05 Things Your Enterprise Needs To Manage IoT Data at Scale

DATASTAX®

IoT and 5G are creating a data explosion through the world's networks.

66%

Organizations plan to deploy 5G by 2020

Smartphones, smart cities, smart homes, and smart cars are reshaping our digital world and creating a massive infrastructure of sensory data that enterprises will be able to capitalize on in ways unimaginable even five years ago.

Offering low latency and higher bandwidth, 5G, the next generation of mobile data infrastructure, will significantly enhance applications benefiting from IoT data. According to Gartner, 66% of organizations plan to deploy 5G by 2020. Meanwhile, 59% say they will include IoT communications in the use cases for 5G.

The next generation of mobile data infrastructure, will significantly enhance applications benefiting from IoT data.

300MB

Per Second (5G)

Twenty years ago, we had 56K modems that brought the internet into our homes. 5G can support up to 300MB per second, over 5,000 times what we achieved at the end of the nineties.

We can only imagine what the next 20 years of technology breakthroughs will bring. Quantum computing and AI will change the rules around what can be achieved in computing and at what speed. Even the digital world is going through a revolution and companies needs to be ready for it.

Quantum computing and AI will change the rules around what can be achieved in computing and at what speed.

175

Zettabyes by 2025 (Global Datasphere)

IDC predicts that the global datasphere will grow from 33 zettabytes in 2018 to 175 zettabytes in 2025. Ninety zettabytes of this will be on IoT devices. IDC also predicts the average person will have nearly 5,000 digital interactions per day by 2025.

But it's not just the volume of data that should concern organizations hoping to reap the benefits of combining IoT and 5G technologies—it's also the velocity and figuring out how to handle that velocity while integrating the data with other internal and external data sources.

But no matter what organizations do with their data, in the end it comes down to the bottom line: business outcomes. Collecting and analyzing data at scale without achieving beneficial end results is pointless, and that's why organizations are going to need all the help they can get to take full advantage of IoT and 5G.

Following are the five key technologies that enable powerful management of IoT data at scale.

01

Time Series Data Modeling

Time series is very different from the more well-known relational data modeling approach. In a time series data model we want to store data in a columnar fashion rather than a traditional row-based model. This allows the database to efficiently write and read data to and from storage in order to speed up the time it takes to return a query.

01 Time Series Data Modeling

KEY BENEFIT

Time series databases are currently the fastest-growing segment of the database industry and are becoming increasingly accurate.

Time series data modeling allows for tracking, monitoring, and aggregating over time. It's specifically designed for handling events that are time-stamped and measuring changes over time. Key features of this type of database include compression, data lifecycle management, and data summarization. You can make requests for summaries of data over a certain period of time and perform computations of certain metrics over periods of time with accuracy and precision.

Time series databases are currently the fastest-growing segment of the database industry and are becoming increasingly accurate. The main benefits of using this type of database is for scalability and usability. Time series modeling is ideal any time you need to view data in the context of time, whether that's to evaluate trends or identify patterns in your data.

02

Real-Time Streaming

IoT systems typically deal with millions or billions of streaming points per day. Some of this data will require real-time action and executions. Building a streaming architecture enables real-time inspection of the data points that are of interest, allowing you to provide immediate feedback if needed.

Streaming data architecture is designed to perform real-time processing and also provide advanced analytics. It can process large volumes of data from not one, but multiple sources. Many organizations are adopting real-time stream processing because it is so effective for managing clickstream analytics.

For example, companies that want to personalize content based on what a web visitor views or clicks on in an ecommerce store can deliver relevant offers or product recommendations as the user navigates the site. Online retailers can also use this for cart abandonment activities. Any data coming from connected devices in IoT can be used in a similar fashion, personalizing the entire customer journey while providing valuable data for analysis.

Some of the key benefits of real-time streaming include:

- **Unlimited data streams** and capabilities to manage large volumes of data
- **Real-time or near real-time processing capabilities**
- **Easy identification** of patterns and other events for analysis
- **Hyper-scalability** and easy development without major infrastructure changes
- **Variety of tools** available for streaming data analytics

With all of the above at your disposal, you are making real-time streaming an incredibly powerful tool for managing IoT data at the enterprise level.

KEY BENEFIT

Unlimited data streams and capabilities to manage large volumes of data

03

Data Tiering

IoT systems have huge amounts of data, much more than any traditional database, so data management and data tiering are important to ensure a lower total cost of ownership. Time series data has a half-life of relevancy so tiered storage increases efficiency.

03 Data Tiering

KEY BENEFIT

Data tiering can help move data between different storage tiers without losing key data and without increasing costs.

