DATASTAX

The power behind the moment.

DataStax Enterprise and Microsoft Azure

THE HYBRID CLOUD DATA LAYER BUILT FOR DISRUPTORS

Microsoft

CONTENTS

ENTERPRISE OPERATIONAL DATA MANAGEMENT REQUIREMENTS HAVE EVOLVED

Digital disruptors are giving customers the information they want, when, where, and how they want it, resulting in a fundamental shift in customer expectations and how enterprises deliver value.

Enterprises are in a race to set themselves apart from the competition and meet soaring customer demands, resulting in the rise of modern applications. Whether it's for purchasing clothes, groceries, or software, companies and customers alike expect these applications to perform at a certain standard all the time.

However, building and managing these applications isn't as easy as it looks. These applications collect data from a variety of sources, which requires them to be able to handle a rapidly increasing data volume and complexity. To add to the challenge, these applications need to be highly relevant to the user, always available, instantly responsive, and accessible anytime, anywhere. Also, many enterprises are now opting to go "hybrid cloud" for these applications because it provides more flexibility while still allowing them to retain control of their data.

Not surprisingly, as enterprises move towards adopting these modern applications and achieving data autonomy, the data layer and cloud infrastructure they run on can make a huge impact.

In this eBook, we'll take a look at why DataStax and Azure combine to make the ideal data layer for contextual, always on, real-time, distributed, and scalable applications.

Now, let's get started.

INTRODUCING DATASTAX ENTERPRISE

The data requirements of today's applications are challenging, and this makes it critical for enterprises to have the right data layer foundation.

DataStax Enterprise (DSE)—a unified layer of database, search, and analytics.is the only hybrid cloud data management layer that allows enterprises to achieve data autonomy and meet the contextual, always-on, real-time, distributed, and scalable requirements of today's most powerful applications.

Built on the best distribution of Apache Cassandra™, DSE offers the complete selection of data management capabilities:

- **Multi model -** Powerful multi-model layer with support for tabular, graph, key-value, and JSON/ document data models allows architects to write data to a single solution and access it using a variety of methods based on the needs of the application.
- **Search and indexing -** Quickly and easily find data using complex queries, including support for sub-string, fuzzy, and full-text search, enabling you to search data at cloud scale in real time.
- **Analytics -** Real-time streaming and batch operational analytics, including ad hoc reports, enables you to make better and faster decisions.
- **Advanced replication -** Single cluster with primary hub and multiple spokes allows configurable, distributed data replication from source clusters to destination clusters bi-directionally.
- **Tiered storage -** Allows you to automatically move data to different storage media based on defined criteria and helps reduce storage costs by relegating lesser-used or older data to less expensive storage devices.
- **Security -** Consistent security model across entire data layer, including support for authentication, authorization, data auditing, and end-to-end encryption to ensure your data is safe both in-flight and at-rest.

INTRODUCING MICROSOFT AZURE— THE CLOUD TRUSTED BY ENTERPRISES

Microsoft Azure is an open, flexible, enterprise-grade cloud computing platform providing an array of services, from Infrastructure as a Service (IaaS) to Platform as a Service (PaaS). 90% of the Fortune 500 companies use the Microsoft Cloud, which underscores the amount of trust customers have in running their business on Azure.

WHY DSE AND MICROSOFT AZURE COMBINE TO MAKE THE IDEAL HYBRID CLOUD DATA LAYER

What's unique about DSE on Azure and why is it the ideal data management layer for your applications?

Scale, context, and distribution—always-on and in real time

Only DSE can deliver applications that are contextual, always on, real time, distributed, and scalable:

- **Contextual -** With integrated search, analytics, and graph capabilities, DSE allows businesses to unlock the value contained in the data's relationships to deliver highly relevant customer experience.
- **Always on -** DSE, due to its masterless architecture, offers no single point of failure and ensures 100% uptime. This means your mission-critical applications will remain always on no matter what.
- **Real time -** With instantaneous responsiveness, including real-time search and analytics capabilities, DSE enables businesses to make faster and better decisions, and serve customers in real time.
- **Distributed -** DSE provides real-time active-everywhere replication, enabling businesses to provide consistent customer experience no matter where they are.
- **Scalable -** DSE offers linear scalability, allowing businesses to deliver predictable performance regardless of the data volume or the number of users.

