IDC

ANALYST CONNECTION

Sponsored by: BMC

DataOps is a new method that has emerged in the modern organization to manage and control modern and legacy technologies in the data environment.

Improving Business Outcomes with DataOps Orchestration

April 2022

Questions posed by: BMC

Answers by: Stewart Bond, Research Director, Data Integration and Intelligence Software

Q. The world has changed a lot in the past few years, and modernization initiatives are driving new business models and outcomes. How does this relate to data in the modern enterprise?

A. The digital-first world is here, and organizations and individuals are always asking, "Is there some digital-based capability or enhancement that could improve our lives and desired outcomes?" To be competitive in the digital-first world, organizations need to build capabilities in enterprise intelligence, fueled by data. We are at a point in history when there is universal acceptance that data — its creation, management, analysis, and use — is a foundational asset for every organization. IDC research regularly shows that executives openly articulate the need for their organizations to be more data driven, in effect to be "data companies." With the desire to be more data driven, there is an expectation that nearly two-thirds (62%) of employees are using data to make decisions (source: IDC's *Data Culture Survey,* December 2020). But only one-third of people strongly believe their actions are driven by data. In a December 2021 IDC survey on the topic of data trust, only one-quarter of the people (27%) noted that they completely trust data. Clearly, gaps exist between expectations placed on the use of data and the ability to trust and use data in delivering data-driven business outcomes. Business modernization initiatives are also creating streams of new data (a digital exhaust) that can be used to trigger automation, drive engagement, and/or create new sources of business value.

Q. As the complexity of infrastructure, application, and data environments explodes, what are the operational challenges organizations are facing in becoming data driven?

A. Foundational to data modernization is data logistics, meaning how data is captured, stored, moved, protected, secured, and governed. Indeed, many data modernization efforts start with improved data logistics. Data orchestration is a key component, as orchestration determines the data flow from the time the data enters the system until the time it is leveraged in production for business results. Data availability and reliability are key to any function that intends to leverage

the value of data. As organizations seek to leverage data in this digital-first world, data environment modernization looks to use the latest in cloud, open source technologies as well as storage and organization solutions that enable more flexibility because fewer constraints are enforced against schema, format, and freedom of usage. But freedom without a framework is chaos. Data logistics is experiencing a renaissance. Many of the capabilities provided by legacy data management and automation solutions that provided control and governance are being refactored, reimagined, and modernized to accelerate work in the modern data environment and to help rein in the chaos. Organizations need to take a balanced approach with data logistics tooling so as to not stifle innovation or slow down progress in becoming data driven, or people will once again find their way around the controls that need to be in place to reduce data risk.

Q. What are the most common functions organizations deploy to deal with operational challenges as well as address gaps between expectations and reality in the use of data?

A. Commonly used functions in data management range from storage logistics to applications, enabling capabilities from data creation to consumption. In storage logistics, data protection functions ensure data is backed up and recoverable. Orchestration ensures data is captured, moved, and stored with integrity and then integrated into larger application workflows in production. Data availability and reliability functions ensure storage systems are reliably operational to serve data application analytics and processing. Applications deploy data engineering functions such as ingestion, transformation, cleansing, and federation to unify disparate data sources for analytics, visualization, and consumption. Metadata management systems capture intelligence about data so that applications understand data in context, including business and technical lineage, and to inform policy by providing security and privacy classification parameters. Policies are implemented via a controller to ensure the right data is being used for the right reason by authorized people and machines, governed to increase trust in the data and the outcome. This range of capabilities and functions introduces even more complexity into the operational environment, requiring a new approach to working with modern data. This new approach is known as DataOps, and it is helping organizations address the distribution, diversity, and dynamics of modern data environments.

