You Can't Secure What You Can't See: Why Data Visibility Matters

In the federal government, agencies are struggling to manage and secure vast troves of data. With disaggregated systems, they may not know where the important data resides, or understand where and how it's being used.

Data visibility is a critical foundation both in supporting any zero-trust strategy and in helping ensure agencies meet compliance requirements. A data observability pipeline will help protect an organization from cyber threats, while enabling collaboration and controlling cost. It offers agencies the means to:

- Implement zero trust
- Optimize security information and event management
- Manage sensitive data appropriately
- Share data more effectively

govloop

Cribl

The Challenge:
Lack of Data Visibility

To make effective use of data and meet compliance requirements, agencies need insight into their data. They need to know where it resides, and they need contextual information to understand what that data represents and what security controls are required. They require visibility into where and when it was created, by whom and for what purpose. The challenges include:

- **Disaggregated data stores:** When data resides in multiple systems, those responsible for security and compliance may not know what they have or where it lives.
- **Lack of context:** Without knowing where data was created, by whom and for what purposes, agencies will struggle to implement appropriate controls.
- **Weak management tools and policies:** Many agencies find they are lacking the appropriate data management tools to effectively implement zero trust across their systems, or to meet their regulatory obligations. Nor do they have policies in place to effectively implement appropriate data governance.

The Solution:
A Data Observability Pipeline

A data observability pipeline offers a way forward. This approach gives agencies deeper insight into the location and disposition of their data stores, along with the context to implement appropriate governance. A data pipeline delivers:

- **Ingestion:** Proper onboarding from all sources is key, giving IT leaders visibility and control over system performance, troubleshooting and ultimately the quality of data and data-based decisions.
- **Processing:** To make effective use of data and ensure appropriate security, the pipeline optimizes, enriches and harmonizes the data. This ensures needed context: things like IP address, geolocation, standardized data formats or standardized time stamps. All these help ensure agencies can manage that data appropriately.
- **Routing:** Data is sent where it needs to be based on its content, value and purpose. Some data, for example, should be stored as metrics, while other data is sent to low-cost storage. Proper routing makes data available and helps control costs.

"Data used to be the byproduct of doing business. Now it's become a strategic asset — and also sometimes a liability to agencies, with a whole host of new regulations, new frameworks, new considerations for what to do with data."

— **Jackie McGuire**, Senior Market Strategy Manager, Cribl

Federal Use Cases

A data observability pipeline helps support a number of key federal requirements:

Implementing Zero Trust

Agencies have prioritized zero trust, a security strategy that requires all users to be authenticated, authorized and continuously validated. "Zero trust depends in part on context. Does this individual typically access this data, at this time, from this computer?" said Jackie McGuire, Senior Market Strategy Manager at Cribl.

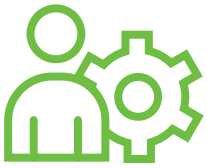
Optimizing SIEM

"In security information and event management, the name of the game is boosting signal-to-noise ratio," McGuire said.

An IT security log, for instance, might have 200 lines of information, of which only 20 actually matter. With proper formatting, the observability pipeline reduces the noise going to the SIEM, making it far easier to identify and remediate anomalies.

"Your machine learning and your artificial intelligence are picking it up faster, because they're not sorting through all those garbage fields in the logs, trying to figure out what actually matters," McGuire said.

Managing Sensitive Data



Agencies are especially interested in how personally identifiable information and other forms of sensitive data are handled. An observability pipeline is critical to meeting these needs.

"This starts at the ingest, with the ability to mask and obfuscate data at the source," McGuire said. "Data processing and data routing are also important. If you're never moving that sensitive data in the first place, or if you're breaking it off and only moving it to the places where you have authority to hold that sensitive data, that in and of itself solves a lot of those problems."

Sharing Data

Agencies require the means to share data easily and appropriately among relevant stakeholders.

With an observability pipeline, "you can put data into file systems and datastores where they can be shared," McGuire said. Administrators can create "very easy and fluid roles-based processes to determine how and where to share that data," she said.

Case Study: A Federal Agency Tackles Cyber Requirements

The challenge: A federal agency needed greater visibility in order to ensure effective uses of its data for both the agency and component stakeholders, as well as to meet compliance obligations such as the cybersecurity requirements laid out in the recent Executive Order on Improving the Nation's Cybersecurity.

The solution: The agency implemented the Cribl Stream data observability solution as on-premises software running in its Linux environments. The observability pipeline sits between where the data is created and where it's needed. It supports data collection and ingestion, as well as processing to refine, reduce, enrich and ultimately route data – getting the information into the hands of those who need it.

The outcomes: The data pipeline gave the agency greater visibility and manageability over its disparate data stores, enabling it to meet its cybersecurity obligations more effectively. The agency saved money by storing data in the most appropriate locations, and made its data more readily shareable in support of mission needs.

"They were able to easily get the data in, shape it for multiple different purposes including enriching the data with the context of whose systems produced the data, and route it to where it was most useful. This enabled them to solve some of their biggest data challenges, resulting in better threat detection and enabling sharing with multiple stakeholders."

— **Jackie McGuire**, Senior Market Strategy Manager, Cribl

How Cribl Helps

The Cribl Stream data observability pipeline delivers greater visibility and manageability. It's simple to get up and running, with graphical interfaces for easy configuration, and it is vendor agnostic: It works with any data, anywhere. Cribl supports its tools with free education for life, including a range of training and certification programs.

govloop

Cribl