Data tiering can help move data between different storage tiers without losing key data and without increasing costs. Some of the most common data tiers include flash storage, traditional SAN/NAS storage arrays, object storage, and public cloud.

Public cloud services for data tiering are becoming increasingly attractive for unstructured data, especially in the era of IoT. However, the public cloud is not without its limitations, including bandwidth costs, security risks, access patterns, and performance issues. This is why a hybrid cloud solution is becoming an ideal option for more balanced data tiering.

With hybrid cloud, it's easier to push all types of data to the public cloud while still having control over the data. It allows data to be moved easily between tiers and integrates with several destination cloud storage platforms.

04

Hybrid Cloud

As the use of IoT devices increases in the next few years, organizations will be increasingly dependent on cloud solutions that can effectively aggregate, organize, and store data efficiently, without compromising security and increasing costs. As more “things” start to deliver more data to computing devices, organizations need a way to handle the incoming workload and data processing efficiently.

Hybrid and multi-cloud solutions will be impacted by IoT as more organizations have new data requirements and higher volumes of data to manage.

04 Hybrid Cloud

Strategic objectives associated with a hybrid cloud solution would include:

- **Reducing the risk** of exposure to sensitive data sets in private clouds
- **Optimizing on-premises databases** and cloud-based solutions with distributed user bases
- **Effectively managing public clouds** for testing purposes and private clouds for production

Cloud-based and edge computing are the two advanced computing models organizations are turning to in the era of IoT. While cloud computing has become increasingly popular and widely adopted, there are still network latency issues to manage. This has led to the development of edge computing where processing of data happens at a point closer to where the data was generated—at the edge.

A good example of an edge use case would be self-driving cars. Since real-time response is critical for these vehicles to operate efficiently, edge computing allows for efficient and precise data processing. Edge computing also allows for IoT devices to communicate even under offline or low bandwidth conditions.

Hybrid cloud models combine the benefits of traditional cloud and edge computing. Hybrid cloud computing provides a faster and more secure experience with data processing happening closer to the source and also in a centralized repository when needed. This affords more flexibility when moving data.

KEY BENEFIT

Hybrid cloud computing provides a faster and more secure experience with data processing happening closer to the source...

05

Advanced Replication

When scaling databases, replication allows for the simultaneous transfer of data between two or more master sites. Data replication applications will be increasingly important for applying advanced analytics to IoT data. This will allow businesses to easily discover usage patterns, identify any weaknesses in connected devices, and, ultimately, help create better connected products.

05 Advanced Replication

KEY BENEFIT

More and more organizations will be using advanced replication in IoT environments to provide a consistent copy of data across all nodes, increase data availability, and improve reliability.

There are also advanced replication strategies like a hub-and-spoke model which allows data to be replicated from remote sites to a central hub. This allows the central hub to have a replica of all the remote sites and for each remote site to only have to look after its own data. This strategy is gaining more momentum with the evolution of edge and near-edge computing.

More and more organizations will be using advanced replication in IoT environments to provide a consistent copy of data across all nodes, increase data availability, and improve reliability. Advanced replication also supports multiple users while maintaining high performance.

Conclusion

Like today, data management will be the key to enterprises' success in the future. Data management has never been easy, but in the coming years it will get even more complicated, with sensors and devices around the world creating trillions of data points daily and each needing to be stored in a database.

To succeed, organizations will need to manage two issues:

01 How to create a data architecture that can manage the speed at which data will arrive from IoT devices, for availability and resiliency.

02 How to integrate that data with other internal and external data sources.

In both industrial and consumer settings, organizations will need to prepare their data architecture to cope with the scale of data from IoT devices created and distributed across 5G networks. Teams have to plan ahead around consolidating and analyzing this data so that they can make the most of their results, even as they spread it across multiple locations for availability and resiliency.

Be ready for the future, and for IoT, with DataStax

CONTACT DATASTAX

DOWNLOAD PRODUCT BRIEF

DataStax has helped some of the world's largest companies with their IoT and device-based platforms, paving the way for the data management of the future.

DataStax Enterprise (DSE) was natively built to deploy modern applications in hybrid cloud and to consume time series and sensor-based information faster than any other database. Based on Apache Cassandra™, DSE provides a contextual, always-on, real-time data management platform which can grow to unlimited scale. The platform is complemented by DSE Search, which provides real-time indexing, DSE Analytics, providing streaming and batch processing, and DSE Graph, enabling users to derive powerful insights from graph data.

KOMATSU

kapsch

EQUINIX

NTT DATA

Objenious by Bouygues Telecom