These events can range from technical events like a machine failure to a customer's expression of a buying interest in a retail application.

In order to efficiently manage different types of assets and events, the Bosch SI platform has a built-in mechanism for creating hierarchies of assets and the events that need to be communicated between these different hierarchies.

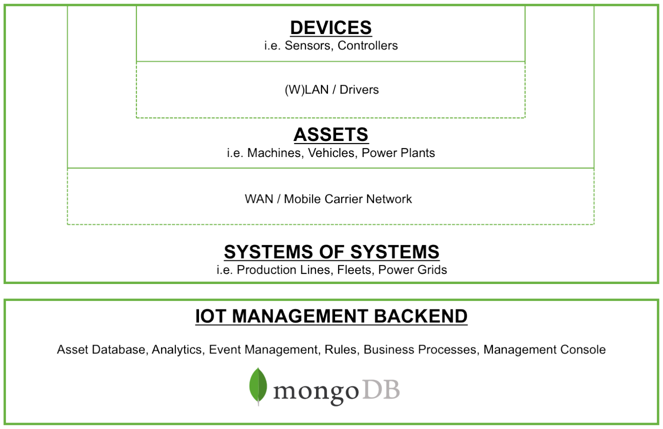
DEVICES
i.e. Sensors, Controllers
(W)LAN / Drivers
ASSETS
i.e. Machines, Vehicles, Power Plants
WAN / Mobile Carrier Network
SYSTEMS OF SYSTEMS
i.e. Production Lines, Fleets, Power Grids
IOT MANAGEMENT BACKEND
Asset Database, Analytics, Event Management, Rules, Business Processes, Management Console
mongoDB

Figure 2: Asset hierarchy in Bosch SI's IoT platform

The three hierarchy levels supported are:

- Devices: typically small, self-contained hardware elements like sensors, controllers, etc.
- Assets: typically a larger, more powerful entity, like a machine, a vehicle, or a power plant. An asset can connect to multiple devices through a local network, using specialized software drivers for different types of devices. The asset typically connects to the backend via a mobile carrier network (4G, GSM, EDGE) or via a direct Internet connection.
- Systems of Systems: Multiple assets can be grouped together to form what is called a System-of-Systems, e.g. an assembly line consisting of multiple machines, a power grid consisting of multiple power plants, etc.

The benefit of using well-defined hierarchies of assets in an IoT application is that the underlying application platform can provide many required features out-of-the-box, regardless of the specific type of asset. These features include:

- Asset Communication: The Bosch SI platform enables efficient communication between the different hierarchy levels, including events sent from remote assets to the backend, as well as operations initiated from the backend and sent to, and executed by, the remote assets.

- Management Backend: The Bosch SI management backend provides a central asset database, data analytics, secure event management, business rules execution, business process execution, and a management console for technical asset administrators.

From a data management point of view, the IoT application platform must manage two distinct types of data in the backend:

- Asset Instances and Hierarchy: Each asset becomes an individual entry in the central asset database, including information about its position in the asset hierarchy, as well as additional attributes and properties. The information about the asset hierarchy is critical in order to enable efficient communication with all assets.

- Events related to individual assets: Each asset entry contains a complete recording of all events related to the asset, including technical events (e.g. failure notifications, etc.), as well as business events (e.g. a vehicle is crossing a geo-fence).

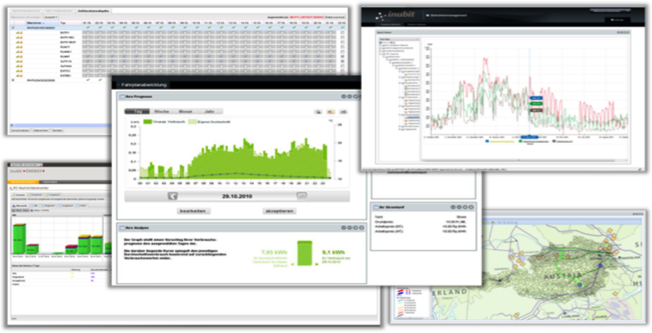

Figure 3: Examples for asset and event data from Bosch SI platform

Asset and Event data is imposing huge challenges on the data management strategies for IoT applications. Some of these challenges include:

- Data Volumes: Depending on time intervals and numbers of assets, data volumes for the event history can become overwhelmingly large. Imagine 10 million smart meters submitting a meter reading every 15 minutes for 20 years

- Data Schemas: In addition to structured data such as meter readings, many IoT applications also have to deal with semi-structured and unstructured data. For example, an event might include additional information such as a digital image taken to document the event, or an unstructured reading in an unknown technical format submitted by a specialized sensor.

