



Highlights

- Data generated by the Internet of Things, Web and mobile-based systems of engagement should be managed and maintained in its native environment—the cloud—rather than moved to on-premises systems for analytics.
 - Retaining your data in the cloud allows for on-demand analytics, leveraging a simple integration between a JSON database service and a relational cloud data warehouse, with minimal data movement or transformation required.
 - By leaving database management to expert service providers, you can focus your attention on activities that provide a competitive advantage in the marketplace.
-

Analyze JSON data for 360-degree customer insights

Bring analytics to IoT, web and mobile data to boost your business

Overview

Today's Internet of Things (IoT), Web and mobile applications can collect more data about customers than ever before. This mass of information presents an opportunity for you to study your users, their needs and their habits with a level of granularity previously considered impossible.

However, the sheer volume, velocity and variety of this data makes it inherently difficult to analyze. How can semi-structured NoSQL data forms, like Java Script Object Notation (JSON) documents from web and mobile apps, be efficiently and safely stored—let alone generate meaningful insights that will boost your business?

There are a number of underlying factors contributing to this challenge:

- Systems of engagement create large stores of semi-structured data (from IoT, web and mobile sources) which complicate traditional data management. Planning, building and maintaining an infrastructure for that data consumes resources that are better spent engaging customers.
- Data generated by systems of engagement tends to be “a mile wide and an inch deep,” and stored in a simplified schema that lacks the robustness required to apply extensive models and predictive algorithms. The lack of compatibility between systems of engagement and systems of record—such as customer relationship management (CRM), master data management (MDM), inventory and other transactional systems—is an obstacle to analytics that would otherwise provide a “360-degree view” of the customer.
- Traditional, on-premises systems of record tend to grow and change at a predictable and stable rate, but data produced by systems of engagement is more fluid. As new technologies emerge and produce semi-structured data, storing that data and drawing useful insights from it becomes more difficult.



Fortunately, the recent explosion of composable cloud services offers some compelling ways to resolve the NoSQL analytics challenge, for three key reasons:

- Data generated by IoT, web and mobile-based systems of engagement should be managed and maintained in its native environment—the cloud—rather than moved to on-premises systems for analytics.
- Data generated by systems of engagement lacks the sophisticated schema of traditional systems of record, which makes it difficult to store this data for analytics using a relational data warehouse. However, it is ideal for storing in cloud-native solutions that leverage both SQL and NoSQL database systems.
- With the right platform, the data of the future—the data created by new, disruptive technologies—can be fully exploited in a cloud-based data warehouse.

Data silos: An obstacle to analyzing NoSQL

Data stores produced by web or mobile apps can be sizable and ever-growing. Keeping ahead of this data influx by planning, building and maintaining a scalable infrastructure requires considerable resources. For example, you may find yourself in the unfortunate position of having to bypass a new IoT technology or delay the launch of a web or mobile application because your servers are not yet available to support additional traffic.

Then, once the data has been collected, how do you draw meaningful insights from it? Much of the data from IoT, web and mobile sources is semi-structured. For example, JSON, stored in massive repositories with changing schemas, lacks the relational structure of data in most systems of record, such as CRM, MDM, inventory or financial systems. JSON data is therefore difficult to explore using business intelligence and analytics tools, which tend to rely on traditional schemas.

Further, the data collected from IoT, web and mobile apps presents a somewhat narrow window into customer data, based on their online interactions and behaviors. To draw deeper, more meaningful insights, customer data from systems of engagement needs to be compared and contrasted with data from systems of record to develop far more nuanced insights into future customer demand and behaviors.

The challenges to closing the gap between systems of engagement and systems of record are numerous. Differences exist in the following areas:

- **Cultural.** Data from systems of engagement belongs to web developers, while systems of record are the domain of traditional IT.
- **Practical.** It is difficult to merge semi-structured data with a flexible schema from a NoSQL database into highly structured data with rigid schemas in a SQL database.
- **Logistical.** Servers supporting systems of engagement could be located anywhere, but data from systems of record is often stored in a secure location behind a firewall, like in an on-premises data center or private cloud.

In the end, most businesses have silos of data, with a limited ability to combine and analyze information from their various systems of engagement and systems of record. But what if you could take the semi-structured data in JSON documents, tease out its structure, and easily export it into a cloud-based data warehouse? There, the data could be analyzed in its natural habitat—in the cloud—even in combination with data from your back-office systems. Bringing analytics to JSON data would be a big leap toward rounding out your knowledge of customers.

