Summary Insights: Data Protection at the Edge Is Essential for the Distributed Enterprise

RFG Perspective: Securing applications and data at the Edge requires an expanded, encompassing, and holistic security strategy, reaching from the Core to the Cloud to the Edge. Data protection and storage at the Edge, in addition to transmission of sensitive data to and from the Edge, is an evolving challenge for IT.

Transparency into the operations of compute, IoT, network, storage, and OT activities must take the appropriate privacy concerns into account – and it must be coupled with security and compliance enforcement capabilities.

Methodologies for protecting Edge devices must incorporate threat detection capabilities, application and perimeter security toolsets, vulnerability management processes and procedures, and patch management strategies.

Edge devices are now established as part of the IT landscape. That means that IT executives, who are already at work on updating their organization's infrastructure, must ensure that all Edge devices comply with data protection and corporate security standards.

On Nov. 24, 2021, the Robert Frances Group hosted a panel and discussion about the IT challenges and business challenges, entitled: "Data Protection at the Edge." This event examined best practices and recommendations for applications and data protection of sensitive data residing at Edge locations, and in-transit to and from Edge devices.

The panel for this RFG 100 event included:

- Lucy Shenton, cyber and fraud risk expert, McKinsey & Co.
- Deal Daly, now CIO of Toysmith, previously CIO/CISO for HammerSpace
- Kimberly Lancaster, privacy, GRC [governance, risk management and compliance] and security expert, Marqeta, Inc.

INTRODUCTION

Edge devices are playing an increasingly important role in the way that companies create, move, and connect with their applications and their data. Clearly, Edge devices and Edge systems must be protected, even as new and potential attack vectors become more widespread.

46 Kent Hills Lane, Wilton, CT. 06897; (203) 429 8951;
http://www.rfgonline.com/; *Contact: inquiry@rfgonline.com*

Edge device strategies should focus on computing, storage, and analytics wherever rapid, real-time processing generates operational data. The ability to protect the full spectrum of distributed data improves data-visibility for the entire organization – and enhances the organization's ability to generate new revenue streams.

The plethora of Edge devices, including sensors and Edge systems, is projected to result in 55.9 billion connected devices with 79 ZB (zettabytes of data) worldwide by 2025, according to IDC. In comparison, there was 13 ZB of Edge data worldwide in 2019, according to IDC. The IDC report includes data that was generated by work-from-home (WFH) environments, along with data generated by branch operations, factories, retail stores, and other work sites.

KEY QUESTIONS

Key questions for the data-protection discussion included:

- Is there a best practice for protecting data at the Edge, especially in a work-from-home (WFH) environment?
- 2. Is SASE (Secure Access Service Edge) mature enough to provide the level of data protection that is needed at the Edge or are other tools still required?
- 3. What can companies do to prevent data extraction from Edge devices?
- 4. How does an enterprise ensure that its DR/BC process provides the level of data protection required for data at the Edge? Can it protect against ransomware or are other actions required?
- 5. How should business executives think of the risk exposure when thinking about data at the Edge?

Key Considerations for Protecting Edge Data

The best place to start with this expansive data landscape is to identify current areas of data loss or data leakage. To accomplish this task, IT managers must look at the DLP (data loss prevention) landscape – and they must do this at-scale, to ensure that they are taking a holistic view of data protection across the entire enterprise.

46 Kent Hills Lane, Wilton, CT. 06897; (203) 429 8951;
http://www.rfgonline.com/; *Contact: inquiry@rfgonline.com*

In many cases, IT staffers may find it useful to deploy tools that can help them to identify Edge locations where data is not moving as it should – and to point to places in the infrastructure where that can be addressed. The pandemic era's WFH practices, which are blurring the lines between work data and personal data, must be examined for potential data leaks. That is why organizations must make discovery of data leakage a top priority – leveraging AI/ML software to help them widen their search, and to apply some automation to do that more effectively.

Some tips and techniques discussed during the RFG 100 videoconference, include "auto-tagging" of sensitive data; auto-discovery of end-devices and sensors; contextual heuristics, applying analytics at the Edge – and automation of the end-to-end dataflow, wherever possible. This portfolio of software tools to accomplish the mission of protecting data is growing, giving IT organizations more options to strengthen their overall data-protection strategy.

"Only the metadata needs to go everywhere," one CISO observed during the RFG 100 discussion. "Customers need visibility into the data, and they must avoid data silos, wherever possible, as a potential attack vector," he said, adding: "They need to know where the data goes, and who accesses it. Only the metadata should go everywhere – not the data. That minimizes the attack surface."

ANALYSIS

Maximum visibility and minimal risk are the twin goals of an effective data-protection policy. Access controls are a big part of an IT organization's constant oversight of data protection from consumer breaches, ransomware, and cyberattacks. Policies are needed to control data flow between on-premises sites, SaaS (software as a service) and clouds.