- Data Schema Evolution: Especially in large IoT deployments it will be impossible to ensure that all assets in the field will always have the same configurations or versions of hardware and/or software. Consequently, the data management solution must be able to handle different versions of asset and event data in parallel.

To finish the discussion of asset and event management as the foundation for IoT applications, let's take a look at how an application must be able to support the analysis of individual events, as well as the correlation between different events.

Figure 4 provides an example. The asset management platform contains information about the different system entities, in this example a machine consisting of two components, A and B. The IoT platform is recording a time series for a given attribute of component B, e.g. engine temperature. If the temperature exceeds a certain threshold, a "temperature out of range" event is submitted. The operator would now shut down the machine. A more detailed analysis might actually reveal the following: Machine component A had a spike in the time series for a critical attribute of component A. This spike might have been above the norm, but within an acceptable range. However, time correlation of the two events reveals that this spike in the behavior of component A might be the root cause for the failure of component B.

This is a good example for the type of real-time data analysis for asset and event data in an IoT application which must be provided by the underlying data management platform.

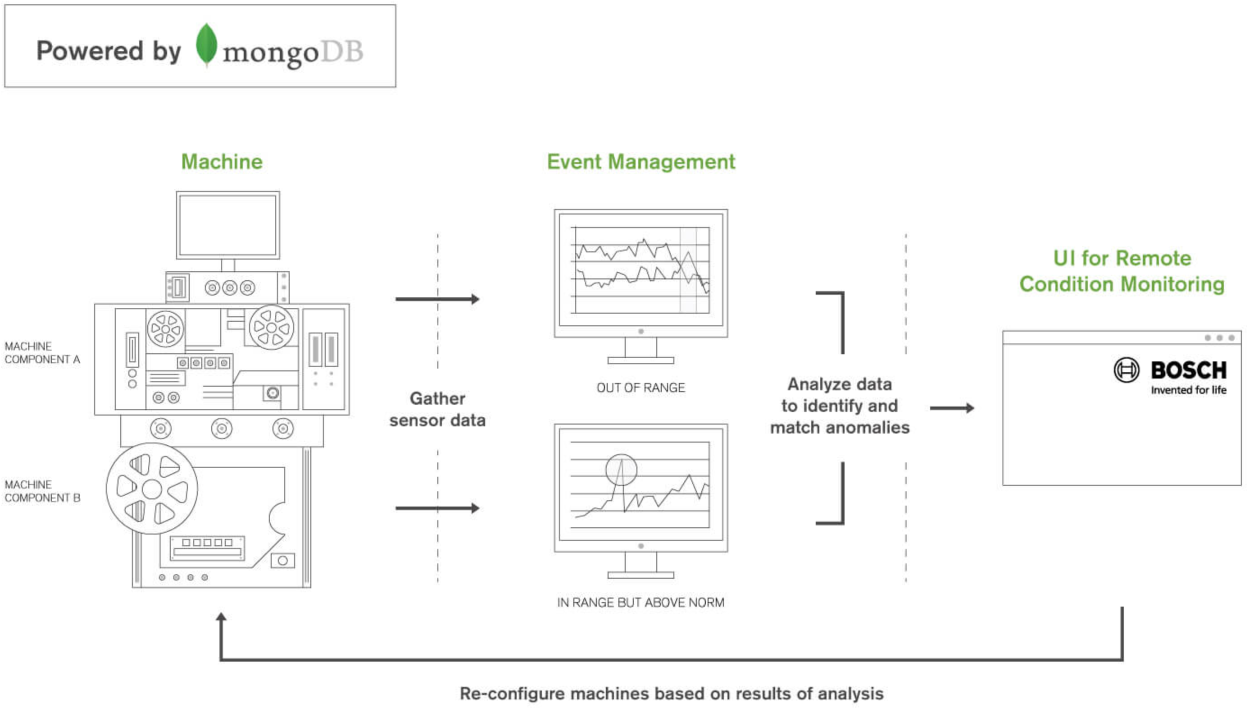

Figure 4: Event Correlation

Big Data: 5 Key Capabilities for IoT

Some of the basic asset and event management capabilities described above can already be found in established M2M application platforms. M2M – or Machine2Machine – is seen by many in the industry as the predecessor of the IoT. However, the evolution from M2M to the IoT is not only a question of adding more devices.

