National Cybersecurity Protection System Cloud Interface Reference Architecture

Volume Two: Reporting Pattern Catalog

December 2020

Version 1.0 (Draft)

Cybersecurity and Infrastructure Security Agency
Cybersecurity Division
Capability Delivery Subdivision
NCPS Program Management Office

Revision/Change Record

Version	Date	Revision Description	Section/Page Affected
Version 1.0	12/22/2020	Initial Release Version	All

EXECUTIVE SUMMARY

The National Cybersecurity Protection System (NCPS) Program is evolving to ensure that security information about cloud-based traffic can be captured and analyzed and Cybersecurity and Infrastructure Security Agency (CISA) analysts can continue to provide situational awareness and support to the agencies. To support this goal, CISA is developing a cloud-based architecture to collect and analyze agency cloud security data. This reference architecture explains how agencies can interact with that system. It includes background about how the cloud impacts NCPS, discusses what security information needs to be captured in the cloud and how it can be captured, and provides reporting patterns to explain how that information can be sent to CISA.

The *NCPS Cloud Interface Reference Architecture* is being released as two individual volumes. The first volume provides an overview of changes to NCPS to accommodate the collection of relevant data from agencies' cloud environments and provides general reporting patterns for sending cloud telemetry to CISA. This second volume builds upon the concepts presented in *NCPS Cloud Interface Reference Architecture: Volume One* and provides an index of common cloud telemetry reporting patterns and characteristics for how agencies can send cloud-specific data to the NCPS cloud-based architecture. Individual cloud service providers (CSPs) can refer to the reporting patterns in this volume to offer guidance on their solutions that allow agencies to send cloud telemetry to CISA in fulfillment of NCPS requirements.

A cloud-based NCPS architecture is currently in development at CISA. This *NCPS Cloud Interface Reference Architecture* is being released to Federal Civilian Agencies in advance of a deployed system to accomplish the following:

- Notify agencies about changes in the NCPS Program and give them time to plan.
- Solicit feedback from agencies so that a final version of this reference architecture provides desired content and meets the needs of agencies.
- Gather requirements from agencies to ensure the cloud-based NCPS architecture can support agency use cases.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Federal civilian departments and agencies[1] must meet the requirements of the National Cybersecurity Protection System (NCPS).[2] CISA analysts use this data for 24/7 situational awareness, analysis, and incident response. Traditionally, NCPS sensors located at Trusted Internet Connections (TIC) and Managed Trusted Internet Protocol Service (MTIPS) gateways capture security information as traffic passes between the agency and the Internet. As agencies move their information technology (IT) infrastructure to the cloud, some network traffic no longer traverses traditional NCPS sensors, and security information about that traffic is no longer captured by NCPS.

The NCPS Program is evolving to ensure that security information about cloud-based traffic can be captured and analyzed and CISA analysts can continue to provide situational awareness and support to the agencies. To support this goal, CISA is deploying a cloud-based architecture, the Cloud Log Aggregation Warehouse (CLAW), to collect and analyze agency cloud security data. CISA has released the *NCPS Cloud Interface Reference Architecture (NCIRA)* as a two-volume document set to explain how agencies can provide cloud-generated security information to the CLAW. Volume One introduces fundamental concepts about cloud data aggregation and reporting patterns (including attributes and options for how agencies can send cloud telemetry to NCPS). Volume Two provides a catalog of common reporting patterns based on the reporting pattern framework developed in Volume One**.**

NCIRA Volume Two (this document) is a continuation of NCIRA Volume One and builds on the concepts presented in that document. In order to understand and implement the reporting patterns presented in this document, agencies must be familiar with the concepts introduced in NCIRA Volume One.

1.1 Document Organization

This document is structured to facilitate readability and ease of use. *NCPS Cloud Interface Reference Architecture: Volume Two* consists of five sections and three appendices.

- Section 1 provides a document overview and a guide on how to use this volume in conjunction with Volume One.
- Section 2 describes the cloud telemetry reporting pattern characteristics and their implications.
- Section 3 is a catalog of simple reporting patterns that can be mapped to common agency cloud use cases.
- Section 4 is a catalog of more complex reporting patterns that combine one or more of the individual patterns developed in Section 3.
- Section 5 discusses conclusions and future work.
- Appendix A provides in-depth analysis of the Cloud Telemetry Timeliness characteristic.
- Appendix B provides in-depth analysis of the Cloud Telemetry Timing Coordination characteristic.
- Appendix C provides in-depth analysis of the Cloud Telemetry Provenance characteristic.

[1] For the purposes of this document, the term "agency" will hereinafter be used to refer to all federal civilian executive branch departments and agencies.
[2] https://www.dhs.gov/cisa/national-cybersecurity-protection-system-ncps

1.2 Purpose

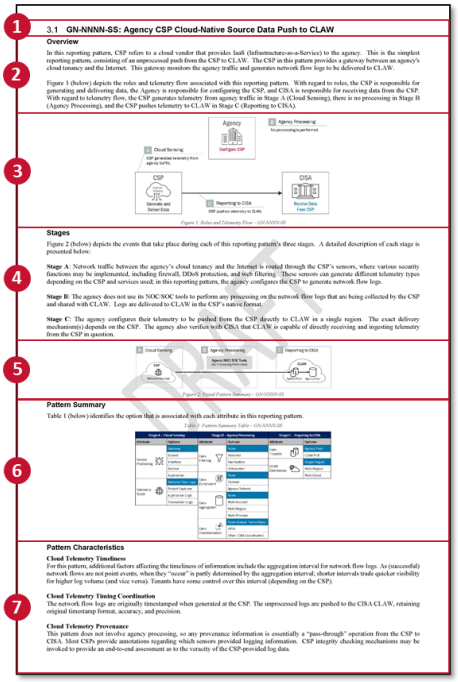
A reference architecture is an authoritative source of information about a specific subject area that guides and constrains the instantiations of multiple architectures and solutions. The purpose of this reference architecture is to explain what information agencies need to capture in the cloud for NCPS, how that information can be captured, and how it can be sent to CISA. This reference architecture is divided into two volumes.

1. Volume One of the *NCPS Cloud Interface Reference Architecture* provides general guidance for agencies on participating in NCPS in the cloud. The information provided includes the introduction of general reporting patterns. The discussion in Volume One is vendor-agnostic and not specific to any particular CSP.
2. Volume Two of the *NCPS Cloud Interface Reference Architecture* contains a catalog of reporting patterns for how agencies can participate in NCPS in the cloud under different cloud service models. The catalog includes individual reporting patterns (typical of an agency using a single CSP) as well as complex reporting patterns (illustrating how an agency can use several cloud service models and providers and send cloud security data to NCPS in the cloud).

1.3 Document Guide

Section 2 of this document discusses the NCPS cloud telemetry characteristics that are common across the reporting patterns in Sections 3 and 4. Sections 3 and 4 of this document are intended to serve as an index to common reporting patterns; it is not necessary for the document to be read in its entirety. Agencies should identify which reporting patterns apply to their cloud use cases and use these patterns to implement NCPS in the Cloud. As shown in Figure 1, each reporting pattern in this document is presented in the following format.

1. **Identifier and Title**: The naming scheme and title description provides a high-level summary of the reporting pattern and the attribute options leveraged.
2. **Overview**: An overview that provides the reader with a brief summary of the reporting pattern, including information used to understand its context and application.
3. **Roles and Telemetry Flow Figure**: This figure depicts which entity is responsible for each of the three telemetry reporting stages and what functions they perform.
4. **Stage Summary**: The stage summary provides an explanation about how each of the three telemetry reporting stages are performed.
5. **Visual Pattern Summary Figure**: This figure provides a visual summary of options selected for each of the attributes in each stage of the pattern.
6. **Pattern Summary Table**: This table articulates the option selected for each of the attributes in each stage of the pattern.
7. **Pattern Characteristics**: Additional details describe the pattern-level characteristics for full agency cloud telemetry sharing.



2 REPORTING PATTERN-LEVEL CHARACTERISTICS

As NCPS evolves to accommodate cloud services, agencies will be required to implement reporting patterns and maintain telemetry sharing with CISA. NCIRA Volume One introduced the three stage reporting pattern concept and discussed a series of attributes and options that agencies needed to select in each stage of a reporting pattern. NCIRA Volume Two (this document) introduces six high-level characteristics that apply to individual reporting patterns as a whole (including all three of their stages) that agencies need to evaluate when selecting which reporting pattern(s) to employ. These characteristics will be negotiated between agencies, CSPs, and CISA during CLAW integration activities. These six reporting pattern-level characteristics are as follows:

1. **Cloud Telemetry Timeliness**: The duration between cloud telemetry creation and presentation of that information to CISA analysts (to accommodate response within cyber-relevant time).
2. **Cloud Telemetry Timing Coordination**: Telemetry timestamp labelling mechanism in use (to enable CISA processing and proper record sequencing).
3. **Cloud Telemetry Provenance**: Telemetry source attribution and labelling (which may be complicated by agency aggregation and processing).
4. **Reporting Connection Administration**: Data transfer initiation, maintenance, and retirement execution.
5. **Cloud Telemetry Sharing Cost**: Expenses incurred for cloud sensing, agency processing, and reporting to CISA (based on attribute options selected).
6. **Agency Data Retention and Use Constraints:** Any additional data handling, retention, and use constraints, as captured in agency and CISA MOUs.

Each of these characteristics have different implications that agencies will need to weigh when selecting a reporting pattern. The first three characteristics (cloud telemetry timeliness, cloud telemetry timing coordination, and cloud telemetry provenance) involve nuanced technical implications and are discussed individually below at a high level. A more detailed discussion of these items is presented in Appendix A (Cloud Telemetry Timeliness), Appendix B (Cloud Telemetry Provenance), and Appendix C (Cloud Telemetry Timing Coordination).

2.1 Cloud Telemetry Timeliness

Different CSPs have different timeframes for log delivery. While typical values range between a few minutes to fifteen minutes of event occurrence, agencies must confirm the timeliness of a CSP's log delivery through discussions with the CSP and their own testing. In most cases, neither the service documentation nor the published Service Level Agreements make a concrete statement regarding the timeliness of log delivery. While one CSP might claim that "events are delivered within five minutes of occurrence," another might claim that "events are delivered in real-time," and another might only provide hints via screenshots. Even within a CSPs offerings, more common/popular services are likely to have better documentation around timeliness than other services.