Q. What is DataOps, and how is IDC seeing it addressing the challenges of modern data environments?

A. IDC defines DataOps as a combination of technologies and methods with a focus on quality for consistent and continuous delivery of data value. DataOps connects producers and consumers of data to extract value for any data-driven outcome, including analytical models, dashboards, algorithmic results, and operational or strategic business decisions. DataOps is a new method that has emerged in the modern organization to manage and control modern and legacy technologies in the data environment. The modern data environment is not just new technology. It includes new data technology alongside legacy data technologies inherited by organizational technical data debt. A December 2020 IDC DataOps survey found that 65% of organizations are using, at a minimum, 10 different data engineering and intelligence tools. According to the survey results, 54% are using one tool for each data engineering or intelligence capability, 33% are using two tools for at least one capability, and 13% are using three different tools for at least one capability.

DataOps is not DevOps for data. It borrows from the lean and agile principles of DevOps, including continuous development, testing, and deployment of applications, but adds statistical process control to continuously test data. Just as continuous testing of code in DevOps stops bugs from getting into production, continuous testing of data in DataOps stops bad data from being consumed and helps prevent analytical models from producing bad results because of bad data. A key tenet of the DataOps method includes orchestrating the highly distributed, diverse, and dynamic data management functions, from data logistics to orchestration, engineering, intelligence, and governance. This, too, is where a renaissance is being experienced, as enterprise-class operational orchestration is a gap in the market for solutions in the modern data environment. Organizations that have deployed open source DevOps orchestration solutions to manage DataOps pipelines are realizing limitations, and custom solutions aren't meeting the scale, performance, flexibility, reliability, and security requirements of enterprise workloads. Security, quality, remote work, and constantly changing business priorities are reported as the top challenges to DataOps implementation.

Q. What are the benefits of DataOps?

A. Principal drivers for adopting DataOps are focused on productivity and quality of data and data outcomes. Organizations that have implemented DataOps have seen a 44% reduction in the number of times data is causing exceptions in applications, a 38% reduction in the number of times applications are causing exceptions in data, and a 49% improvement in the ability to deliver data projects on time. These results are helping organizations build quality and trust back into the modern data environment. There are many positive correlations between improvement in data quality and trust and business outcomes, according to the previously mentioned IDC data trust survey. Nearly two-thirds (62%) of respondents indicated an improvement in business metrics: Customer satisfaction (42%), industry innovation (40%), and time to market (39%) were the metrics that experienced the greatest improvement. The opposite is also true, with 11% of respondents reporting that lower levels of data quality and trust have a negative impact on business metrics, with operational costs mentioned more than any other metric. Modern data environments are demanding, but so too is trying to remain competitive in this digital-first world. The ability to manage, control, and orchestrate data and data functions within the DataOps discipline will be a key success factor.

About the Analyst

Stewart Bond, *Research Director, Data Integration and Data Intelligence Service*

Stewart Bond is Research Director of IDC's Data Integration and Intelligence Software service. Mr. Bond's core research coverage includes watching emerging trends that are shaping and changing data movement, ingestion, transformation, mastering, cleansing, and consumption in the era of digital transformation. Having worked in the IT industry for over 25 years, from early experience in database and application development, through solution design and deployment, to strategic architectural consulting, Stewart has worked through some significant changes in the IT industry. His depth of field experience coupled with market insight gives him a unique perspective, valued by his customers and peers.

MESSAGE FROM THE SPONSOR

BMC's approach to DataOps

At BMC, we know that the modern business landscape presents teams and leaders with a growing list of challenges and problems to solve. Learning to harness your data is critical to evolving and staying competitive in an ever-shifting, disruptive world. BMC has an extensive history of empowering companies to become a Data-Driven Business. Our solutions are built to support the complex environments that span your entire ecosystem—from cloud to on-premises data centers to edge and everywhere in between.

IDC Custom Solutions

IDC Research, Inc.
140 Kendrick Street
Building B
Needham, MA 02494
T 508.872.8200
F 508.935.4015
Twitter @IDC
idc-insights-community.com
www.idc.com