IDC TECHNOLOGY SPOTLIGHT

Sponsored by: Fivetran

Enterprise-grade cloud data integration capabilities are required to match the qualities of service available in target systems to help these data integration pipelines avoid being the weakest link in the modern data stack.

Enterprise Security and Compliance for the Modern Data Stack

October 2021

Written by: Stewart Bond, Research Director, and Ryan O'Leary, Research Manager

Introduction

Every organization — big or small, private or public, old or new — is evaluating how data is being used today and exploring opportunities to leverage it to compete in the digital economy. Modern data environments have become highly distributed, diverse, and dynamic (the three Ds) as organizations migrate data and applications to the cloud and deploy digital capabilities in cloud-native architectures. This new "three D" reality includes not only the data but also the tools used in the modern data stack and the people who work with and consume data. Digital transformation is happening, but the compliance and security risks associated with data — from data creation to data consumption —in a digital-first world are increasing.

Figure 1 illustrates the frequency of security events that organizations are facing in the modern IT environment, including data. Fifty-seven percent of organizations are experiencing the loss or leakage of personally identifiable information annually, with nearly half of those organizations experiencing it monthly. Security events, including business email compromise, ransomware-infected endpoints, and malware-infected endpoints, are occurring at least quarterly for most organizations, which could also be contributing factors to the loss of personally identifiable information (PII) and other business-sensitive information. Cyberthreat actors are becoming more sophisticated and aggressive as the world shifts to digital-first enterprises.

AT A GLANCE

KEY STATS

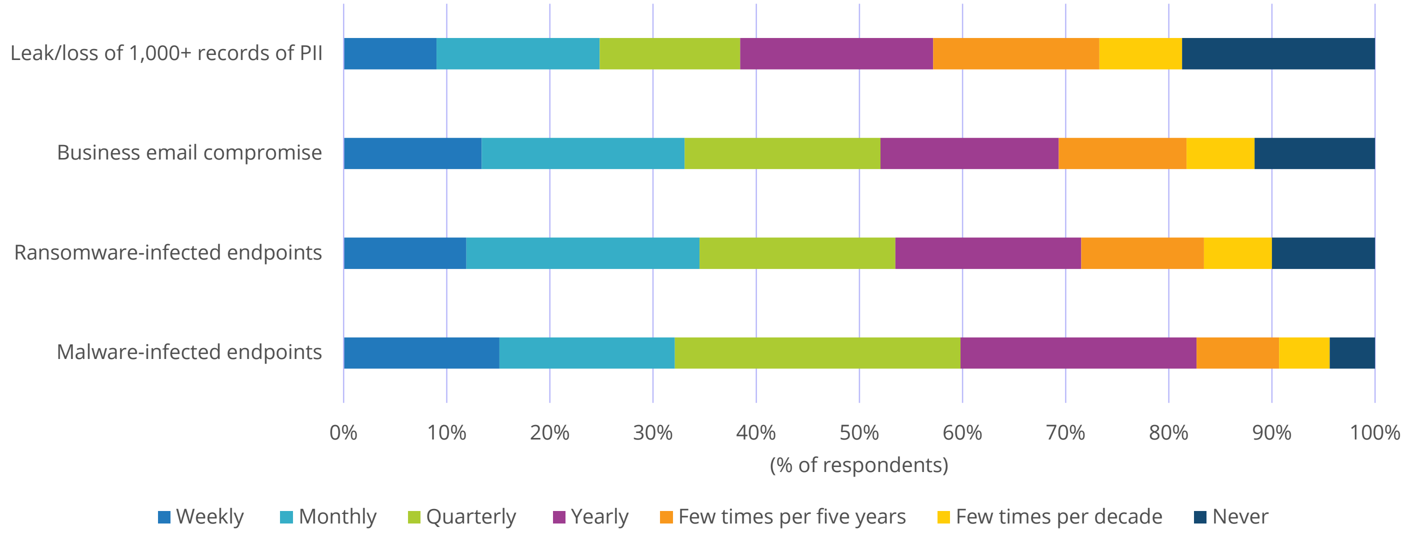
- 57% of organizations are experiencing the loss or leakage of personally identifiable information annually.
- Only 24% of organizations have indicated that security is leading to digital transformation success.

KEY TAKEAWAY

Security and compliance risks are increasing as data, data tools, and the people who work with data become more distributed, diverse, and dynamic in a digital-first world.