Some generalizations may be made based on log type. Logs concerning point-in-time events (e.g., transaction logs for auditing Application Programming Interface (API) calls) can be delivered quickly, whereas those concerning continuous events (e.g., network flow logs or application metrics detailing resource usage) have some interval that must transpire before the event is recorded and delivered. In the latter case, tenants may be given some control over the interval, with the caveat that shorter intervals incur greater costs than the default/free interval.

CISA's goal is to detect, investigate, and respond to *any* threat before it has time to evolve and progress. Although CISA acknowledges that an agency has limited control over the timeliness of a CSP's delivery of raw logs, once the logs are received from the CSP, it is the agency that largely determines how long it takes to process the logs and deliver them to CLAW. Agencies should ensure that the time between raw logs release to the agency tenant from the CSP and the delivery of the processed logs to CLAW is within 30 minutes.

Additional cloud telemetry timeliness details can be found in Appendix A.

CISA Preference

When agency processing is performed, CISA expects the time between receiving raw logs from the CSP and the delivery of processed logs to CLAW to not exceed 30 minutes.

2.2 Cloud Telemetry Timing Coordination

Time synchronization is the coordination of the system clocks (agency, branch, and remote) and the components that comprise the systems (servers, workstations, network devices, etc.). Modern network infrastructures may have multiple links, network tiers, or data centers between the point the data is captured and the point where the analysis is performed. The insertion of a standardized timestamp is a common method for preserving the telemetry generation times. This method is widely used in the industry, but the implementation specifics (timestamp accuracy, format, etc.) vary based on the application. One of the key requirements for accuracy when performing any kind of analytics is understanding precisely when an event generating an observable record took place. Telemetry timestamping is essential for data analysis in modern networks (network troubleshooting, application performance tracking, security or threat analysis and legal compliance). Any time-specific analysis performed is dependent upon the accuracy and precision of the timestamps of data being analyzed.

In the simplest case, the source (the CSP) and the destination (the CISA CLAW) both influence timing synchronization, and discrepancies may occur between the systems. The cloud telemetry logs are timestamped when the log entries are generated. The logs are available for examination with the agency processing tools, where the original telemetry timestamps can be viewed but must not be altered. When the logs are pushed to the CISA CLAW, the originally generated log timestamps are retained. The cloud telemetry timestamp format must be coordinated between agencies and CISA to ensure compatibility and accurate processing.

Additional cloud telemetry timing coordination details can be found in Appendix B.

CISA Preference

When feasible, cloud-native telemetry timestamp format, precision, and accuracy should be preserved by agency processing to ensure accurate processing and use by CLAW systems and analysts.

2.3 Cloud Telemetry Provenance

Data provenance of an information object refers to the process of tracing and recording the object's origin and history. Generally, provenance will include author identification, modification times, and some degree of activities performed that have affected the object's content or handling, as well as a method to ensure integrity of the history and object itself. Provenance information can help analysts and systems trace telemetry sources, track changes in data richness over time, identify the scope of untrustworthy data (if a telemetry source is misbehaving for whatever reason), and track updates as new telemetry versions become available, etc.

Provenance of cloud telemetry must be conveyed by agencies to CISA at sharing initiation and on an ongoing basis. When an agency inhabits multiple tenancies and reports information to CISA in a push form, log information may require a form of fusing and editing. Assuming a common log type and format across each data source in each tenant, the agency may aggregate the sources either by interleaving or combining them in some other fashion (e.g., data from one tenancy might precede that from another). In this case, provenance claims are likely to be made by the agency above and beyond each data source. The agency would be responsible for asserting that it provided the aggregation of the multiple streams, and constituent streams may retain sufficient provenance information to be checked end-to-end by CISA when no agency filtration is performed. A multi-tenant agency responding to CISA pull requests may be able to enable custom-tailored responses.

As an agency performs additional levels of processing, data filtering and enrichment may occur. In this case, the agency is an author of log information, as it is providing enrichment and editing. Agency processing should be arranged to convey both the nature of the modifications (e.g., enrichment) performed, the type of information removed, and the processing mechanisms (e.g., software artifacts) used in performing the processing.

Additional cloud telemetry provenance details can be found in Appendix C.

CISA Preference

Provenance of cloud telemetry must be conveyed by agencies to CISA at sharing initiation and on an ongoing basis.

3 GENERIC REPORTING PATTERNS

The selection of a combination of Stage A, B, and C options constitutes a reporting pattern. Stage A addresses Cloud Sensing, Stage B addresses Agency Processing, and Stage C covers Reporting to CISA. Given that there are three stages in any reporting pattern, each with multiple attributes and options (as listed in Table 1), there are many possible reporting patterns. Therefore, it is desirable to have a scheme for easily identifying generic reporting patterns. Each generic reporting pattern will be identified by an eight-character identifier in the format shown in Figure 2.

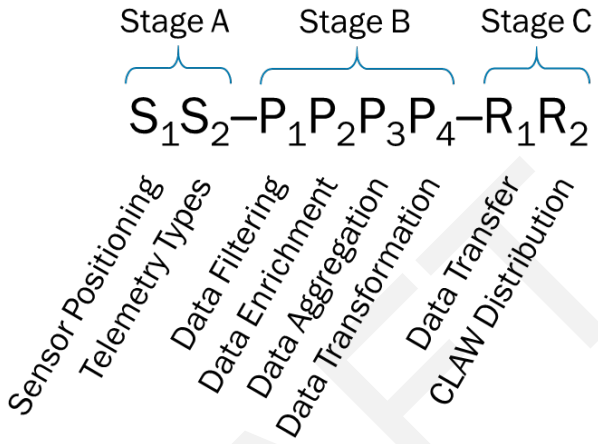


Figure 2: Reporting Pattern Identifier Format

The acceptable values for each character position, and their corresponding option, are listed in Table 1.

Table 1: Reporting Pattern Identification

Stage A – Cloud Sensing		Stage B – Agency Processing		Stage C – Reporting to CISA	
Attribute	**Option and Letter Code**	**Attribute**	**Option and Letter Code**	**Attribute**	**Option and Letter Code**
Sensor Positioning (S_1)	Gateway (G)	**Data Filtering (P_1)**	None (N)	**Data Transfer (R_1)**	Agency Push (S)
	Subnet (N)		Removal (R)		CLAW Pull (L)
	Interface (I)		Sanitization (S)	**CLAW Distribution (R_2)**	Single Region (S)
	Service (S)		Obfuscation (O)		Multi-Region (R)
	Application (A)	**Data Enrichment (P_2)**	None (N)		Multi-Cloud (C)
Telemetry Types (S_2)	Network Flow Logs (N)		Derived (D)		
	Packet Captures (P)		Agency-Defined (A)		
	Application Logs (A)	**Data Aggregation (P_3)**	None (N)		
	Transaction Logs (T)		Multi-Account (A)		
			Multi-Region (R)		
			Multi-Provider (P)		
		Data Transformation (P_4)	None (N)		
			IPFIX (I)		
			CISA Coordinated (C)		

For example, using the preceding table, the identifier "*GN-NNNN-SS*" indicates that there is a gateway sensor sending network flow logs (Stage A), with no additional processing (Stage B), and an agency push to CLAW in a single region (Stage C).

Each generic reporting pattern also includes a short name to better accommodate conversation. These short names will be provided as part of the reporting pattern title.

Based on the variety of options available, there are many reporting pattern permutations, and it is not practical to discuss every possible permutation within this document. Instead, this document will focus on a small set of reporting patterns that represent current and planned CLAW pilot activities. Patterns not shown here may still be viable alternatives and should be discussed with CISA on a case-by-case basis for adoption and possible inclusion in future versions of this volume.

While these patterns focus on the process of delivering telemetry to CLAW, agencies may use the same telemetry in their own analytics process. When considering how well each pattern would satisfy an agency's need to share telemetry with CISA, agencies should also consider how well the pattern overlaps with their existing analytics process; leveraging this overlap may result in significant cost savings.

3.1 GN-NNNN-SS: Agency CSP Cloud-Native Source Data Push to CLAW

Overview

In this reporting pattern, CSP refers to a cloud vendor that provides Infrastructure-as-a-Service (IaaS) to the agency. This is the simplest reporting pattern, consisting of an unprocessed push from the CSP to CLAW. The CSP in this pattern provides a gateway between an agency's cloud tenancy and the Internet. This gateway monitors the agency traffic and generates network flow logs to be delivered to CLAW.[3]

Figure 3 (below) shows the roles and telemetry flow associated with this reporting pattern. With regard to roles, the CSP is responsible for generating and delivering data, the Agency is responsible for configuring the CSP, and CISA is responsible for receiving data from the CSP. With regard to telemetry flow, the CSP generates telemetry from agency traffic in Stage A (Cloud Sensing), there is no processing in Stage B (Agency Processing), and the CSP pushes telemetry to CLAW in Stage C (Reporting to CISA).

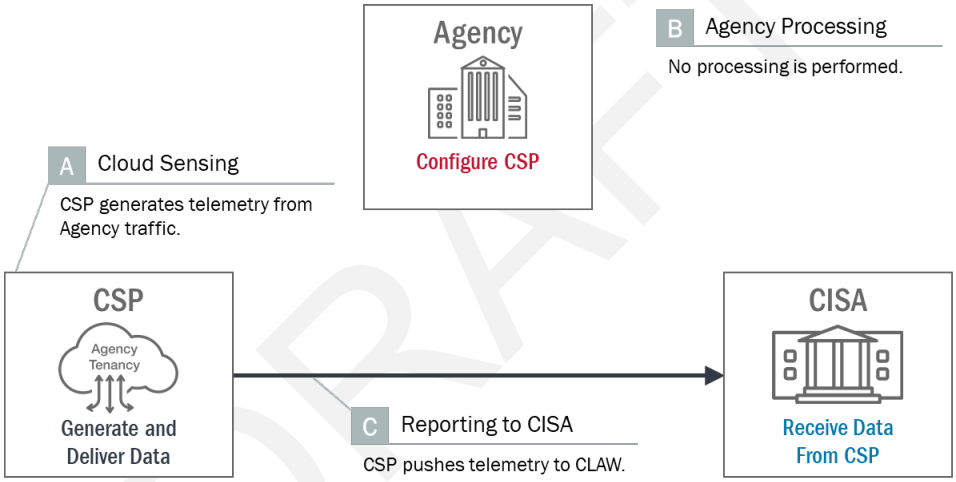


Figure 3: Roles and Telemetry Flow – GN-NNNN-SS

Stages

Figure 4 (below) shows the events that take place during each of this reporting pattern's three stages. A detailed description of each stage is presented below:

Stage A: Network traffic between the agency's cloud tenancy and the Internet is routed through the CSP's sensors, where various security functions may be implemented, including firewall, Distributed Denial of Service (DDoS) protection, and web filtering. These sensors can generate different telemetry types depending on the CSP and services used; in this reporting pattern, the agency configures the CSP to generate network flow logs.