From a report published by Machina Research, an analyst firm specializing on M2M, IoT and Big Data: “The significance of the Internet of Things is not that more and more devices, people and systems are ‘connected’ with one another. It is that the data generated from these ‘things’ is shared, processed, analysed and acted upon through new and innovative applications, applying completely new analysis methods and within significantly altered timeframes. The Internet of Things will drive Big Data, providing more information, from many different sources, in real-time, and allow us to gain completely new perspectives on the environments around us.”

Machine to Machine (M2M) applications are well established in a range of industrial applications, they are typically characterized by managing well defined data sets from specific classes of device, encoded in vertically integrated technology stacks designed to enable monitoring and alerts.

The evolution from M2M to IoT fundamentally changes these characteristics. Data needs to be aggregated from multiple, disparate devices in addition to data from other physical assets and even from enterprise systems. The data arrives at higher speeds, in greater volumes and variability of structure. The data must be analyzed in real time to deliver richer applications enabling enhanced operational insight - examples of which are included in the following section of the whitepaper. The arrival of IoT places new demands on all components of the technology stack – and especially in the underlying databases used to store, manage, process and analyze the data.

To learn more about the challenges of IoT data storage, download the Machina Research report Why NoSQL databases are needed for the Internet of Things.

Based on our own research and experience, we have identified the following 5 key capabilities for data management in IoT:

1. **Creating rich, functional applications**: Data management must support the development of functionally rich applications with complex data and algorithms, with fast time to market and at low cost.
2. **Unlocking business agility**: The ability to support many new and frequently changing business requirements, causing fast and continuous evolution of the underlying data model.
3. **Enabling a single point of truth & Business Convergence**: Aggregate multiple views of related data from multiple systems into one consistent version of the data.
4. **Real-time operational insight**: Support both operational as well as analytical applications from the same data source.
5. **Enterprise-grade platform**: Provide highly scalable, cloud-based, robust and secure applications. In the following section, we will discuss each of these 5 key capabilities in context of our IoT business use cases.

1: Creating Rich, Functional Applications

Today's applications now incorporate a wide variety of data, bringing structured, semi-structured and unstructured data together to yield deeper operational insight into all areas of the business:

- Retail: "farm-to-fork" initiatives and increasing regulatory requirements to prove food lineage require the addition of sensors to generate audit trails tracking food production and transportation through the supply chain.
- Manufacturing: Capturing time-series, event based sensor data directly from the production line enables manufacturers to detect when processes are exceeding predefined tolerances and quickly take corrective action to avoid product wastage.
- Telematics and Mobility: The engine data bus is already established as a standard in luxury cars and trucks, consolidating individual events from sensors for engine diagnostics. The addition of new sensors enables richer applications and an increasing amount of data is being pushed back to central servers, enabling a fleet management company to start building new asset management applications that compare drivers and vehicles across their fleet, identifying best — and worst — practices.

The variability of data generated by IoT applications is a far cry from the simple general ledger and address book applications that helped to popularize the relational database. Trying to force-fit database technologies designed decades ago to support these new types of applications inhibits agility and drives up cost and complexity.

Semi-structured and unstructured data does not lend itself to be stored and processed in the rigid row and column format imposed by relational databases, and cannot be fully harnessed for analytics if stored in BLOBS or flat files. Organizations must therefore embrace database technologies that provide the flexibility to model, store, process and analyze new complex data types typical in sensor-based applications.

Flexible Data Model: JSON Documents

Complex data types typical in IoT applications can be modeled and represented more efficiently using JSON (JavaScript Object Notation) documents, rather than tables. MongoDB stores JSON documents in a binary representation called BSON (Binary JSON). BSON encoding extends the popular JSON representation to include additional data types.