Featured service: IBM Cloudant



IBM® Cloudant® is a fully managed JSON document store that is optimized for handling many concurrent reads and writes, which is a workload that is typical for large, global mobile app usage.

Cloudant functions as a NoSQL Database-as-a-Service (DBaaS) for fast provisioning and worry-free data management. It is also available as Cloudant Local, which puts the power of the Cloudant platform in the privacy of your own data centers. You can even link Cloudant Local and Cloudant DBaaS together to form a hybrid cloud system for the greatest balance of cost, reach, performance and compliance control.

Flexible NoSQL data layer

Cloudant has a robust, RESTful API built on Apache CouchDB, which makes it easy to access using any language or Platform-as-a-Service (PaaS). It contains integrated APIs for advanced geospatial indexes and full-text search based on Apache Lucene. Cloudant provides unique multi-master replication and geoload balancing to spread data across global data centers and devices, enable extreme scalability and deliver data non-stop—whether users are online or offline.

With Cloudant, IBM experts manage the NoSQL data layer 24x7 so you don't have to, with higher-than-industry-standard service level offerings thanks to its support SLAs, high availability and uptime.

Elastic global scalability

Instead of a massive upfront capital expense for infrastructure, Cloudant offers pay-as-you-grow capacity so you can add and remove computing resources to match demand. In addition to elasticity for data volumes, Cloudant also has the flexibility to spread your data over multiple servers or multiple data centers to be closer to your users. It can even confine data to a specific geographic location to comply with a nation's unique data regulations.

Step one: Keep data in the cloud

An important first consideration when building an analytics solution for NoSQL is where to host the data. While some organizations remain committed to building server farms and hosting cloud-generated data in-house, that approach is proving to be costly, cumbersome and a drain on resources. A better approach is to keep cloud-generated data in the cloud.

Cloud computing is now a commodity, with many available cloud service providers, large and small. However, commodification does not excuse the need for due diligence and planning when choosing a cloud service provider. Decisions about who to partner with should align with the following requirements:

- **Service levels.** Establish the service levels you require, especially for uptime and network reliability, then compare and contrast them with the service level agreement (SLA) the cloud service provider will guarantee.
- **Scalability.** As you grow your business, your cloud service provider must be equally capable of growing with you. Confirm the availability of server space to meet your growth needs and the additional financial cost of scaling out as needed.
- **Security.** Verify which security measures the cloud service provider has in place to protect sensitive customer data. How does it manage encryption, replication and disaster recovery? Will those security measures meet your needs?
- **Support.** Some cloud service providers offer a turnkey solution, while others provide a multitiered support structure with experts on hand to respond to escalations and special needs. Check to make certain that any potential cloud service provider can offer the level of support you need.

One area of cost that can be frequently overlooked is in transmission and network fees. Many companies mistakenly look at data generated by IoT, web and mobile services as something to be stored temporarily on the cloud. They believe that to make that data useful requires exporting it into a local repository and absorbing the related costs to transfer that data. That belief, as you will see below, is changing.

Step two: Simplify analytics for JSON data

As mentioned previously, when semi-structured data is stored as JSON documents in a NoSQL database, it is effectively “resistant” to many sophisticated business intelligence tools. Previously, it required the use of multiple point solutions to export data from NoSQL databases, convert and scrub the data, and then store it in an on-premises data warehouse. However, a new hybrid solution, that integrates SQL and NoSQL capabilities, allows cloud-based data to be retained in the cloud for on-demand analytics, with minimal data movement or transformation required.

This solution leverages a simple integration between a JSON database service (Cloudant)—acting as the operational data store—and a relational cloud data warehouse (IBM Db2® products), where the semi-structured data is given structure

in order to be analyzed. This solution leverages Cloudant automated schema discovery process, which scans the JSON database and intuits the implicit structures in the data. It then creates a corresponding schema in the relational data warehouse, copying the data over in a format that business intelligence and analytics tools can easily understand (See Figure 1).

The benefits of such an approach are many, including:

- Easy visualization and reporting on the very latest data produced by IoT, web and mobile systems.
- On-the-fly data analysis of large, amorphous, fast-changing cloud-based data sets to identify trends and their related business opportunities in a timely manner.
- Eliminating data transmission and network costs.