FIGURE 1: ***Frequency of Security Events***

Q ***Please provide your best estimate for the frequency with which the following events occur or are expected to occur in your entire organization.***



n = 411

Source: IDC's 2020 Data Protection and Privacy Survey, December 2020

These security events are symptoms of the underlying "three D" reality related to data, tooling, and people. The results of a 2021 IDC DataOps survey of 401 organizations illustrated that 95% of respondents have data pipelines running across hybrid cloud environments. These pipelines are handling up to 10 different types of data (e.g., relational, analytical, structured, and unstructured) coming from or going to up to nine different data management technologies, including mainframes, relational databases, streaming, NoSQL, and cloud-native data warehouses. Tooling used in the modern data stack includes custom and third-party tools to ingest, transform, federate, cleanse, master, define, and secure data. Nearly half (47%) of organizations that responded to the same survey are using more than one type of tool across a spectrum of custom code, commercial on-premises software, and software-as-a-service deployments to meet their data integration requirements.

The distribution and diversity of people who work with data are also increasing. Data engineers, data analysts, and database administrators are becoming more distributed throughout organizations, spreading skills thin and causing disconnects while increasing data security and compliance risks. An ever-growing number of organizational roles — including those in technical and business departments, from operations to strategy and from the back office to the front office — work with data daily to complete tasks, make decisions, and affect business outcomes. Data must be accessible from anywhere people are working, but access to data must be controlled to ensure the data is securely delivered to the right resource and for the right reason.

The pace of cloud migration has been accelerated as employees have become more distributed, digital experiences are expected by employees and consumers, and the seemingly infinite scale of compute and storage can be provisioned within moments instead of months, within flexible licensing models that enable cost optimization. Cloud is also offering higher qualities of service for operational and analytical databases in the modern data stack. In another recent IDC survey of 406 organizations that are actively migrating databases to the cloud, respondents cited high availability, security, scalability, flexibility, and agility as the top benefits achieved after cloud migration was completed. This list is likely a mirror image of why, up to a few years ago, organizations didn't move to the cloud. There remains a problem, however, in that solutions are only as strong as the weakest link.

To avoid being the weakest link in the modern data stack, cloud-native and cloud-resident data integration services also must provide enterprise-level qualities of service to meet security and compliance requirements.

Benefits of Enterprise-Grade Cloud Data Integration Services

While the COVID-19 pandemic continues to race across the globe, digital transformation has evolved into business resiliency. As enterprises attempt to address the security talent shortage and resulting demands on staff, many are searching for ways to safely leverage the cloud, including protection of sensitive enterprise data. Strict endpoint access controls, simple and secure access, and taking the guesswork out of deployments are some of the modern and core zero-trust attributes. Enterprises are tightening their fists and playing cards closer to the chest when it comes to data security and compliance. Zero trust extends to requiring private connections to and from cloud service providers to avoid unnecessary leaks caused by cloud provider misconfiguration. Cloud data integration services create that secure connection and protect data movement, severely mitigating the risk posed by data in transit.

50% of respondents to an IDC survey indicated that security and compliance need the most focus and development in the next six months.

Organizations taking a resilient posture strive for more control over data access. In the current zero-trust climate, enterprises are managing or bringing their own keys to data encryption processes. The stakes for security are high and tie directly into requirements laid out in most data privacy regulations, including the ability to shield access to data even from vendors that are transporting it. In IDC's 2021 *Future of Enterprise Resiliency and Spending Survey, Wave 3*, 50% of respondents indicated that security and compliance need the most focus and development in the next six months. Investment in cloud data integration services will only further this mandate and mitigate risk for the organization. Security and compliance go hand in hand.

Data privacy regulations have been popping up around the globe, with strong penalties for noncompliance. Multimillion-dollar fines have been levied in Europe for GDPR violations. While GDPR doesn't have specific data residency requirements, many other regulations do, such as Singapore's PDPA and China's Cybersecurity Law. GDPR requires that any jurisdiction in which the data is stored must have similar levels of data protection. The United States is not counted among similar jurisdictions, so it is critical that European Union (EU) citizen data not be stored in the United States. Cloud data integration service providers can enable the flexibility needed to ensure that data is properly geofenced and stored in the appropriate jurisdiction. This is a major benefit for multinationals and saves the cost and effort of setting up datacenters in every jurisdiction. Data sovereignty and GDPR-focused data practices have network benefits for the protection of PII.

Further, industry-specific regulations, such as HIPAA and PCI, require strong levels of protection. Cloud data integration services do not generally store any data, so the risk when it comes to HIPAA and disclosure of health information is low, especially with the encryption and key management capabilities that are features of many data integration solution providers. Enterprises can work with cloud vendors on the receiving end of the data integration process to ensure that encryption that is compliant with PCI DSS and HIPAA is carried through to the data destination.