With sub-documents and arrays, JSON documents also align with the data structure of objects at the application level. This makes it easy for developers to map the information model of the device or asset to its associated document in the database.

```
{
  vehicle_id: ‘123abc’
  vehicle_driver: ‘Miller’,
  base: ‘London’,
  tracking: [
    { timestamp: ‘2017-04-17-12:05:00’,
      location: [51.123,-0.232],
      speed: 55,….}
        { timestamp:
‘2017-04-17-12:10:00’,
      location: [[51.224,-0.238],
      speed: 5,….}]
}
```

Figure 5: MongoDB’s intuitive document data model

In contrast, trying to map the same object representation of the data to the tabular representation of an RDBMS slows down development. Adding Object Relational Mappers (ORMs) can create additional complexity by reducing the flexibility to evolve schemas and to optimize queries to meet new application requirements.

Instead of spending a lot of time dealing with the impedance mismatch between the programming language and the database, IoT developers must be enabled to focus on creating rich, functional applications.

2: Unlocking Business Agility

IoT is still in relatively early stages of development. Changes in customer requirements, emerging standards and new use-cases demand flexible and dynamic development methodologies and data storage architecture.

- Retail: Technologies such as NFC and beacons enable retailers to derive as much insight from customer movement around their physical stores as they are used to getting from tracking customer movement around their digital eCommerce stores. By capturing and visualizing data from these location-based sensors, retailers can build heat maps to optimize the placement of high-margin products.
- Manufacturing: Smart Factory concepts are proposing more flexible assembly lines and support for smaller batch sizes by moving away from centrally controlled systems towards chains of intelligent and more autonomous production modules which interact with each other directly via the products, e.g. using RFID to create a “product memory”.
- Telematics and Mobility: Rapid developments in automation and fleet management see each new generation of vehicle bristling with more sensors! Established telematics applications such as geo-location and engine diagnostics are being complemented by new services extending to areas such as fuel efficiency, autonomous control, driver safety, theft prevention, entertainment and more.

The rapid evolution of IoT applications can be constrained by traditional software development methodologies — for example, the waterfall approach places enormous dependency on the requirements defined upfront. In modern IoT applications, these requirements are simply not known at the start of many projects. Instead, organizations need flexible, iterative development practices to make it easy for teams to respond to new business and market demands, without being held back by rigid and static data models.

Continuous Integration: Dynamic Database Schema

MongoDB’s dynamic schema means that application development and ongoing evolution are straightforward, enabling continuous integration as developers add new features.

MongoDB enables developers to evolve the database schema through iterative and agile methodologies. Developers can start writing code and persist the objects without first pre-defining their structure. Each document (analogous to a row in a relational database) can have its own set of fields. Users can adapt the structure of a document’s schema just by adding new fields or deleting existing ones, making it very simple to handle the rapidly changing data generated by fast moving IoT applications.

Contrast this with a traditional relational database — the developer and DBA working on a new project must first start by specifying the database schema, before any code is written. At minimum this will take days; it often takes weeks or months.

The benefits of a flexible data model also extend to maintaining the application in production. When using a relational database, an application upgrade may require the DBA to add or modify fields in the database. These changes require planning across development, DBA and operations teams to synchronize application and database upgrades, agreeing on when to schedule the necessary ALTER TABLE operations. As MongoDB allows schemas to evolve dynamically, such an upgrade requires modifying just the application, with typically no action required for MongoDB.

Dynamic schemas bring great agility, but it is also important that controls can be implemented to maintain data quality. Unlike NoSQL databases that push enforcement of these controls back into application code, MongoDB provides schema validation within the database. Users can enforce checks on document structure, data types, data ranges and the presence of mandatory fields. As a result, DBAs can apply data governance standards, while developers maintain the benefits of a flexible document model.

High Productivity Application Development: Native Drivers

One fundamental difference with SQL-based relational databases is that the MongoDB interface is implemented as methods (or functions) within the API of a specific programming language, as opposed to a completely separate language like SQL, making MongoDB simple and natural to use. Developers interact with the database through native libraries that are integrated with their respective development environments, enabling much higher development productivity as they build new, or update existing, applications.

Creating Data Pipelines that React to Events in Real Time

MongoDB change streams enable developers to build reactive, real-time IoT services that can act on sensor data changes as they are ingested into the database. Change streams enable seamless data movement across distributed database and IOT application estates, making it simple to stream data changes and trigger actions wherever they are needed, using a fully reactive programming style.