Stage B: The agency does not use its NOC/SOC tools to perform any processing on the network flow logs that are being collected by the CSP and shared with CLAW. Logs are delivered to CLAW in the CSP's native format.

[3] The CSP may also provide gateways between the agency's cloud tenancy and external networks that are not the Internet, such as the agency's on-premise network. These gateways provide similar monitoring capabilities.

Stage C: The agency configures their telemetry to be pushed from the CSP directly to CLAW in a single region. The exact delivery mechanism(s) depends on the CSP. The agency also verifies with CISA that CLAW is capable of directly receiving and ingesting telemetry from the CSP in question.[4]

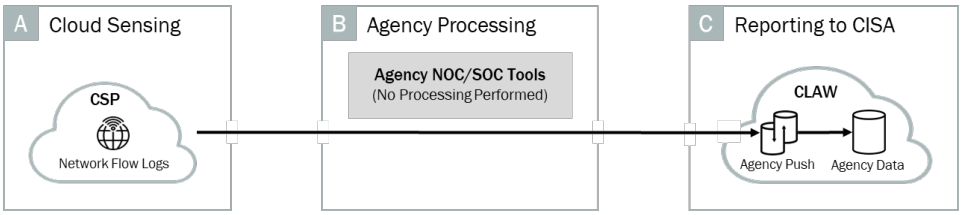


Figure 4: Visual Pattern Summary – GN-NNNN-SS

Pattern Summary

Table 2 (below) lists the option that is associated with each attribute in this reporting pattern.

Table 2: Pattern Summary Table – GN-NNNN-SS

Stage A – Cloud Sensing		Stage B – Agency Processing		Stage C – Reporting to CISA	
Attribute	**Options**	**Attribute**	**Options**	**Attribute**	**Options**
Sensor Positioning	Gateway	Data Filtering	None	Data Transfer	Agency Push
	Subnet		Removal		CLAW Pull
	Interface		Sanitization	CLAW Distribution	Single Region
	Service		Obfuscation		Multi-Region
	Application	Data Enrichment	None		Multi-Cloud
Telemetry Types	Network Flow Logs		Derived		
	Packet Captures		Agency-Defined		
	Application Logs	Data Aggregation	None		
	Transaction Logs		Multi-Account		
			Multi-Region		
			Multi-Provider		
		Data Transformation	None (Native Forms Align)		
			IPFIX		
			Other (CISA Coordinated)		

Pattern Characteristics

Cloud Telemetry Timeliness

For this pattern, additional factors affecting the timeliness of information include the aggregation interval for network flow logs. As (successful) network flows are not point events, when they "occur" is partly determined by the aggregation interval; shorter intervals trade quicker visibility for higher log volume (and vice versa). Tenants have some control over this interval (depending on the CSP).

Cloud Telemetry Timing Coordination

The network flow logs are originally timestamped when generated at the CSP. The unprocessed logs are pushed to the CISA CLAW, retaining original timestamp format, accuracy, and precision.

[4] If this is not the case, e.g. for CSP's that are not commonly used by agencies, the agency may request CISA to add support for this CSP. Alternatively, the agency themselves can transform the data into IPFIX or another format that they negotiate with CISA; refer to Patterns 6-8 for examples.

Cloud Telemetry Provenance

This pattern does not involve agency processing, so any provenance information is essentially a "pass-through" operation from the CSP to CISA. Most CSPs provide annotations regarding which sensors provided logging information. CSP integrity checking mechanisms may be invoked to provide an end-to-end assessment as to the veracity of the CSP-provided log data.

3.2 SN-NNNN-LS: CLAW Pull from Agency CSP Cloud-Native Source

Overview

In this reporting pattern, CSP refers to an agency's Platform-as-a-Service (PaaS) cloud vendor. CLAW sends requests to pull data from the CSP, which in turn responds with the desired telemetry. The CSP in this pattern may provide various services, such as load balancing, network/application firewalls, Domain Name System (DNS), identity/authentication, key management, web hosting, etc. These services each generate telemetry, which is made available through an API (either the service's own, or, if the telemetry is exported to a CSP storage service, then that service's API).

Figure 5 (below) shows the roles and telemetry flow associated with this reporting pattern. With regard to roles, the CSP is responsible for generating and storing data, the Agency is responsible for configuring the CSP, and CISA is responsible for retrieving data from the CSP. With regard to telemetry flow, the CSP generates telemetry from agency traffic in Stage A (Cloud Sensing), there is no processing in Stage B (Agency Processing), and CLAW pulls telemetry from the CSP in Stage C (Reporting to CISA).

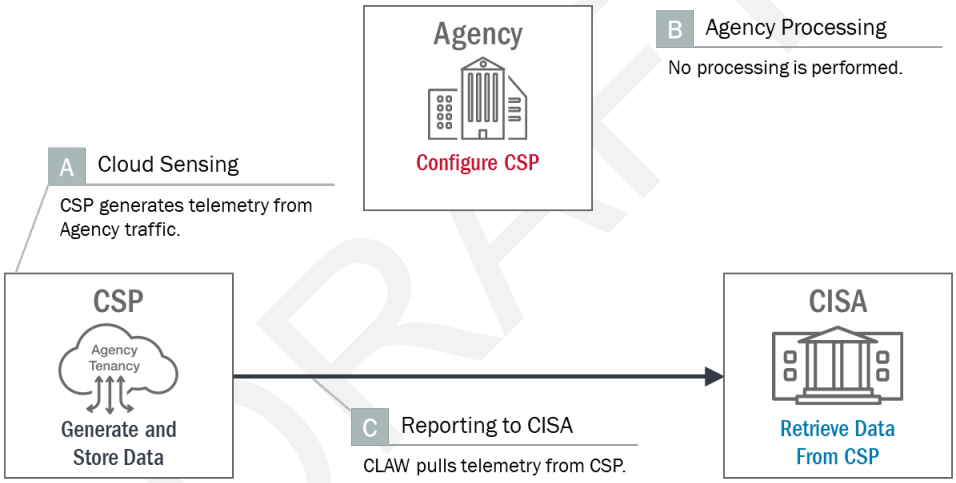


Figure 5: Roles and Telemetry Flow – SN-NNNN-LS

Stages

Figure 6 (below) shows the events that take place during each of this reporting pattern's three stages. A detailed description of each stage is presented below:

Stage A: Network traffic between the agency's cloud tenancy and the internet is routed through the CSP services; the agency configures one or more of these services to generate network flow logs.

Stage B: The agency does not use its NOC/SOC tools to perform any processing on the network flow logs that are being collected by the CSP and shared with CLAW. Logs are delivered to CLAW in the CSP's native format.

Stage C: The agency configures their telemetry to be supplied from the CSP service directly to CLAW in a single region. This involves configuring permissions on the CSP such that CLAW[5] has the proper

[5] That is, an authenticated identity principal corresponding to the CLAW instance in the selected region.

pull credentials to make the necessary requests and pull the network flow logs from the CSP. The exact delivery mechanism(s) depends on the CSP.

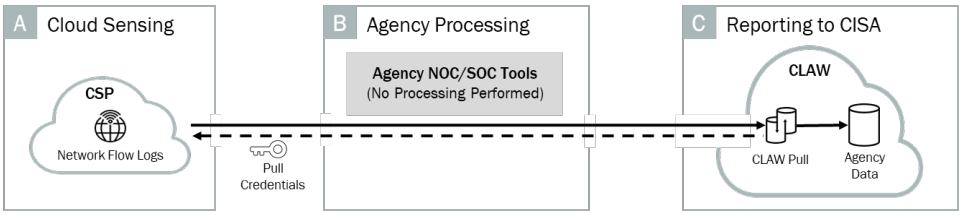


Figure 6: Visual Pattern Summary – SN-NNNN-LS

Pattern Summary

Table 3 (below) lists the option that is associated with each attribute in this reporting pattern.

Table 3: Pattern Summary Table – SN-NNNN-LS

Stage A – Cloud Sensing		Stage B – Agency Processing		Stage C – Reporting to CISA	
Attribute	**Options**	**Attribute**	**Options**	**Attribute**	**Options**
Sensor Positioning	Gateway	Data Filtering	**None**	Data Transfer	Agency Push
	Subnet		Removal		**CLAW Pull**
	Interface		Sanitization	CLAW Distribution	**Single Region**
	Service		Obfuscation		Multi-Region
	Application	Data Enrichment	**None**		Multi-Cloud
Telemetry Types	**Network Flow Logs**		Derived		
	Packet Captures		Agency-Defined		
	Application Logs	Data Aggregation	**None**		
	Transaction Logs		Multi-Account		
			Multi-Region		
			Multi-Provider		
		Data Transformation	**None (Native Forms Align)**		
			IPFIX		
			Other (CISA Coordinated)		

Pattern Characteristics

Cloud Telemetry Timeliness

For this pattern, factors affecting the timeliness of information include the aggregation interval for network flow logs and the polling frequency of CLAW. As (successful) network flows are not point events, when they "occur" is partly determined by the aggregation interval; shorter intervals trade quicker visibility for higher log volume (and vice versa). Tenants have some control over this interval (depending on the CSP). The frequency with which CLAW checks for and pulls new data adds delay and is limited by mechanisms such as API request throttling.

Cloud Telemetry Timing Coordination

The network flow logs are originally timestamped when generated at the CSP. The unprocessed logs are pulled by CLAW, retaining original timestamp format, accuracy, and precision.

Cloud Telemetry Provenance

This pattern does not involve agency processing, so any provenance information is essentially a "pass-through" operation from the CSP to CISA. Most CSPs provide annotations regarding which sensors provided logging information. CSP integrity checking mechanisms may be invoked to provide an end-to-end assessment as to the veracity of the CSP-provided log data.

3.3 GT-NNAN-SS: Agency Aggregated Data Push to CLAW

Overview

In this reporting pattern, CSP refers to an agency's IaaS cloud vendor. The agency has multiple accounts with the CSP and aggregates data from each before sending to CLAW. The CSP in this pattern provides gateways between each of the agency's cloud tenancies and the Internet. These gateways monitor the agency traffic and generate transaction logs to be delivered to CLAW. The transaction log types are the same for each agency tenancy.