Microsoft Azure—The only consistent hybrid cloud you can trust

- Azure is the only consistent hybrid cloud on the market that allows you to build and deploy applications wherever you want. Connect data and apps in the cloud and/or on-premise for maximum portability and value from your existing investments. Azure offers hybrid consistency in application development; management and security; identity management; and across the data platform.
- Azure has the most compliance certifications of any public cloud provider, including national, regional, and industry-specific certifications. Also, Microsoft is committed to the highest levels of trust, transparency, and regulatory compliance to ensure customer data is protected. More information and the full list of certifications can be found at the Microsoft Trust Center.
- Azure has over three dozen datacenter regions across various continents, giving customers the choice and flexibility to choose where their data resides.
- With Azure, customers can leverage private, direct network connectivity with ExpressRoute, allowing them to move data from cloud to on-premise securely. Because this private connection does not leverage the internet, the data stays securely on a customer-controlled network.
- Azure also has a variety of virtual machine (VM) sizes, which gives customers the flexibility to pick the amount of central processing unit (CPU) cores, memory, solid-state drive (SSD), and local disks they want to provision and the right infrastructure for their DSE nodes.

Data Autonomy

Data autonomy is the ability to retain control over your data and manage your data and apps on a public cloud like Azure or on the hybrid cloud, independent of the underlying infrastructure. It improves business agility, compliance, data sovereignty, and data governance.

DSE allows businesses to achieve data autonomy by providing a consistent data management layer built for on-premise and/or hybrid-cloud data infrastructures that can be either self-managed or fully managed by DataStax.

DATASTAX MANAGED CLOUD ON AZURE

A fully managed, white-glove service, DataStax Managed Cloud on Azure allows businesses to focus on business innovation instead of worrying about the database management operations. It allows businesses to completely offload operational tasks—including backup and recovery, upgrades, security patches, and performance optimization—to DataStax, alleviating operational headaches and eliminating the various risks associated with doing things yourself.

To learn more, click **here**.

DATASTAX ENTERPRISE AND AZURE REFERENCE ARCHITECTURE

DSE can be deployed to a single Azure region or to multiple regions based on the distributed nature of the workload. Having a multi-region deployment model allows for high availability, low latency, and having data available in closer proximity of the users.

DSE's built-in node replication capability allows the synchronization of multi-master writes originating from multiple data centers easily and presents a consistent view of the data to applications. Also, DSE's masterless architecture, which allows a client to read/write data to any node in the cluster, ensures that your applications keeps serving the end users without any disruption in the event a node or a region is down.

Below is a reference architecture of a typical DSE deployment leveraging multiple Azure regions and interconnecting the networks with secure Virtual Private Networks (VPN)s. ExpressRoute can also be used instead of VPNs, when combining Azure regions with on-premise or co-location data centers.