Change streams can be used to create powerful IoT data pipelines that can react whenever the state of physical objects change. For example, generating alarms whenever a connected vehicle moves outside of a geo-fenced area, or requesting a service event whenever diagnostics sensors in a production line detect a failing component.

3: Enabling A Single Point of Truth and Business Convergence

Building a single digital view of a business entity — whether a physical asset or a customer — can deliver a range of benefits, from improved cross-sell and up-sell to enhanced operational efficiency and reduced costs:

- Retail: Building on the NFC and Beacon example earlier, retailers can instantiate a single view of their customers in real time, converging actual location in the store with their profile, purchase history and loyalty card details in order to deliver timely and targeted promotions.
- Manufacturing: Production line machines contain many discrete components, each with their own sensors. Bringing these together, along with the relevant service history can ensure optimum asset utilization and production line efficiency.
- Telematics / Mobility: Fleet managers can blend views of a vehicle's real time operational performance and diagnostics against asset registers that track service history to optimize preventative maintenance schedules.

Creating this "single point of truth" requires aggregating multiple views of related data distributed across different source systems into one consistent view. Using a relational database, the development and DBA team would first have to undertake lengthy design reviews in order to pre-define a common schema. Subsequent changes to the any of the source schemas would then necessitate associated changes to the single view schema.

MongoDB's dynamic schema and flexible document model do not impose the same constraints, enabling source systems to continuously evolve without impacting the single view needed by the business.

4: Real-Time Operational Insight

IoT applications enable new levels of operational insight and business discovery, but their value can only be fully realized when analysis is delivered in real time — providing the ability to react and respond as processes are in-flight.

- Retail: Inventory is tracked as it moves from shelf to basket while the retailer concurrently performs analytics that attempt to match available supply to predicted demand, adjusting for any deviation automatically through warehouse operations and the supply chain.
- Manufacturing: Sensor data from robotic systems is persisted to the database while analytics processes in the background to identify optimizations to the production line.
- Telematics / Mobility: Engine diagnostics is enhanced by writing a continuous stream of sensor data to the database while simultaneously performing analytics comparing current status to historical baseline readings in order to proactively identify deviations and potential faults. For example, changes in oil or engine temperature may indicate the need to perform preventative maintenance.

Many traditional databases support operational applications by capturing structured data as it is generated. They then rely on complex, slow moving batch ETL (Extract Transform Load) processes to replicate the data to the Enterprise Data Warehouse (EDW) where it is blended with semi-and-unstructured data for OLAP (OnLine Analytical Processing). To eliminate the analytics latency that inhibits real time business insight, it is necessary for the database to support both operational and analytical processes across the same data source handling structured, semi-structured and unstructured data.

Running Complex Analytics across Operational Data

MongoDB's replica sets allow live operational data to be written to a primary node before then being replicated transparently to secondary nodes against, which analytic processes can be run. This separation of concerns ensures that complex analytical queries do not take computational resource from the operational workload.

With its rich index and query support, including secondary, geospatial and text search indexes, and the Aggregation Pipeline, MongoDB can run complex ad-hoc or reporting analytics in-place against multi-structured data to deliver real-time operational insight. The MongoDB Connector for BI allows SQL-based business intelligence tools to natively query and visualise IoT data, blending it with data from internal relational systems. The MongoDB Connector for Apache Spark allows sophisticated machine learning processes to generate predictive analytics against MongoDB data.

5: Enterprise-Grade Platform

As IoT applications become embedded within the operational fabric of the business, they must deliver the scalability, availability and security demanded by any enterprise application. Business continuity and security are typically governed by strict mandates in every industry vertical. Specific examples include:

- Retail: With the average cost of security breaches reaching $4m in 2016, retailers are especially sensitive to the direct and reputational damage incurred from cyber attacks. As IoT applications are integrated into retail operations, they must provide security against multiple threat actors. Standards such as PCI-DSS and HIPAA (for those retailers selling pharmaceutical products) are also top of mind
- Manufacturing: If a production line were to stop due to an unplanned failure — even for a short period of time — the costs can be significant, including lost production capacity, idle workers and scrapped product. As IoT is at the very heart of many smart factory systems, continuous availability should be a prime concern.
- Telematics / Mobility: As drivers and fleet management companies come to rely on in-vehicle sensors for engine diagnostics, driver aids and regulatory compliance (i.e. self-certification of driver hours), downtime of IoT systems can cause critical business disruption and heavy costs.