Figure 7 (below) shows the roles and telemetry flow associated with this reporting pattern. With regard to roles, the CSP is responsible for generating and delivering data, the Agency is responsible for configuring the CSP and combining data, and CISA is responsible for receiving data from the Agency. With regard to telemetry flow, the CSP generates telemetry from agency activity on multiple accounts in Stage A (Cloud Sensing), the Agency aggregates telemetry received from the CSP in Stage B (Agency Processing), and the Agency pushes telemetry to CLAW in Stage C (Reporting to CISA).

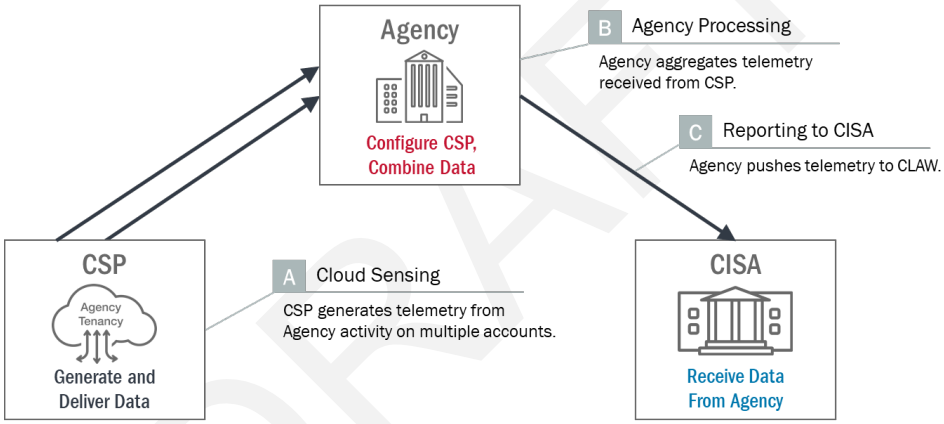


Figure 7: Roles and Telemetry Flow – GT-NNAN-SS

Stages

Figure 8 (below) shows the events that take place during each of this reporting pattern's three stages. A detailed description of each stage is presented below:

Stage A: Network traffic between the agency's cloud tenancies and the Internet is routed through the CSP's gateway sensors, where various security functions may be implemented, including firewall, DDoS protection, and web filtering. These gateway sensors can generate different telemetry types depending on the CSP and services used; in this reporting pattern, the agency configures the CSP to generate transaction logs for each account.

Stage B: The agency uses its NOC/SOC tools to aggregate the transaction logs from multiple CSP accounts into a single stream. The format of the logs (i.e., the CSP's native format) is preserved, and the agency does not perform any filtering or enrichment.

Stage C: The agency pushes the aggregated telemetry to CLAW in a single region. The exact delivery mechanism(s) depends on the CSP.

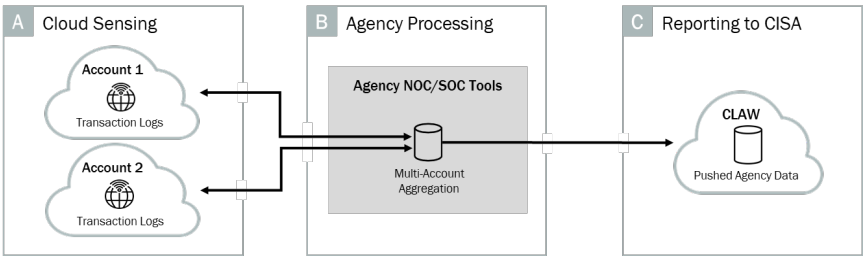


Figure 8: Visual Pattern Summary – GT-NNAN-SS

Pattern Summary

Table 4 (below) lists the option that is associated with each attribute in this reporting pattern.

Table 4: Pattern Summary Table – GT-NNAN-SS

Stage A – Cloud Sensing		Stage B – Agency Processing		Stage C – Reporting to CISA	
Attribute	**Options**	**Attribute**	**Options**	**Attribute**	**Options**
Sensor Positioning	Gateway	Data Filtering	None	Data Transfer	Agency Push
	Subnet		Removal		CLAW Pull
	Interface		Sanitization	CLAW Distribution	Single Region
	Service		Obfuscation		Multi-Region
	Application	Data Enrichment	None		Multi-Cloud
Telemetry Types	Network Flow Logs		Derived		
	Packet Captures		Agency-Defined		
	Application Logs	Data Aggregation	None		
	Transaction Logs		Multi-Account		
			Multi-Region		
			Multi-Provider		
		Data Transformation	None (Native Forms Align)		
			IPFIX		
			Other (CISA Coordinated)		

Pattern Characteristics

Cloud Telemetry Timeliness

For this pattern, factors affecting the timeliness of information include the agency's own policy for delivery to CLAW. Agencies can delay delivering individual records/objects (e.g., as part of a batching policy) and may do so if they do not exceed the maximum processing delay parameters.

Agency processing itself should not significantly affect timeliness; aggregating a common log type from multiple sources is expected to be a low complexity operation, facilitating rapid execution.

Cloud Telemetry Timing Coordination

In this case, the processing stage will have an opportunity to introduce its own timestamps into the overall chain that originates from the source and terminates at the CISA CLAW. However, the gateway transaction logs are still timestamped when the log entry is generated at the CSP. Additional timestamps may be added at the time when the log entries are aggregated. However, the original log entries' timestamps should be preserved.

Cloud Telemetry Provenance

This pattern involves agency processing on multiple log streams across multiple tenancies. As a common log type and format is assumed across each data source, the agency is able to aggregate the sources either by interleaving or combining them in some other fashion (e.g., data from one tenancy

might precede that from another). In this case, provenance claims are likely to be made by the agency. In particular, although multiple streams may arrive at the agency labeled and integrity-protected, the process of interleaving would create a new stream that itself requires provenance metadata. In short, the agency would be responsible for asserting that it provided the aggregation of the multiple streams, and constituent streams may retain sufficient provenance information to be checked end-to-end by CISA.

3.4 ST-NNAN-LS: CLAW Pull of Agency Aggregated Service Data

Overview

In this reporting pattern, CSP refers to an agency's IaaS cloud vendor. The CSP in this pattern provides various services, such as load balancing, network/application firewalls, DNS, identity/authentication, key management, etc. These services generate separate telemetry streams for each account. The agency then gathers data from the multiple accounts and aggregates it. CLAW sends requests to pull data from the agency, which in turn responds with the desired telemetry.

Figure 9 (below) shows the roles and telemetry flow associated with this reporting pattern. With regard to roles, the CSP is responsible for generating and delivering data, the Agency is responsible for configuring the CSP and combining data, and CISA is responsible for retrieving data from the Agency. With regard to telemetry flow, the CSP generates telemetry from agency activity on multiple accounts in Stage A (Cloud Sensing), the Agency aggregates telemetry received from the CSP in Stage B (Agency Processing), and CLAW pulls telemetry from the Agency in Stage C (Reporting to CISA).

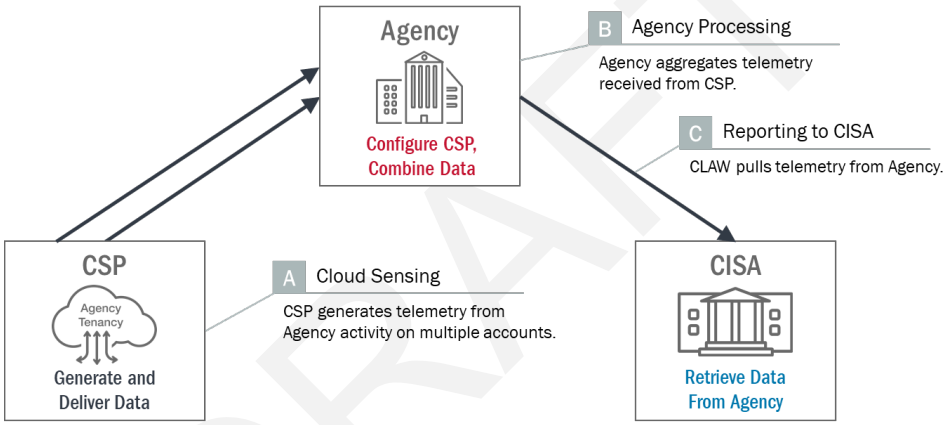


Figure 9: Roles and Telemetry Flow – ST-NNAN-LS

Stages

Figure 10 (below) shows the events that take place during each of this reporting pattern's three stages. A detailed description of each stage is presented below:

Stage A: Network traffic between the agency's cloud tenancies and the internet is routed through the CSP services; the agency configures one or more of these services in each account to generate transaction logs.

Stage B: The agency uses its NOC/SOC tools to aggregate the transaction logs from multiple CSP accounts. Once merged into a single stream, the aggregated logs may then be stored for later retrieval by CLAW. The aggregation may take place on agency premise equipment or may occur on agency-configured CSP infrastructure. The format of the logs (i.e., the CSP's native format) is preserved and the agency does not perform any filtering or enrichment.

Stage C: The agency supplies the aggregated telemetry to be pulled by CLAW in a single region. This involves configuring pull credentials such that CLAW[6] is authorized to make the necessary requests. The exact delivery mechanism(s) depends on the CSP.

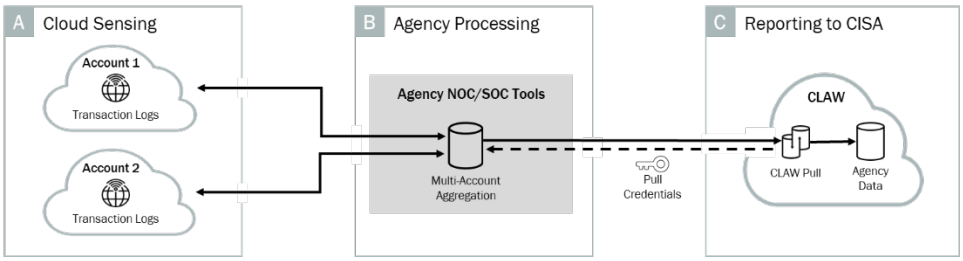


Figure 10: Visual Pattern Summary – ST-NNAN-LS

Pattern Summary

Table 5 (below) lists the option that is associated with each attribute in this reporting pattern.