Ensuring that data complies with data-protection regulations and policies must be a top IT priority, especially with respect to PII (Personal Identifiable Information) – and compliance with governmental regulations, including GDPR (General Data Protection Regulation, from the European Union), and CCPA (the California Consumer Privacy Act).

A thorough process of discovery will reveal structural data-protection problems across the organization. Granular controls are needed, panelists warned. Controls for data access must be organized by personas, or roles – rather than by individual names, because individuals can have multiple personas or roles with different access rights. "Tiered" data access as you go "up the

stack," along with a policy supporting encryption of enterprise data, are powerful techniques to protect data – wherever it exists.

But the data landscape is a moving target – a temporal landscape that changes all the time. For this reason, customers need to discard or remove data, as needed, from this active landscape. One example given cited cars that are being tested in Germany, although the test data generated is reviewed by people in Singapore. Policies must be written that map when – and where – the data is to be used. In addition, preparation for ransomware attacks is needed, so that the data is protected in two or more sites, even when those attacks occur.

That complexity – involving multiple sites on a round-the-clock basis – is what makes the data-protection policies so challenging to frame. To be effective, data protection must be done at-scale, to capture the broadest view of the data. As part of this process, data leakage must be minimized. IT organizations need to have the visibility to identify where data protection problems occur – and to address the way the data flows across the network, as needed.

Executives will find that a new generation of AI/ML software tools can help to discover data leakages – and to enforce data-access policies on a global basis. These "smart" tools will help IT control the data flow between on-premises, SaaS, Clouds – and the Edge.

Continuous monitoring of data access – all the way across the data landscape – is the best way to ensure compliance with security standards (e.g., SOC, ISO). The monitoring will include audit trails and logs, on a 24 x 7 x 365 basis. "It's about people, policy, process and accountability," said another RFG 100 panelist. And it's important to make sure that data protection is part of the mindset of employees. That's why training is key to data protection success."

The Edge Requires a New View of Data Management

At the Edge, IT organizations must pay special attention to data-protection policies and procedures. That is partly due to fewer IT resources in remote sites, and partly due to compensating for latency in transmitting data to remote Edge sites. In many cases, access controls for Edge systems have been implemented differently than in core systems. Even though Edge systems have a smaller footprint than datacenter (core) systems, they must support enterprise-wide security and data-protection standards. Of course, different implementations for data protection and data management are acceptable, provided that the overall corporate controls and security/compliance requirements are met.

There is no "magic bullet" that protects all Edge data – at rest, or in-transit. That's why many customers are moving to "Zero-Trust" systems, in which all security for enterprise data is challenged, across the board. However, it's worth noting that some supporting technologies are still evolving – including SASE and the implementation of open-source codes, which may invite more hacking and attacking than traditional Core computer systems and storage systems.

Another option for data-protection is to leverage VDI solutions, with data stored at the central (Core) site, even though it is being accessed from Edge sites, closer to customers and to consumers.

Importantly, some customers look to cloud service provider (CSP) data protection software for data-in-flight and data-at-rest for those times when data that is stored within the Cloud. However, it should be noted that IT executives should not assume CSP data protection software meets corporate requirements. That kind of certification must be verified because data protection accountability resides with the corporation – and not a third-party provider.

RFG POV: The move to leverage data resources at the Edge is undeniable. Today, Edge data is an important source for enterprise-wide data, along with Core data-center data and Cloud data. All three sources contribute to the overall trove of data available for enterprise analytics – and for subsequent decision-making by business executives.

The Edge has differentiating characteristics, which must be taken into consideration for data-protection policies that span the organization or enterprise. Significant amounts of operational data are being generated by factories and stores – and even more real-time data is being generated by Edge events, remote IoT sensors and consumers.

That's why there is a need to build bridges to Cloud and Core data to assemble a more complete picture of the organization's business activities and processes. Deep analytics based on rich resources of data will lead to better business outcomes.

Finally, in a business world that is being re-shaped by the COVID-19 pandemic and WFH conditions, Edge data is vital – and that's why Edge data must be protected by a comprehensive data strategy. Given the rapid growth in Edge data, IT and business executives can no longer afford to analyze it separately – because it contains so many "real-time" aspects of overall business dynamics throughout the enterprise.

Additional relevant research and consulting services are available. Interested readers should contact Client Services to arrange further discussion or interview with Cal Braunstein, CEO and

Executive Director of Research. Jean S. Bozman, president of Cloud Architects llc, co-authored this report.

46 Kent Hills Lane, Wilton, CT. 06897; (203) 429 8951;
http://www.rfgonline.com/; *Contact: inquiry@rfgonline.com*