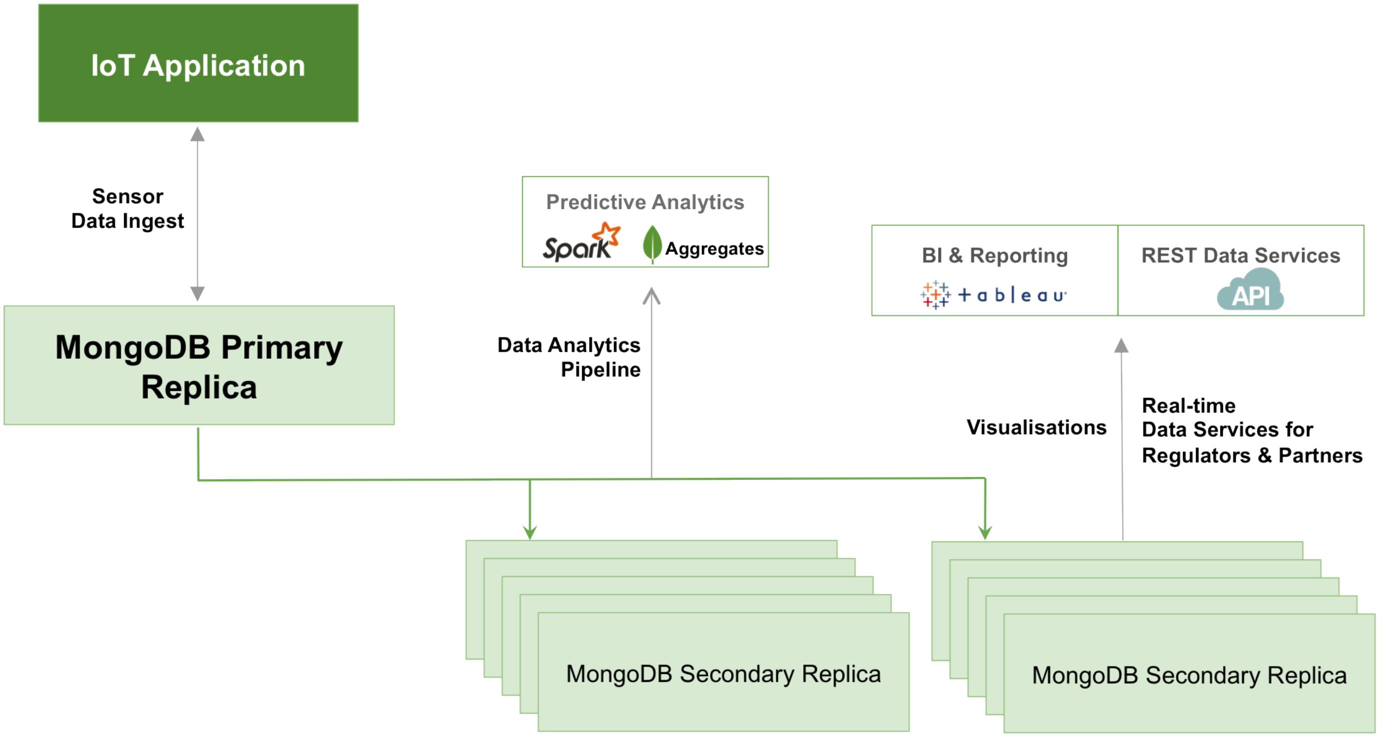

Figure 6: Single IoT platform converging real-time and analytic workloads against asset and event data

While databases such as MongoDB offer new capabilities for flexible data management and agile development methodologies, they cannot compromise on the enterprise-grade capabilities of traditional relational databases. Using a distributed database such as MongoDB organizations can build fault tolerant and secure applications that scale-out on commodity hardware as data volumes generated by sensors continues to explode.

Business Continuity with MongoDB Replica Sets

MongoDB maintains multiple copies of data within replica sets, using native replication to mirror data between replica set members. Replica sets can be deployed both within and across geographically distributed data centers. A replica set is self-healing, eliminating the need for administrators to manually intervene in the event of failure.