Table 5: Pattern Summary Table – ST-NNAN-LS

Stage A – Cloud Sensing		Stage B – Agency Processing		Stage C – Reporting to CISA	
Attribute	**Options**	**Attribute**	**Options**	**Attribute**	**Options**
Sensor Positioning	Gateway	Data Filtering	None	Data Transfer	Agency Push
	Subnet		Removal		CLAW Pull
	Interface		Sanitization	CLAW Distribution	Single Region
	Service		Obfuscation		Multi-Region
	Application	Data Enrichment	None		Multi-Cloud
Telemetry Types	Network Flow Logs		Derived		
	Packet Captures		Agency-Defined		
	Application Logs	Data Aggregation	None		
	Transaction Logs		Multi-Account		
			Multi-Region		
			Multi-Provider		
		Data Transformation	None (Native Forms Align)		
			IPFIX		
			Other (CISA Coordinated)		

Pattern Characteristics

Cloud Telemetry Timeliness

For this pattern, factors affecting the timeliness of information include the polling frequency of CLAW. The frequency with which CLAW checks for and pulls new data adds delay and is limited by mechanisms such as API request throttling.

Agency processing itself should not significantly affect timeliness; aggregating a common log type from multiple sources is expected to be a low complexity operation, facilitating rapid execution.

Cloud Telemetry Timing Coordination

In this case, the processing stage will have an opportunity to introduce its own timestamps into the overall chain that originates from the source and terminates at the CLAW. However, the service transaction logs are still timestamped when the log entry is generated at the CSP. Additional timestamps

[6] That is, an authenticated identity principal corresponding to the CLAW instance in the selected region.

may be added at the time when the log entries are aggregated. However, the original log entries' timestamps should be preserved.

Cloud Telemetry Provenance

This pattern involves agency processing on multiple log streams across multiple tenancies. As a common log type and format is assumed across each data source, the agency is able to aggregate the sources either by interleaving or combining them in some other fashion (e.g., data from one tenancy might precede that from another). In this case, provenance claims are likely to be made by the agency. In particular, although multiple streams may arrive at the agency labeled and integrity-protected, the process of interleaving would create a new stream that itself requires provenance metadata. In short, the agency would be responsible for asserting that it provided the aggregation of the multiple streams, and constituent streams may retain sufficient provenance information to be checked end-to-end by CISA because in this pattern no agency filtration is performed. In addition, as the agency is not necessarily guaranteed to receive incoming telemetry requests at a predetermined rate, the agency may need to decide which data to retain or discard. Should it be necessary for the agency to discard data, this fact should be noted and integrity protected as part of the ordinary provenance processing.

3.5 SA-SDNN-SS: Agency Filtered Data Push to CLAW

Overview

In this reporting pattern, CSP refers to an agency's SaaS cloud vendor. The CSP provides various application services, such as customer relations, email, or support service delivery tracking. These services each generate telemetry that may contain sensitive information; the agency gathers the telemetry and then filters and enriches it before sending to CLAW.

Figure 11 (below) shows the roles and telemetry flow associated with this reporting pattern. With regard to roles, the CSP is responsible for generating and delivering data, the Agency is responsible for configuring the CSP and filtering data, and CISA is responsible for receiving data from the Agency. With regard to telemetry flow, the CSP generates telemetry from agency applications in Stage A (Cloud Sensing), the Agency filters telemetry received from the CSP in Stage B (Agency Processing), and the Agency pushes telemetry to CLAW in Stage C (Reporting to CISA).

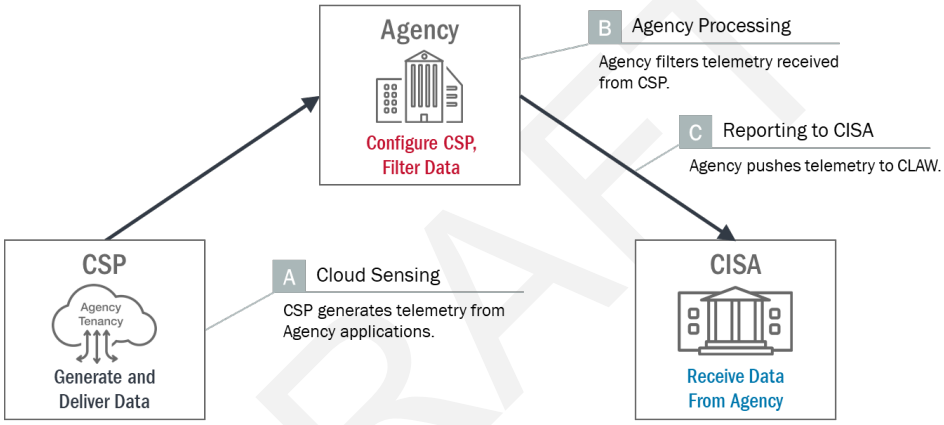


Figure 11: Roles and Telemetry Flow – SA-SDNN-SS

Stages

Figure 12 (below) shows the events that take place during each of this reporting pattern's three stages. A detailed description of each stage is presented below:

Stage A: Network traffic between the agency's cloud tenancy and the Internet is handled by the CSP services. The agency configures one or more of these services to generate application logs.

Stage B: The agency uses its NOC/SOC tools to perform data sanitization and enrichment functions to process the raw data. The raw data may be filtered to remove agency "private/internal" sources, personally identifiable information (PII), or other sensitive information in conformance with agency sanitization and sharing requirements. The agency then enriches some fields with derived information (e.g., destination country). The agency does not perform any aggregation or transformation.

Stage C: The agency pushes the processed telemetry to CLAW in a single region. The exact delivery mechanism(s) depends on the CSP.

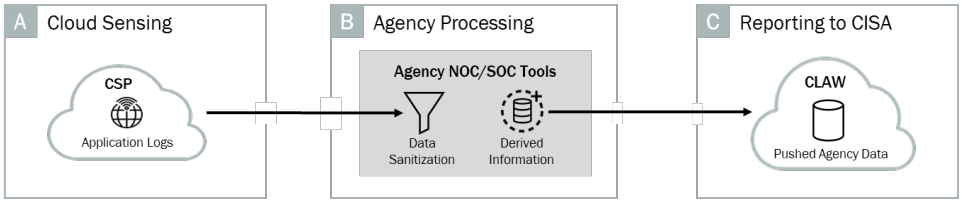


Figure 12: Visual Pattern Summary – SA-SDNN-SS

Pattern Summary

Table 6 (below) lists the option that is associated with each attribute in this reporting pattern.

Table 6: Pattern Summary Table – SA-SDNN-SS

Stage A – Cloud Sensing		Stage B – Agency Processing		Stage C – Reporting to CISA	
Attribute	**Options**	**Attribute**	**Options**	**Attribute**	**Options**
Sensor Positioning	Gateway	Data Filtering	None	Data Transfer	Agency Push
	Subnet		Removal		CLAW Pull
	Interface		Sanitization	CLAW Distribution	Single Region
	Service		Obfuscation		Multi-Region
	Application	Data Enrichment	None		Multi-Cloud
Telemetry Types	Network Flow Logs		Derived		
	Packet Captures		Agency-Defined		
	Application Logs	Data Aggregation	None		
	Transaction Logs		Multi-Account		
			Multi-Region		
			Multi-Provider		
		Data Transformation	None (Native Forms Align)		
			IPFIX		
			Other (CISA Coordinated)		

Pattern Characteristics

Cloud Telemetry Timeliness

For this pattern, factors affecting the timeliness of information include the application and the agency's own policy for delivery to CLAW. Agencies should consult application-specific documentation and determine which fields might have sensitive information that is not trivial to detect and remove. Agencies can delay delivering individual records/objects (e.g., as part of a batching policy) and may do so if they do not exceed the maximum delay parameters.

Agency processing may significantly affect timeliness. Some log fields may be sanitized by withholding them while others may require deep scanning (e.g., PII embedded in a Uniform Resource Locator (URL) field). Agencies should also characterize the performance of different methods of cross-referencing the relevant data for enrichment.

Cloud Telemetry Timing Coordination

In this case, the processing stage will have an opportunity to introduce its own timestamps into the overall chain that originates from the source and terminates at the CLAW. However, the service application logs are still timestamped when the log entry is generated at the CSP. Additional timestamps may be added at the time when the log entries are processed. However, the original log entries' timestamps should be preserved.

Cloud Telemetry Provenance

This pattern involves agency processing on log content, including data removal and addition. In this case, the agency is an author of log information, as it is providing enrichment and editing. Provenance claims in this context are three-fold: the origin of the information from the SaaS service, the origin of the information used in performing the enrichment, and resulting stream provided to CISA by the agency. Agency processing should be arranged to convey both the nature of the modifications (e.g., enrichment) performed, the type of information removed, and the processing mechanisms (e.g., software artifacts) used in performing the processing.

3.6 NN-SDNI-LS: CLAW Pull of Agency Filtered Data

Overview

In this reporting pattern, CSP refers to an agency's IaaS cloud vendor. The agency configures sensors for specific subnets within its cloud tenancy. These sensors monitor the agency traffic to/from those subnets and generate network flow logs, which are processed extensively by the agency prior to being retrieved by CLAW. The agency processing is done at a single location, so retrieval by a single region of CLAW is utilized.

Figure 13 (below) shows the roles and telemetry flow associated with this reporting pattern. With regard to roles, the CSP is responsible for generating and delivering data, the Agency is responsible for configuring the CSP and filtering data, and CISA is responsible for retrieving data from the Agency. With regard to telemetry flow, the CSP generates telemetry from agency traffic in Stage A (Cloud Sensing), the Agency filters telemetry received from the CSP in Stage B (Agency Processing), and CLAW pulls telemetry from the Agency in Stage C (Reporting to CISA).

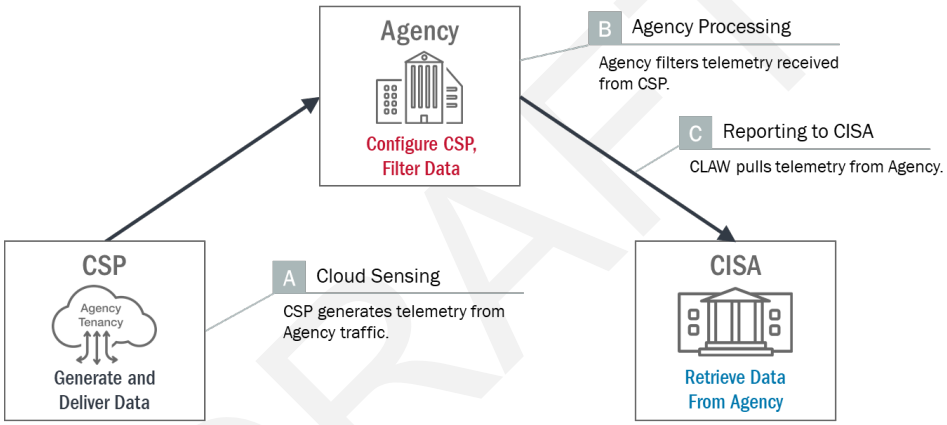


Figure 13: Roles and Telemetry Flow – NN-SDNI-LS

Stages

Figure 14 (below) shows the events that take place during each of this reporting pattern's three stages. A detailed description of each stage is presented below:

Stage A: Network traffic to and from the agency's chosen[7] subnets pass through the CSP's sensors, where security functions (e.g., firewall) are implemented. In addition to executing their security functions, these subnet-level sensors also generate network flow logs.