Enterprise-grade cloud data integration services offer high levels of availability, including loss-tolerant pipelines to mitigate any data loss caused by inadvertent failures, including unavailability of source and target data stores. Higher levels of availability support the ability for always-on, high-frequency synchronization so that organizations can analyze what is happening today rather than wait until tomorrow. The endgame of highly available, highly secure, and compliant cloud data integration services is improved business outcomes, delivering faster time to insight while avoiding risk.

Trends Associated with the Modern Data Stack

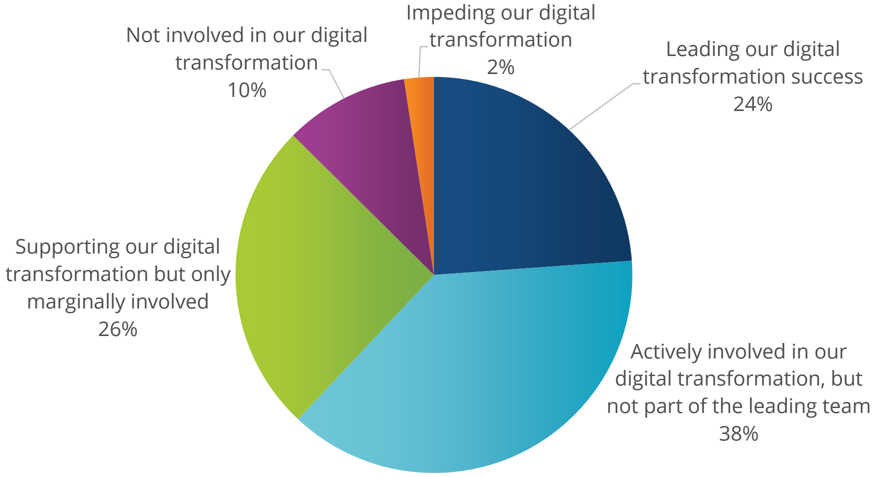
It is well understood that the global pandemic brought on by COVID-19 accelerated digital transformation and cloud migration in everything from infrastructure to applications and databases — both operational and analytical. In an IDC cloud data migration survey, 63% of respondents indicated that they were actively migrating data to the cloud, and only 10% indicated the pandemic had slowed down their migration efforts. The biggest challenge survey respondents reported was in selecting the most appropriate cloud database, data lake, or data warehouse service. This reflects the reality of many options being available, each with its own strengths and weaknesses for meeting different needs in the market. Enterprise-grade data integration services used for cloud data migration and/or steady-state cloud data integration need to provide openness and breadth of connectivity qualities of service to support multiple source and target alternatives.

The increasing number of threats and concerns when it comes to data security has in many ways been an accelerator of cloud data migration. Organizations are no longer interested in being security experts on top of their regular business. As vendors move from on-premises to infrastructure, platform, and software as a service, they become less responsible for the entire security posture and can shift the security risk to their vendors. In an on-premises world, the enterprise maintains the physical security, data security, network security, application security, and more. With "as a service" options, the security responsibility for one or more aspects of the data stack is shifted to the vendor. Further, many data privacy regulations require data security to be "state of the art" but don't specify what that means. Shifting risk to cloud suppliers and integrators has become more popular as a result of this ambiguity.

The digital-first world is upon us, and the gap between digital leaders and digital laggards is widening, separated along the dimensions of culture, strategy, financials, and platform. Data is the lifeblood of the digital transformation platform as it flows across lines of business, informing operational and strategic decisions that lead to action in the enterprise. Given the frequency of data breaches, business email compromises, and malware and ransomware attacks documented in Figure 1, security (and privacy) by design should be a focus of digital transformation initiatives; yet only 24% of organizations have identified security as leading their success, as illustrated in Figure 2.

FIGURE 2: ***Security in Digital Transformation Initiatives***

Q. *Please rate the extent to which security is involved in digital transformation.*



n = 341

Source: IDC's COVID-19 Impact on IT Spending, December 2020

DataOps has emerged as a method to help organizations with data modernization efforts alongside application modernization. DataOps is not DevOps for data; rather, it is a combination of technologies and methods with a focus on quality for consistent and continuous delivery of data value. DataOps is bridging the gap between data pipelines and business value with high-quality and high-fidelity data in near real time, accelerating the time to value of data, data analytics, data science, and machine learning. Data governance and security are key tenets of the DataOps method and accelerators for DataOps adoption, according to 74% of respondents to a 2021 IDC survey of 401 DataOps practitioners. A key differentiator between DevOps and DataOps is separation of continuous testing of the data (including testing data security and compliance controls) from continuous testing of the application. DataOps stops bad data from being consumed and security and compliance exceptions from occurring, just as DevOps stops bugs from being delivered in application code and implements control of application deployment.