The number of members in a MongoDB replica set is configurable, and a larger number of replica members provides increased data durability and protection against database downtime (e.g., in case of multiple machine failures, rack failures, data center failures, or network partitions). Optionally, operations can be configured to write to multiple replica set members before returning to the application, thereby providing functionality that is similar to synchronous replication. Replica sets also provide operational flexibility by providing a way to upgrade hardware and software without requiring the database to go offline.

IOT applications cannot afford any loss of write availability when ingesting sensor data. MongoDB offers retryable writes, so in the event a network or node failure, the MongoDB driver automatically resubmits the write, and he MongoDB server enforces exactly-once processing semantics. Coupled with self-healing replica sets, MongoDB delivers always-on, global availability of write operations to ensure sensor data is always captured reliably by the database.

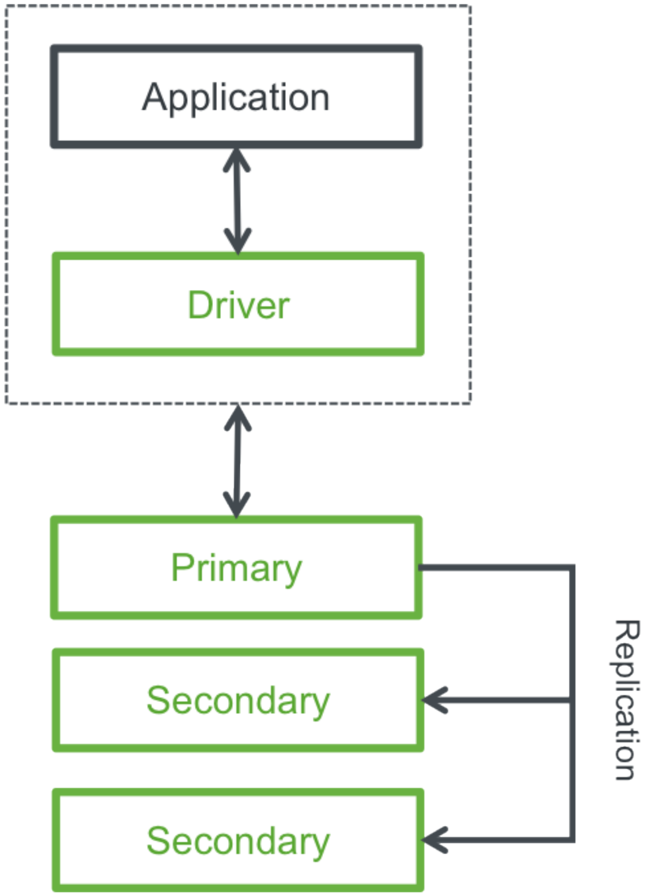

Figure 7: Self-healing MongoDB clusters

Defense in Depth Security Architecture

MongoDB is used to power business-critical applications in over 50% of the Fortune 100. Security is a top consideration for all of these customers.

MongoDB is architected to provide defense in depth at every layer of the database:

- Authentication can be managed from within the database itself or via integration with external security mechanisms including LDAP, Windows Active Directory, Kerberos and PKI x.509 certificates.
- Authorization can be configured with granular user-defined roles, making it possible to implement a separation of duties between different entities accessing and managing the database. Read-only views can be used to manage access to sensitive data at the level of an individual field. A single document can therefore contain data with multiple security levels, avoiding the complexity of separating data across distinct databases.
- Auditing captures access and administrative actions to the database, providing a trail for compliance and forensic analysis.
- Encryption, certified to FIPS 140-2, protects data in motion over the network and at rest on disk and in backups.

Economic Scaling as Sensor Data Grows

As new sensor-rich devices and assets come online, so data volumes can quickly exceed the throughput and capacity requirements of a single server. Relational databases were designed at a time when scaling meant buying larger servers, however most modern applications rely on scaling out by distributing the load across fleets of commodity servers - providing a much lower cost and less complex way of adding capacity.

To try and leverage the flexibility and economics of scale-out computing, developers and DBAs using relational databases spend valuable cycles adding caching layers or building custom database partitioning in their application code and queries — rather than writing business functionality.