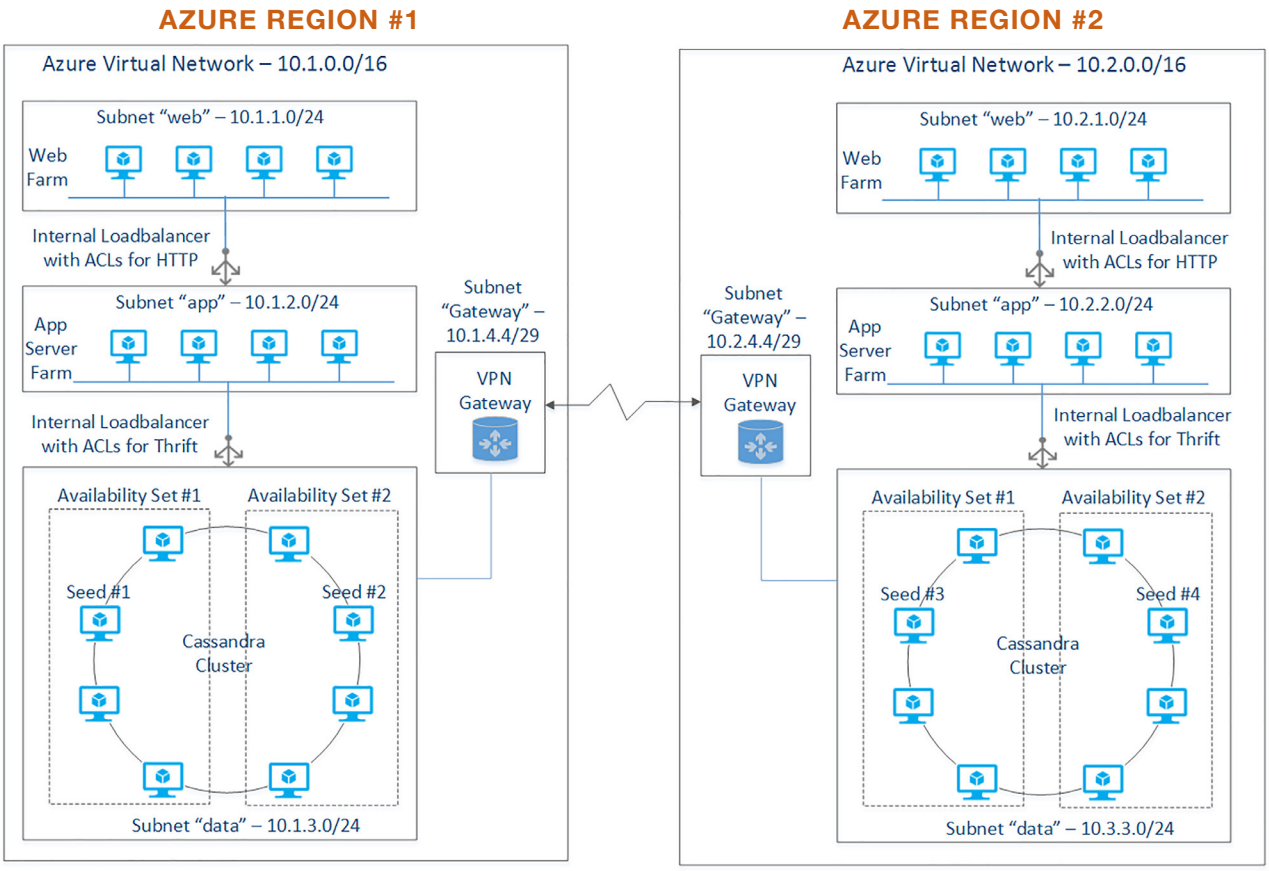

Figure 1: Reference architecture of a typical DSE deployment leveraging multiple Azure regions and interconnecting the networks with secure Virtual Private Networks (VPN)s.

POWERING USE CASES FOR MODERN APPLICATIONS

DSE on Azure is perfect for use cases that require contextual, always-on, real-time data—at cloud scale.

- **Customer 360 -** Get a complete and holistic view of your customers—their preferences, behaviors and interactions—to provide the best customer experience.
- **Personalization and Recommendations -** Provide tailored content and make relevant and timely offers to customers based on their preferences, interests, and online behavior.
- **Loyalty Programs -** Understand real-time and historic customer interactions and purchasing trends to provide instant rewards tailored to your customers' interests and habits.
- **Consumer Fraud Detection -** Understand user behavior to provide contextual warning for fraudulent activities.
- **eCommerce -** Use real-time access to multiple systems of record and always-on data to deliver a robust customer experience.
- **Inventory Management -** Use always-on data to allow shoppers and employees to make meaningful, real-time choices.
- **Asset Monitoring -** Use real-time and historical data, along with advanced analytics, to help customers make better operational decisions.
- **Supply Chain Management -** Understand your end-to-end supply chain network while being able to connect with multiple partners in real time.
- **Logistics -** Provide up-to-the-moment updates on product location and movements for even the most distributed and complex global logistics operations.
- **Security and Compliance -** Process massive amounts of data in real time to identify, stop, or prevent threats and meet compliance requirements.
- **Identity Management -** Guarantee user security via quick identity checks and real-time access controls.