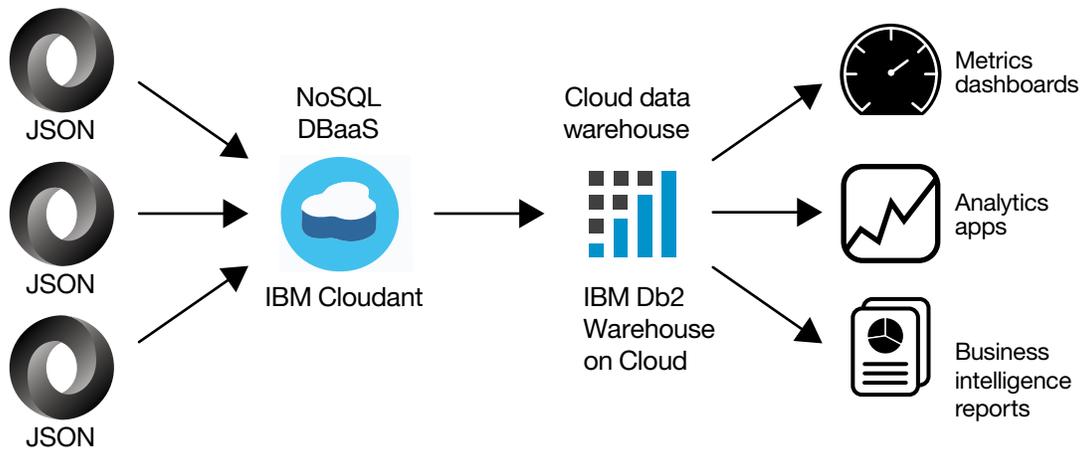


Figure 1: IBM Cloudant and IBM Db2 Warehouse on Cloud integration enables analytics for JSON document data.

An IoT use case: Real-time retail analytics

One of the fastest-growing IoT implementations is the use of Bluetooth low-energy (BLE) beacons in the retail industry. These beacons are unobtrusive, connected devices that can be installed at intervals within an environment and broadcast to receptive mobile devices — either one-way or two-way — with a data connection.

Retailers are finding BLE beacons to be highly effective for tracking customers' in-store behaviors. BLE beacons are an excellent example of new technology that is producing considerable amounts of data best stored on the cloud and demonstrate why that data should remain on the cloud to be analyzed.

For example, a beacon near a sale rack can connect with an app on a customer's mobile device, instantly rewarding the customer with an additional discount e-coupon for a sale item. This type of interaction can be stored as semi-structured JSON documents in a NoSQL data store like Cloudant. Already, the retailer gains some level of intelligence about the customer's in-store behavior as he or she moves within range of one beacon or moves to another, and the customer receives a highly tailored in-store experience.

But the real insight for the retailer comes when it combines semi-structured data generated from the beacons and mobile apps with structured data from its own systems of record. In this case, the retailer can pull profiles of customers from the master data management (MDM) database into a cloud data warehouse, like Db2 Warehouse on Cloud. Profiles show who the customer is, what clothes they have purchased in the past (including style, color and size) and whether this particular visit resulted in a sale.

Featured service: IBM Db2 Warehouse on Cloud

Db2 Warehouse on Cloud is a fully managed cloud data warehouse service that is:

- **Fast.** Can be deployed in minutes with the click of a button.
- **Scalable.** Pay only for what you need, with zero capital out-lay or planning.
- **Nimble.** Load and go with appliance-like simplicity based on IBM in-memory analytics technology.
- **Flexible.** Fully compatible with business intelligence and analytics tools from IBM including IBM Watson® Analytics™ and IBM Cognos®, as well as third- party solutions like Tableau and Looker.

Db2 Warehouse on Cloud is natively compatible with Cloudant, making it easy to bring Cloudant JSON documents into a relational data warehouse in the cloud. This allows you to create modeling and insightful analytics from web and mobile data, so you can focus on improving your business, not building out new infrastructure.

One of the most valuable features of Db2 Warehouse on Cloud is the ability, with the click of a button, to compare, contrast and query data from multiple sources. Not only is Db2 Warehouse on Cloud compatible with Cloudant-stored data in JSON documents, but also with data from your traditional on-premises systems of record.

The beacon and mobile data detailing a customer's most recent visit to the retail store, stored as JSON in Cloudant, can also be moved to Db2 Warehouse on Cloud using the Cloudant schema discovery process capability. Data from the systems of engagement and systems of record, previously held in silos, has now been effectively merged.