MongoDB uses a technique called auto-sharding to automatically distribute data across clusters of commodity nodes in a way that is transparent to the application. This approach allows organizations to cost-effectively scale their IoT database without imposing additional complexity on development teams.

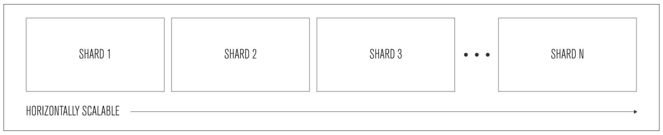

Figure 8: Simple scaling for IoT data growth

Conclusions

Platforms managing data and applications require significantly more flexibility, agility and scalability to meet the requirements of businesses in the Internet of Things.

These requirements will be dynamic and constantly changing as the opportunities from integrated sensors and devices, and their data are identified. Businesses will look to gain and maintain competitive advantage from innovative applications, requiring quicker application development, agile data models coupled with enterprise-grade assurances. Bosch SI and MongoDB have worked closely together to create a solution that helps customers addressing these challenges.

About Bosch Software Innovations

At Bosch Software Innovations we have been active in the Internet of Things for nearly ten years. We know what it can achieve, which goes far beyond technology, changing businesses, nurturing ecosystems and disrupting industries.

We are experts for the Internet of Things. Our team of IoT consultants, software developers, solution architects, project managers, UX designers, business model innovators, and trainers work with you to bring your IoT idea from strategy to implementation.

Bosch Software Innovations has designed, developed, and operated more than 150 international IoT projects. The Bosch IoT Suite is our comprehensive toolbox in the cloud provided as Platform as a Service. It allows the secure and efficient interaction of devices, users, and enterprise systems on a central, open software platform.

Our company vision is 'We connect everyThing'. Through our Bosch IoT Suite, we already connect more than 5.1 million sensors, devices, and machines with their users and enterprise systems today. From asparagus fields to welding robots, from smart home gateways to freight trains, from elevators to EV charging stations.

Learn more

About MongoDB

By offering the best of traditional databases as well as the flexibility, scale and performance today's applications require, we let innovators deploy apps as big as they can possibly dream. From startups to enterprises, for the modern and the mission-critical, MongoDB is the database for giant ideas. For more information, visit www.mongodb.com.

We Can Help

We are the MongoDB experts. Over 4,300 organizations rely on our commercial products, including startups and more than half of the Fortune 100. We offer software and services to make your life easier:

MongoDB Enterprise Advanced is the best way to run MongoDB in your data center. It's a finely-tuned package of advanced software, support, certifications, and other services designed for the way you do business.

MongoDB Atlas is a database as a service for MongoDB, letting you focus on apps instead of ops. With MongoDB Atlas, you only pay for what you use with a convenient hourly billing model. With the click of a button, you can scale up and down when you need to, with no downtime, full security, and high performance.

MongoDB Stitch is a backend as a service (BaaS), giving developers full access to MongoDB, declarative read/write controls, and integration with their choice of services.

MongoDB Cloud Manager is a cloud-based tool that helps you manage MongoDB on your own infrastructure. With automated provisioning, fine-grained monitoring, and continuous backups, you get a full management suite that reduces operational overhead, while maintaining full control over your databases.

MongoDB Professional helps you manage your deployment and keep it running smoothly. It includes support from MongoDB engineers, as well as access to MongoDB Cloud Manager.

Development Support helps you get up and running quickly. It gives you a complete package of software and services for the early stages of your project.

MongoDB Consulting packages get you to production faster, help you tune performance in production, help you scale, and free you up to focus on your next release.

MongoDB Training helps you become a MongoDB expert, from design to operating mission-critical systems at scale. Whether you're a developer, DBA, or architect, we can make you better at MongoDB.

Resources

For more information, please visit mongodb.com or contact us at sales@mongodb.com.

Case Studies (mongodb.com/customers)
Presentations (mongodb.com/presentations)
Free Online Training (university.mongodb.com)
Webinars and Events (mongodb.com/events)
Documentation (docs.mongodb.com)
MongoDB Enterprise Download (mongodb.com/download)
MongoDB Atlas database as a service for MongoDB (mongodb.com/cloud)
MongoDB Stitch backend as a service (mongodb.com/cloud/stitch)