CUSTOMER STORIES

Microsoft Office 365

Enhancing the Office 365 Experience for Millions of Customers

With more than 85 million active users relying on its services, the Microsoft Office 365 team wanted to ensure that its popular productivity platform provided the best possible user experience. Office 365 turned to the Customer Fabrics Platform Team at Microsoft to drive this project.

To deliver, the Customer Fabrics Platform team needed to gain a far better understanding of user behavior and needed to be able to detect any problems users faced before they noticed it themselves. Doing so required tapping into more than 300 terabytes of anonymized data from client and server logs.

The Customer Fabrics Platform team realized it needed a data platform that could scale quickly to handle this huge influx of information coming from every Office 365 server and client device without any downtime.

After evaluating various alternatives, the Customer Fabrics Platform team chose DSE for its always-on, linear scalability and real-time analytics. The pilot project, which started as a 40-node cluster, has now expanded to 400 nodes and handles about 400,000 write requests per second. To simplify implementation, the Customer Fabrics Platform team uses Azure Resource Manager templates to automatically deploy new clusters.

With DSE, the Customer Fabrics Platform team can now process and analyze both server- and client-side log data, allowing it to determine if a problem has originated internally or from tenant administrators and solve that problem before it becomes a crisis. This has reduced the number of support calls—resulting in major cost savings—while protecting tenant administrators from getting barraged with complaints.

Today, the Customer Fabrics Platform team runs multiple DSE clusters on Azure and has expanded its DataStax infrastructure to include systems that monitor Microsoft Exchange activity for performance and security.

Microsoft

> DSE on Azure enables us to do everything we needed, including near-real-time data processing and longer-term batch processes. We're also able to handle linear data growth without having to shard our clusters and manually manage data.
>
> **Sean Usher, Sr. Software Engineer, Customer Fabrics Platform, Microsoft**

IHS Markit

500 Million Components: IHS Markit Converts Data to Value

The UK-based IHS Markit provides critical information, analytics, and solutions to Fortune 500 enterprises across business lines, improving their operational efficiency and providing deep insights that lead to well-informed decisions.

IHS Markit manages a digital catalog of more than 500 million electronic parts and components. Customers can use this database to monitor component status; receive part changes and end-of-life notifications; quickly and accurately identify alternate component sources; optimize spend; meet regulatory compliance requirements; and more.

To improve the catalog's power, speed, and dependability in order to provide its customers with the best possible experience, IHS Markit wanted to unify the database into a single system called the Parts Content Factory. But the company soon realized that its relational database couldn't handle the long-term demand and growth requirements of its parts business cost-effectively.

IHS Markit chose DSE deployed on Azure for its massively linear scalability, masterless architecture, and CQL interface, with DSE winning out over competing technologies such as Apache Hadoop® with Apache HBase™.

With DSE Search, IHS Markit can easily optimize queries across a diverse set of complex attributes to generate fast and reliable results. Additionally, with DSE Analytics, the company can build dashboards for visualizations and real-time business analysis. DSE Analytics also allows IHS Markit to automate content classification and entry, an advancement that will literally cut days out of the intake process, getting data updates to market in less time—in many cases weeks ahead of what could be done prior to DSE.

Backed by DSE on Azure, the Parts Content Factory now helps customers find the parts and insights they need, while IHS Markit increases revenue opportunities and employees do more in less time and at lower cost.

IHS Markit™

> DataStax Enterprise and Microsoft Azure help us collect, validate, and distribute almost a full terabyte of content and convert it to value for our customers.
>
> **Graham Lammers, Director Product Development and Delivery, IHS Markit**

Komatsu

Creating the Construction Sites of the Future with Smart Construction

Komatsu (Komatsu Limited) is a global leader and manufacturer of construction, mining, and military equipment, based in Japan. The company also provides Smart Construction service—a solution for promoting the use of information and communications technology (ICT) across entire construction sites.