Stage B: The agency uses its NOC/Soc tools to perform data sanitization, enrichment, and transformation functions to process the raw data. The raw data is filtered to remove agency "private/internal" sources, PII, and other sensitive information in conformance with agency sanitization requirements. The agency may perform filtering before or after other processing. The data is also transformed to the Internet Protocol Flow Information Export (IPFIX) format (although CLAW is likely capable of ingesting the data in the CSP's native format, the agency may prefer IPFIX for its own

[7] One possible selection of subnets consists of just those publicly accessible from the Internet; this allows the agency to filter out much of the data corresponding to "private/internal" sources even before the Agency Processing stage. CISA is primarily interested in this, as opposed to private traffic between internal components.

analytics). The agency enriches some fields with derived information (e.g., destination country) in the IPFIX format.[8] The agency does not perform any aggregation.

Stage C: The agency supplies the filtered telemetry to be pulled by CLAW in a single region. This involves configuring pull credentials such that CLAW[9] is authorized to make the necessary requests. The exact delivery mechanism(s) depends on the CSP.

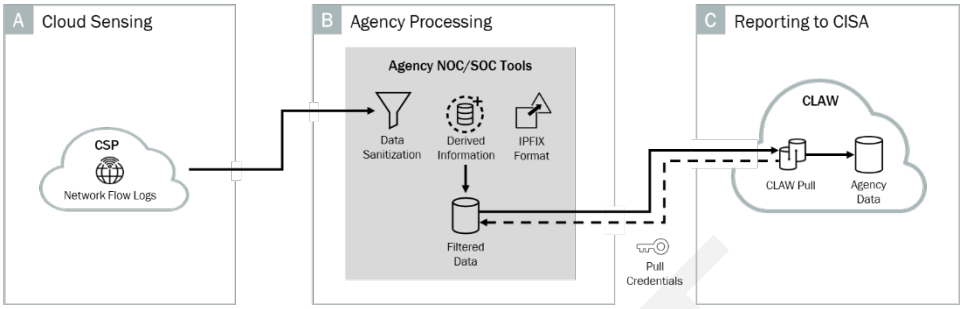


Figure 14: Visual Pattern Summary – NN-SDNI-LS

Pattern Summary

Table 7 (below) lists the option that is associated with each attribute in this reporting pattern.

Table 7: Pattern Summary Table – NN-SDNI-LS

Stage A – Cloud Sensing		Stage B – Agency Processing		Stage C – Reporting to CISA	
Attribute	**Options**	**Attribute**	**Options**	**Attribute**	**Options**
Sensor Positioning	Gateway	Data Filtering	None	Data Transfer	Agency Push
	Subnet		Removal		CLAW Pull
	Interface		Sanitization	CLAW Distribution	Single Region
	Service		Obfuscation		Multi-Region
	Application	Data Enrichment	None		Multi-Cloud
Telemetry Types	Network Flow Logs		Derived		
	Packet Captures		Agency-Defined		
	Application Logs	Data Aggregation	None		
	Transaction Logs		Multi-Account		
			Multi-Region		
			Multi-Provider		
		Data Transformation	None (Native Forms Align)		
			IPFIX		
			Other (CISA Coordinated)		

Pattern Characteristics

Cloud Telemetry Timeliness

For this pattern, factors affecting the timeliness of information include the aggregation interval for network flow logs and the polling frequency of CLAW. As (successful) network flows are not point events, when they "occur" is partly determined by the aggregation interval; shorter intervals trade quicker visibility for higher log volume (and vice versa). Tenants have some control over this interval (depending on the CSP). The frequency with which CLAW checks for and pulls new data adds delay and is limited by mechanisms such as API request throttling.

[8] Potentially in the form of enterprise-specific information elements.
[9] That is, an authenticated identity principal corresponding to the CLAW instance in the selected region.

Agency processing may significantly affect timeliness. As there is more extensive processing than in other patterns, agencies should test and document the end-to-end processing time for logs, ideally under realistic workloads.

Cloud Telemetry Timing Coordination
In this case, the processing stage will have an opportunity to introduce its own timestamps into the overall chain that originates from the source and terminates at the CLAW. However, the service application logs are still timestamped when the log entry is generated at the CSP. Additional timestamps may be added at the time when the log entries are processed. However, the original log entries' timestamps should be preserved.

Cloud Telemetry Provenance
This pattern involves potentially significant processing on log content by the agency, including data removal, data transformation, and enrichment. Agency processing should be arranged to convey both the nature of the modifications (e.g., enrichment) performed, the type of information removed, and the processing mechanisms (e.g., software artifacts) used in performing the processing. In this case, the agency is an author of log information or metadata. Provenance claims in this context are three-fold: the origin of the information from the IaaS service, the origin of the information used in performing the enrichment, and the resulting stream provided to CISA by the agency. The report stream has undergone a transformation so the nature and entity author(s) of the transformations should be captured in the provenance claims.

3.7 SN-NDNC-SS: Agency CSP SECaaS Data Push to CLAW

Overview

In this reporting pattern, the CSP provides Security-as-a-Service (SECaaS) to the agency. In the SECaaS model, the sensors that generate telemetry are managed by the CSP and configured by the agency. This reporting pattern outlines a basic case where telemetry generated by the CSP is delivered directly to CLAW.

Figure 15 (below) shows the roles and telemetry flow associated with this reporting pattern. With regard to roles, the CSP is responsible for generating and delivering data, the Agency is responsible for configuring the CSP, and CISA is responsible for receiving data from the CSP. With regard to telemetry flow, the CSP generates telemetry from agency traffic in Stage A (Cloud Sensing), the CSP processes telemetry based on agency settings in Stage B (Agency Processing), and the CSP pushes telemetry to CLAW in Stage C (Reporting to CISA).

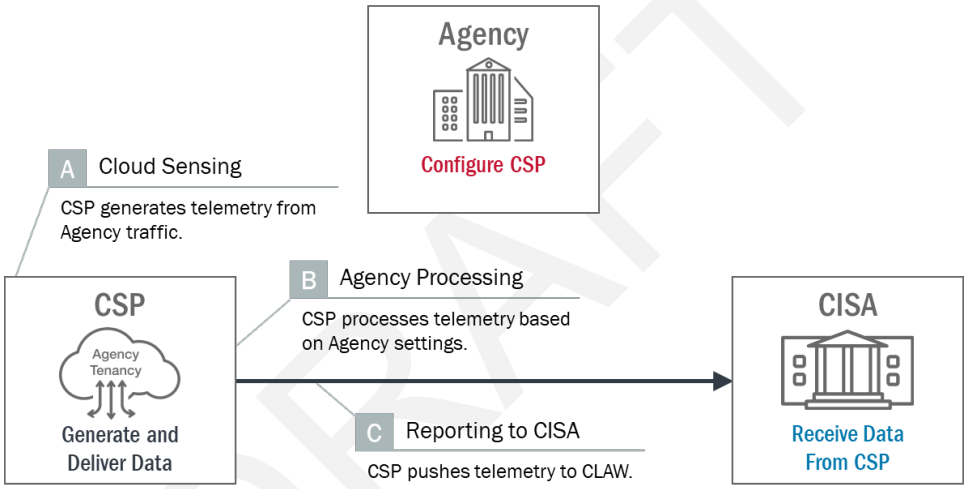


Figure 15: Roles and Telemetry Flow – SN-NDNC-SS

Stages

Figure 16 (below) shows the events that take place during each of this reporting pattern's three stages. A detailed description of each stage is presented below:

Stage A: Network traffic between the agency and the Internet is routed through the CSP's services, where various security functions are implemented, which may include firewall, DDoS protection, or web filtering. These services can generate different telemetry types depending on the CSP and services used. In this reporting pattern, the agency configures the CSP to generate network flow logs.

Stage B: The agency configures the CSP service using NOC/SOC tools. The agency does not configure any filtering[10] but does configure enrichment and transformation. The agency configures the CSP option to include some enrichment fields with derived information (e.g., destination country). Transformation is

[10] Similar to pattern NN-SDNI-LS, the agency can configure the routing so that only traffic between public-facing components and the Internet is routed through the CSP's sensors, removing one of the common drivers of filtering.

necessary as the agency and CISA coordinated a format for the telemetry that differs from the CSP's native format.[11] The agency does not perform any aggregation.

Stage C: The agency configures their telemetry to be pushed from the NOC/SOC tools to CLAW in a single region. The exact delivery mechanism(s) depends on the CSP; while coordinating on the telemetry format, the agency and CISA also work together to ensure that CLAW is capable of directly receiving telemetry from the third-party.

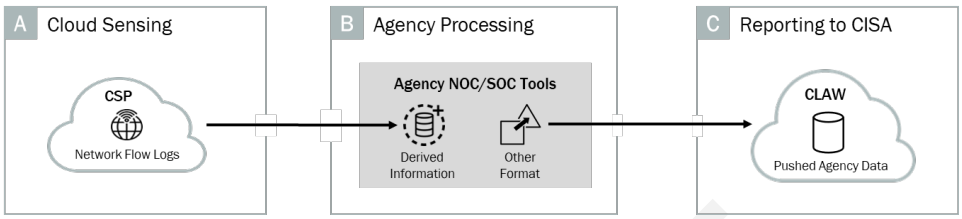


Figure 16: Visual Pattern Summary – SN-NDNC-SS

Pattern Summary

Table 8 (below) lists the option that is associated with each attribute in this reporting pattern.