Considering Fivetran for Enterprise-Grade Cloud Data Integration

The mission of Fivetran is to make access to data as simple and reliable as electricity. The company has been a pioneer in the cloud data ingestion and integration market, delivering data into cloud data warehouses such as Snowflake, Amazon Redshift, and Google BigQuery. Fivetran recently expanded to support additional cloud targets, including Azure Synapse, Amazon S3, and Databricks Delta Lake. The company provides cloud-native managed extract, load, and transform (ELT) capabilities in the cloud, with a large breadth of connectors to source systems, log-based change data capture (CDC), automated schema creation, column promotion, fast append, and propagation of source system changes into the target destination.

As a managed service, Fivetran offers a high level of availability (99.9%) as well as enterprise-level qualities of service focused on security and compliance. Data access is protected by TLS1.2+ and AES 256-bit encryption, 100% customer-controlled access, and single sign-on integration using SAML and role-based access control with full auditability.

Fivetran operations are SOC 2, Type 2 audited with ISO 27001 certification, and compliant with regulations such as GDPR, CCPA, and HIPAA.

The Fivetran Business Critical offering adds another layer of security and compliance controls to further protect data in motion during the ingestion and integration process. Services in the offering include Amazon Web Services (AWS) and Microsoft Azure PrivateLink, customer-managed keys, regional cloud support, and PCI DSS Level 1. AWS and Azure Private Link prevent data from being exposed to the public internet, eliminating opportunities for attacks against in-transit data being transmitted via the Fivetran solution.

Customer-managed keys enable users to flexibly create, own, and manage data access control, mitigating data breaches and limiting data exfiltration while enabling the ability to revert access if required. Regional cloud support allows customers to comply with data residency regulations. The added PCI DSS Level 1 certification available in the Business Critical offering enables customers to connect Fivetran systems within their PCI boundary.

The Fivetran Business Critical offering is also helping organizations comply with governance and privacy policies in DataOps methods by applying the rules, processes, and procedures required to maintain the security, quality, and deliverability of data from source to destination.

The mission of Fivetran to make data access as simple and reliable as electricity is being realized as the company puts more control into the hands of customers and enables enterprises not only to provide easier and secure access but also to harness the power of data in a digital-first world.

Challenges

Fivetran is one of the pioneers in data ingestion and integration in the modern data stack, and adding product features through its Business Critical offering is a positive step toward providing more comprehensive enterprise capabilities. Despite enterprise-grade capabilities being available in many components of the modern data stack, diligence is required on the part of the customer to determine where the weakest link is. From a market perspective, Fivetran faces increased competition from data integration and management software vendors that have been delivering enterprise-grade capabilities in legacy data stacks. Those vendors are now adapting and transforming their solutions and go-to-market activities for the modern data stack.

Conclusion

The reality of highly distributed, diverse, and dynamic data in a digital-first world increases security and compliance risks for organizations operating in hybrid and multicloud environments. Enterprise-grade cloud data integration capabilities are required to match the qualities of service available in source and target systems to help these data integration pipelines avoid being the weakest link in the modern data stack. IDC believes enterprise-grade improvements will continue to be made in cloud-native data integration solutions, increasing the competitive pressures for vendors in the market. To the extent that Fivetran can demonstrate value over and above its competitors, the company has a significant opportunity for success.

About the Analysts

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

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.

Ryan O'Leary, *Research Manager, Privacy and Legal Technology*

Ryan O'Leary is a Research Manager in IDC's Security and Trust research program covering privacy and legal technology. In this role, Mr. O'Leary leverages his legal experience to provide perspective on changes in laws, shifting regulations, and other market forces that affect technology decision making today for both law firms and corporations.

MESSAGE FROM THE SPONSOR

About Fivetran

Fivetran, the leader in automated data integration, delivers ready-to-use connectors that automatically adapt as schemas and APIs change, ensuring consistent, reliable access to data. Fivetran improves the accuracy of data-driven decisions by continuously synchronizing data from source applications to any destination, allowing analysts to work with the freshest possible data. To accelerate analytics, Fivetran automates in-warehouse transformations and programmatically manages ready-to-query schemas. Fivetran is headquartered in Oakland, California, with offices around the globe.

Fivetran's Business Critical plan builds upon its secure cloud data integration platform by providing enterprises with the highest level of protection for sensitive data. Fivetran Business Critical enables customers to create a highly secure modern data stack that meets internal and regulatory requirements.

Please request a demo to learn more about Fivetran and its Business Critical plan for secure data integration.

IDC Custom Solutions

The content in this paper was adapted from existing IDC research published on www.idc.com.

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

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.