Now, the data can be combined and analyzed with business intelligence tools in useful ways to produce new holistic insights. Did the customer try on a garment? Did they buy it? If not, why not? Does the store carry the customer's size? The picture that is starting to take shape is that much-vaunted 360-degree view of the customer (See Figure 2).

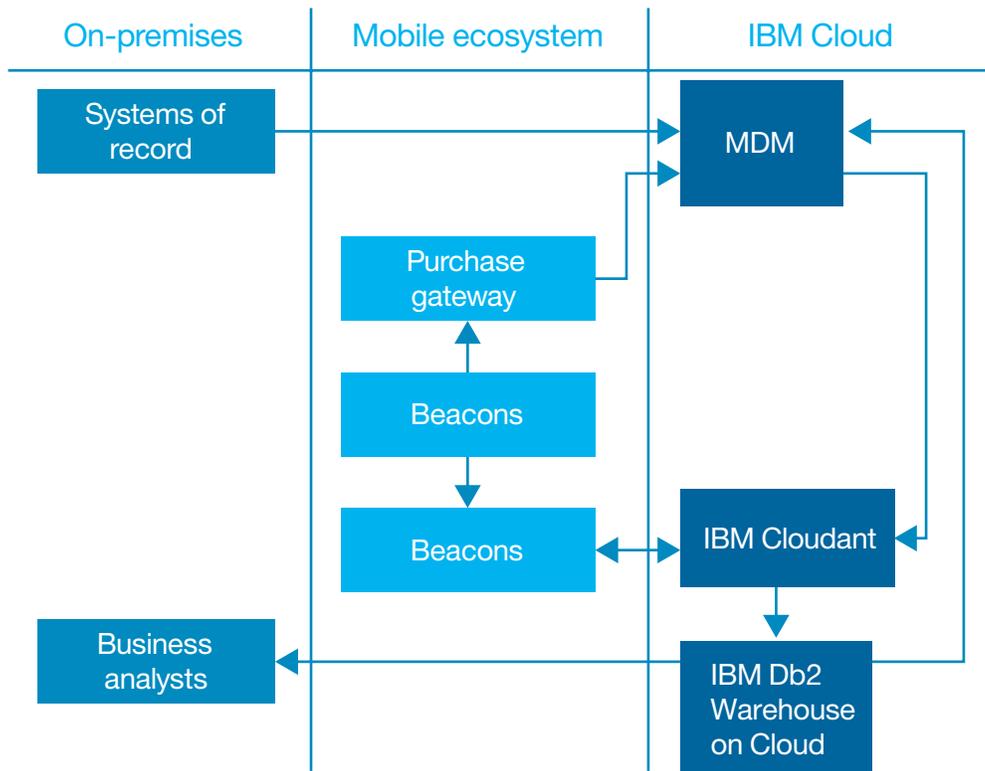


Figure 2: This is one example of an IoT analytics infrastructure using IBM cloud data services.

Do more with your application data

The growth of IoT, web and mobile applications—and the semi-structured data they produce—places a greater emphasis on the need for cloud-based data management. This data, stored in newer formats like JSON, is more efficiently managed by cloud service providers for whom it is their core competency. By leaving cloud storage to expert providers, you can refocus your attention on those activities that will provide a competitive advantage in the marketplace.

But, while traditional thinking has viewed cloud data stores as simply a holding place for web and mobile data until it is downloaded to a local system, to get the most out of data from IoT, web and mobile apps, you must keep that data online and bring the analytics to where the data already resides—in the cloud.

In this manner, semi-structured data—like JSON—can be efficiently incorporated into business processes, and converted far more quickly into actionable insights.

About IBM Cloud Data Services

IBM Cloud Data Services provides developers, data science professionals and analytics architects with a broad portfolio of composable, integrated data services for content, data and analytics. The open, self-service IBM Cloud Data Services offerings speed time to market, improve uptime and deliver higher value—backed by IBM technology leadership. For information about how IBM Cloud Data Services is changing the way services are created for and delivered to developers, follow us on Twitter at [@IBMAalytics](#) and [@IBMcloudant](#), and visit ibm.biz/clouddataservices.

For more information

To learn more about this offering, contact your IBM sales representative or visit the following web site: ibm.com/gbs/contact



© Copyright IBM Corporation 2017

IBM Analytics
Route 100
Somers, NY 10589

Produced in the United States of America
July 2017

IBM, the IBM logo, ibm.com, Cloudant, IBM Watson, Watson Analytics, Cognos and Db2 are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.



Please Recycle