Table 8: Pattern Summary Table – SN-NDNC-SS

Stage A – Cloud Sensing		Stage B – Agency Processing		Stage C – Reporting to CISA	
Attribute	**Options**	**Attribute**	**Options**	**Attribute**	**Options**
Sensor Positioning	Gateway	Data Filtering	**None**	Data Transfer	**Agency Push**
	Subnet		Removal		CLAW Pull
	Interface		Sanitization	CLAW Distribution	**Single Region**
	Service		Obfuscation		Multi-Region
	Application	Data Enrichment	None		Multi-Cloud
Telemetry Types	**Network Flow Logs**		**Derived**		
	Packet Captures		Agency-Defined		
	Application Logs	Data Aggregation	**None**		
	Transaction Logs		Multi-Account		
			Multi-Region		
			Multi-Provider		
		Data Transformation	None (Native Forms Align)		
			IPFIX		
			Other (CISA Coordinated)		

Pattern Characteristics

Cloud Telemetry Timeliness

For this pattern, factors affecting the timeliness of information include the aggregation interval for network flow logs and the agency's own policy for delivery to CLAW. As (successful) network flows are not point events, when they "occur" is partly determined by the aggregation interval; shorter intervals trade quicker visibility for higher log volume (and vice versa). Tenants have some control over this interval (depending on the CSP). Agencies can delay delivering individual records/objects (e.g., as part of a batching policy) and may do so if they do not exceed the maximum delay.

[11] CISA may request modifications to the CSP's default format in order to include required information or to improve ingestion processing. Once CISA decides on a format for a given vendor and service, subsequent agencies who use the same CSP/service may use it as a standard.

Agency processing may significantly affect timeliness. Agencies should characterize the performance of different methods of cross-referencing the relevant data for enrichment.

Cloud Telemetry Timing Coordination
In this case, the processing stage will have an opportunity to introduce its own timestamps into the overall chain that originates from the source and terminates at the CLAW. However, the service application logs are still timestamped when the log entry is generated at the CSP. Additional timestamps may be added at the time when the log entries are processed. However, the original log entries' timestamps should be preserved.

Cloud Telemetry Provenance
This pattern involves the agency applying processing to transform data from a SECaaS vendor format to a CISA-acceptable format, including potential data enrichment. In this case, the agency is the primary author of log information. Provenance claims in this context are three-fold: the origin of the information from the SECaaS service, the origin of the information used in performing the enrichment, and the information regarding the resulting stream provided to CISA and authored by the agency. The stream is being freshly authored based on information provided by enrichment and the SECaaS provider and is not limited to simple transformations. In this pattern, agency processing should be arranged to convey both the nature of the original sources, the processing mechanisms (e.g., software artifacts) used in performing the processing, and an indicator of the agreement between the agency and CISA governing the streams provided.

3.8 ST-SANC-SS: CSP SECaaS Data Push to Agency, Agency Processing and Push Data to CLAW

Overview

In this reporting pattern, the CSP provides Security-as-a-Service (SECaaS) to the agency. Telemetry generated by the CSP is sent to the agency, which processes the data prior to sending it to CLAW. Agencies may choose this pattern if the processing functions provided by the CSP sensors are insufficient.

Figure 17 (below) shows the roles and telemetry flow associated with this reporting pattern. With regard to roles, the CSP is responsible for generating and delivering data, the Agency is responsible for configuring the CSP and processing data, and CISA is responsible for receiving data from the Agency. With regard to telemetry flow, the CSP generates telemetry from agency traffic in Stage A (Cloud Sensing), the Agency processes telemetry received from the CSP in Stage B (Agency Processing), and the Agency pushes telemetry to CLAW in Stage C (Reporting to CISA).

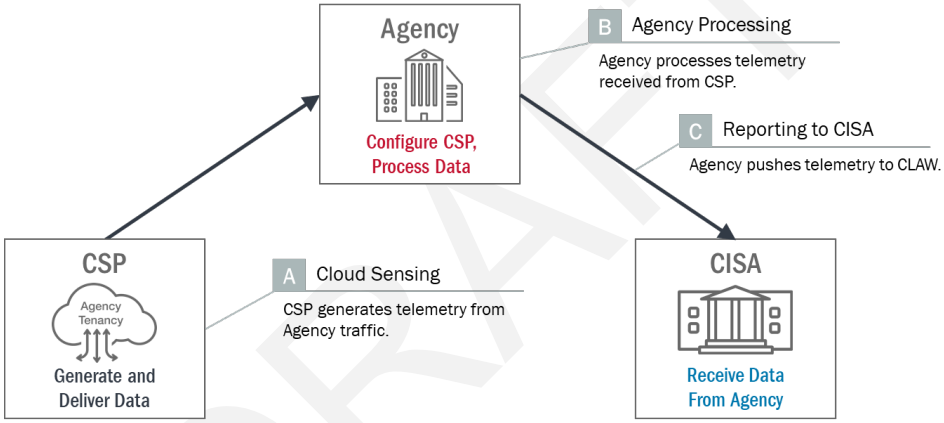


Figure 17: Roles and Telemetry Flow – ST-SANC-SS

Stages

Figure 18 (below) shows the events that take place during each of this reporting pattern's three stages. A detailed description of each stage is presented below:

Stage A: The agency configures the CSP services to generate network flow logs and application logs.

Stage B: The agency performs data filtering, enrichment, and transformation, first by using functions provided by the CSP services (to perform data sanitization and enrichment) and then by using their own NOC/SOC tools (to perform further data sanitization and enrichment, as well as transformation). Factors for selection of where processing occurs include performance, cost, and privacy. The agency processing may include capabilities implemented through self-hosted services or from an agency's cloud telemetry processing service. The agency uses the CSP's service capabilities to pre-process the telemetry to include certain enrichment fields with agency-defined information and exclude certain fields with sensitive information that should not be shared (i.e., data sanitization). Optionally, the agency may also configure the CSP services to output the telemetry in an intermediate format convenient for its own processing. After the processing at the CSP service, the telemetry is delivered to the agency for additional processing. For example, the agency further sanitizes web transaction telemetry by scanning for sensitive data embedded within the URL field, and further enriches firewall transaction logs with

agency-defined data (such as labels identifying resources within their cloud tenancy). As final processing, the agency transforms the data into a format agreed-upon with CISA. No data aggregation is performed.

Stage C: The agency pushes the processed logs to CLAW in a single region.

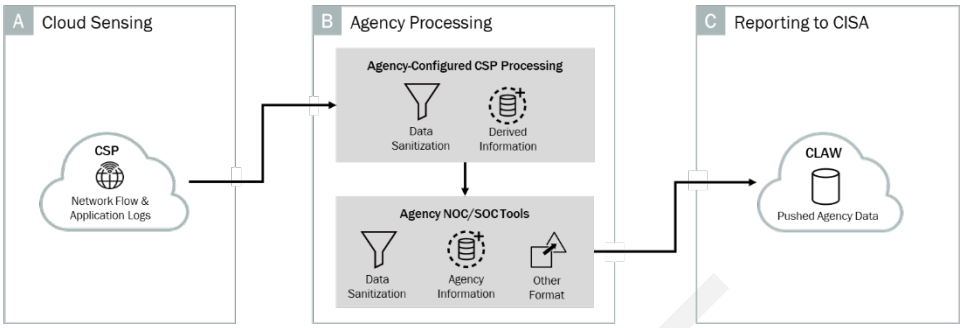


Figure 18: Visual Pattern Summary – ST-SANC-SS

Pattern Summary

Table 9 (below) lists the option that is associated with each attribute in this reporting pattern.

Table 9: Pattern Summary Table – ST-SANC-SS

Stage A – Cloud Sensing		Stage B – Agency Processing		Stage C – Reporting to CISA	
Attribute	**Options**	**Attribute**	**Options**	**Attribute**	**Options**
Sensor Positioning	Gateway	Data Filtering	None	Data Transfer	Agency Push
	Subnet		Removal		CLAW Pull
	Interface		Sanitization	CLAW Distribution	Single Region
	Service		Obfuscation		Multi-Region
	Application	Data Enrichment	None		Multi-Cloud
Telemetry Types	Network Flow Logs		Derived		
	Packet Captures		Agency-Defined		
	Application Logs	Data Aggregation	None		
	Transaction Logs		Multi-Account		
			Multi-Region		
			Multi-Provider		
		Data Transformation	None (Native Forms Align)		
			IPFIX		
			Other (CISA Coordinated)		

Pattern Characteristics

Cloud Telemetry Timeliness

For this pattern, factors affecting the timeliness of information include the agency's own policy for delivery to CLAW. Agencies can delay delivering individual records/objects (e.g., as part of a batching policy) and may do so if they do not exceed the maximum delay parameters.

Agency processing may significantly affect timeliness. As there is more extensive processing than in other patterns, agencies should test and document the end-to-end processing time for logs, ideally under realistic workloads. However, it is expected that the CSP can perform any pre-processing configured by the agency in real-time.

Cloud Telemetry Timing Coordination
In this case, the processing stage will have an opportunity to introduce its own timestamps into the overall chain that originates from the source and terminates at the CLAW. However, the service application logs are still timestamped when the log entry is generated at the CSP. Additional timestamps may be added at the time when the log entries are processed. However, the original log entries' timestamps should be preserved.

Cloud Telemetry Provenance
This pattern involves the agency applying processing to transform data from a SECaaS vendor format to a CISA-acceptable format, along with arbitrary data transformations, filtration, and enrichment decided by the agency. In this case, the agency is the primary author of log information. Provenance claims in this context are multiple (depending on the complexity of the agency processing performed) but include: the origin of the information from the SECaaS and other services, the origin of the information used in performing the enrichment, and information regarding the resulting stream provided to CISA and authored by the agency. The stream is being freshly authored based on information provided by enrichment and the SECaaS provider and may involve nearly arbitrary data processing. In this pattern, agency processing should be arranged to convey the provenance of all original sources, all processing mechanisms (e.g., software artifacts and services) used in performing the processing, and an indicator of the agreement between the agency and CISA demonstrating how the stream provided to CISA is sufficient for NCPS operations.

4 COMBINATION REPORTING PATTERNS

A combination reporting pattern is when two or more existing reporting patterns are selected to be applied in concert. Combination patterns tend to arise when there are multiple sources of raw telemetry and one reporting pattern is not appropriate for all of them. As with Section 3, this document will only focus on a small set of possible combinations. Combinations not shown here may still be viable alternatives and should be discussed with CISA on a case-by-case basis.

A short description is provided for each combination reporting pattern, along with pros, cons, and alternatives to guide characteristics. For brevity, familiarity with Section 3 is assumed and discussion about the attributes and options of each constituent pattern is omitted.

Combination Reporting Pattern Characteristics

Cloud Telemetry Timeliness
The combination reporting patterns mix various details of the previously discussed timeliness characteristic. Agencies should not expect or try to achieve "uniform" timeliness from all sources but instead make sure that the delay from event occurrence to delivery to CLAW is reasonable in all cases; this will require extensive testing.