To deal with the growing shortage of labor on construction sites due to an ageing population and to improve the safety and productivity across construction sites, Komatsu provided Smart Construction service via its KOMTRAX System, a on-premise system comprised of 40 servers. However, as the number of users and construction equipment grew, Komatsu realized it needed a cloud data platform to store massive amounts of data being generated. Also, over 90% of the companies working at construction sites are small with fewer than 10 employees, so it is essential to keep the infrastructure costs as low as possible. At the same time, having cloud availability zones across the globe, compliance with Japanese regulations, and easy deployment and management were other key requirements.

After evaluating various alternatives, Komatsu adopted Microsoft Azure along with DSE as the IoT platform to support its Smart Construction Cloud Service. Azure was selected because it provided easy application development, complied with the Japanese regulations, and was the only cloud provider with two locations in Japan. DSE—an always-on data layer built on the best distribution of Apache Cassandra—was selected because of its linear scalability and enhanced operational management capabilities, including backup, restoring, and visual monitoring. The volume of data stored on DSE reaches several terabytes. There are approximately 60 virtual machines running at present, and the number can be easily increased and decreased as required.

Over 4,000 construction sites have already adopted Smart Construction Cloud Service, powered by Azure and DSE. Komatsu was able to build the Cloud Service in just two months and at a cost that is approximately half of that of an on-premise solution. The Cloud Service has allowed equipment operators even with little experience to perform construction work skillfully, enabling Komatsu to handle the shortage of construction site labor. Looking into the future, Komatsu aims to transform the foundation of Smart Construction into an open platform to create a sharing economy for construction sites, and the IoT platform powered by Azure will play a key role in that aspect as well.

KOMATSU

> Because Smart Construction Cloud Service requires handling a large amount of data, a distributed NoSQL database such as DataStax Enterprise is essential. DataStax Enterprise (DSE) was selected because it was often mentioned as a product with enhanced support functions for operational management, including backup and restoring, visual monitoring, and support for periodic maintenance work, not to mention the general support that is offered alongside it.
>
> **Mr. Hiroki Akanuma, Manager System Development Department Smart Construction Promotion Division, Komatsu Limited**

READY TO GET STARTED OR WANT TO LEARN MORE?

You can learn more more about DSE on Azure here.
And, if you are ready to get started, try DSE on Azure Marketplace.

About the Authors

KARTAVYA JAIN, SR. PRODUCT MARKETING MANAGER AT DATASTAX

Kartavya Jain is a Product Marketing Manager at DataStax. He has over six years of global experience in product marketing, digital marketing, and solution marketing at enterprise software companies. Prior to joining DataStax, he was the Head of Product Marketing team at DataTorrent and Sr. Product Marketing Manager at Hewlett Packard Enterprise.

RON ABELLERA, CLOUD ARCHITECT AT MICROSOFT

Ron is a Cloud Architect on Microsoft's Azure Global Black Belt team, working with customers from the CxO level to the technical architects helping them develop strategy, architecture, and a DevOps mindset to solve business challenges with the Cloud.

ABOUT DATASTAX

DataStax powers the Right-Now Enterprise with an always-on data layer for real-time applications at cloud scale. Built on the industry's best distribution of Apache Cassandra™, DataStax Enterprise makes it possible for companies to exceed expectations through consumer and enterprise applications that provide responsive and meaningful engagement to each customer wherever they go. Our product also gives businesses full data autonomy, allowing them to retain control and strategic ownership of their most valuable asset in a hybrid cloud world. DataStax helps more than 400 of the world's leading brands like Capital One, Cisco, Comcast, eBay, McDonald's, Microsoft, Safeway, Sony, UBS, and Walmart transform their businesses through right-now applications focused on enterprise optimization and customer experience. For more information, visit DataStax.com and follow us on @DataStax.

ABOUT MICROSOFT

Microsoft (Nasdaq "MSFT" @microsoft) is the leading platform and productivity company for the mobile-first, cloud-first world, and its mission is to empower every person and every organization on the planet to achieve more.

DATASTAX