Cloud Telemetry Timing Coordination
In the case of combination reporting patterns, the processing stage will have an opportunity to introduce its own timestamps into the overall chain that originates from the source and terminates at the CLAW. However, the cloud telemetry logs are still timestamped when they are generated at the CSP. Additional timestamps may be added at the time when the log entries are processed. However, the original log entries' timestamps should be preserved.

Cloud Telemetry Provenance
The combination reporting patterns mix various details of the other generic reporting patterns and consequently the provenance concerns vary depending on the specific details. The particular scenarios may be more complex as provenance from different types of systems (e.g., SaaS, IaaS) and locations or administrative controls may be interleaved, each with different levels of abstraction and granularity or capabilities of reporting (e.g., time, identity). In cases where multiple different log types can be aggregated and processed, a common field is typically used to correlate information. A timestamp or transaction identifier is commonly used; note that time should be of sufficient precision and accuracy to make such log aggregation possible.[12]

[12] See, for example, minimum requirements for 1msec granularity in the financial industry (Consolidated Audit Trail NMS).

4.1 Differentiated Processing of Multi-Account Data (GT-NNAN-SS + SN-NNNN-LS)

Description

This combination pattern is for agencies that have multiple accounts and where telemetry from different accounts may have different characteristics. Telemetry from Accounts 1/2 are handled as in pattern GT-NNAN-SS, with some additional processing and data from multiple accounts is aggregated prior to delivery to CLAW. Telemetry from Account 3 is handled independently, pulled directly from the CSP by CLAW (just as in pattern SN-NNNN-LS). This approach can support various use cases, such as bypassing agency processing for streams that do not require it (e.g., no sanitization required for Account 3), or for sending multiple streams to CLAW based on sub-organizations within the agency (e.g., one group owns Accounts 1/2 and another owns Account 3).

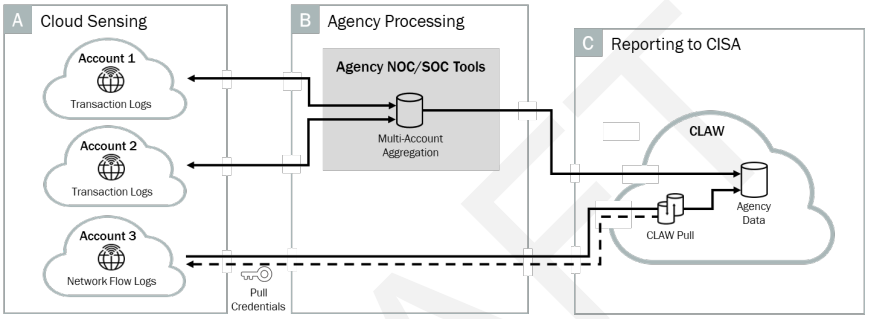


Figure 19: Visual Pattern Summary – Differentiated Processing of Multi-Account Data

Pattern Summary

Table 10: Pattern Summary Table – Differentiated Processing of Multi-Account Data

Stage	Attribute	Account 1 Option	Account 2 Option	Account 3 Option
Stage A: Sensing	Sensor Positioning	Gateway	Gateway	Subnet
	Telemetry Types	Transaction Logs	Transaction Logs	Network Flow Logs
Stage B: Agency Processing	Data Filtering	None	None	None
	Data Enrichment	None	None	None
	Data Aggregation	Multi-Account	Multi-Account	None
	Data Transformation	None	None	None
Stage C: Reporting to CISA	Data Transfer	Agency Push	Agency Push	CLAW Pull
	CLAW Distribution	Single region	Single region	Single region

Pros

- Different input streams are handled naturally according to their needs.
- A "sub-agency" can be assigned to each output stream sent to CLAW, allowing CISA to conduct both whole-agency and more granular analysis.
- Issues pushing Account 1/2 data to CLAW do not necessarily affect CLAW's ability to pull Account 3 data.

Cons

- Account-level granularity may not be enough when differentiating streams.
- Multiple groups may be responsible for sending data to CLAW.
- Without additional configuration, the agency Network Operations Center / Security Operations Center does not have visibility into Account 3.

Alternatives

In the simplest alternative, Account 3 telemetry is aggregated along with Account 1/2 data, reducing this combination pattern into a variant of pattern GT-NNAN-SS. This approach largely inverts the pros/cons listed above.

In another alternative, Account 3 telemetry undergoes a separate and minimal processing pipeline, resulting in a push to CLAW independent of the Account 1/2 telemetry. This approach alleviates some of the cons listed above but results in additional complexity in the Agency Processing stage.

4.2 Per-Region Processing of Multi-Region Data

Description

This combination pattern is for agencies using a single CSP in multiple regions. Like pattern SA-SDNN-SS, the agency has logs that they sanitize prior to delivery to CLAW. This time, the logs are network flow logs and originate from two different regions, which the agency handles entirely in-region; they provision identical processing pipelines in both regions and send the output of each to the "local" CLAW (i.e., the instance of CLAW in the same CSP and region). In other words, two instances of pattern GA-SDNN-SS are combined to handle two regions. This approach can be generalized to any number of regions and can be applied in any instance where similar telemetry is generated in multiple regions.

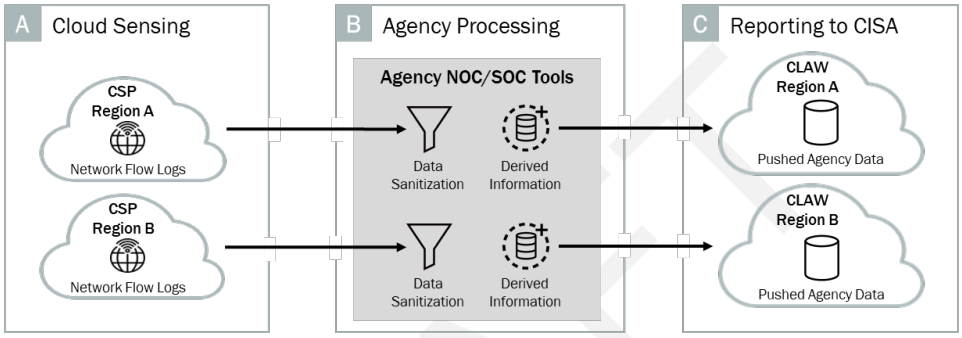


Figure 20: Visual Pattern Summary – Per-Region Processing of Multi-Region Data

Pattern Summary

Table 11: Pattern Summary Table – Per-Region Processing of Multi-Region Data

Stage	Attribute	Region A Option	Region B Option
Stage A: Sensing	Sensor Positioning	Gateway	Gateway
	Telemetry Types	Network Flow Logs	Network Flow Logs
Stage B: Agency Processing	Data Filtering	Sanitization	Sanitization
	Data Enrichment	Derived	Derived
	Data Aggregation	None	None
	Data Transformation	None	None
Stage C: Reporting to CISA	Data Transfer	Agency Push	Agency Push
	CLAW Distribution	Multi-Region	Multi-Region

Pros

- Data is kept within one region, minimizing data transfer costs.
- Infrastructure-as-code services can be used so the agency only implements a pipeline template once.
- Issues in one pipeline do not necessarily affect others.

Cons

- Cost of operating multiple pipelines may exceed the cost of a single pipeline capable of handling all the data.
- In the absence of infrastructure-as-code services, changes need to be applied independently to each pipeline.
- A local CLAW may not be present in each region where telemetry is generated.

Alternatives

Agencies may instead conduct multi-region aggregation to produce a single stream of data, processed by a single pipeline and delivered to a single CLAW. This approach largely inverts the pros/cons listed above.

4.3 Push from Integrated Sharing Solution

Description

In the integrated sharing solution, an agency is already performing robust cloud telemetry processing and is extending the output of their tools to now include reporting to CISA via CLAW. This pattern takes an "all of the above" approach to the breadth of input and processing. Input sources may include telemetry from the local CSP, other CSPs, on-premise analytics, mobile device management systems, and CSP or third-party threat intelligence. The cloud sensing may include multiple CSP sensor positions with multiple telemetry types. The resulting information is aggregated together with other (possibly non-security) information for subsequent filtration, enrichment, transformation, and export as selected by the agency. CISA is one consumer; others may include the agency's own risk management, security, and operational personnel.

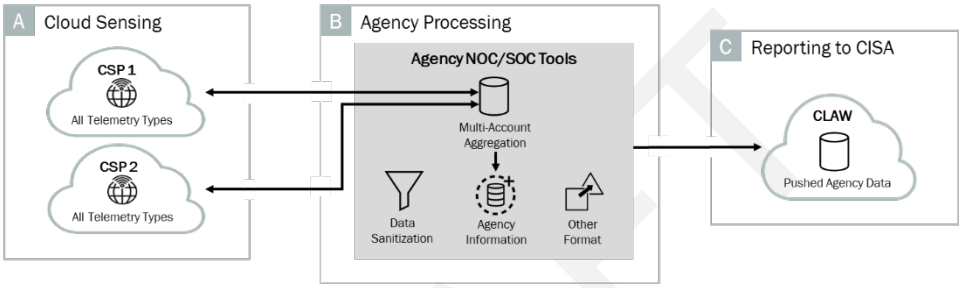


Figure 21: Visual Pattern Summary – Push from Integrated Sharing Solution

Pattern Summary

Table 12: Pattern Summary Table – Push from Integrated Sharing Solution

Stage	Attribute	Tenancy Options
Stage A: Sensing	Sensor Positioning	Gateway, Subnet, Interface, Service, Application
	Telemetry Types	Network Flow Logs, Packet Captures, Application Logs, Transaction Logs
Stage B: Agency Processing	Data Filtering	Removal, Sanitization, Obfuscation
	Data Enrichment	Derived, Agency-Defined
	Data Aggregation	Multi-Account, Multi-Region, Multi-Provider
	Data Transformation	CISA Coordinated
Stage C: Reporting to CISA	Data Transfer	Agency Push
	CLAW Distribution	Multi-Region, Multi-Cloud

Pros

- Visibility is broad due to multiple input streams.
- Leverages existing agency capabilities and integration.
- Simplified CLAW attribution and coordination, as all telemetry for the protected entity is originating from a single source system.

Cons

- Complex mechanisms to ensure unique identifiers for all physical and logical resources in both on-premise (if applicable) and cloud environments including computing resources, person and non-person accounts, and IP addressable infrastructure components.
- Complex Data Model and Reporting Architecture vs. Cloud Native Telemetry and Visualization capabilities by identifying the means to accommodate inherent differences in underlying data types and attributes between cloud, on-premise (if applicable), and CISA/CLAW environments.
- Complex requirements and supported capabilities for ingestion of information from other CSPs.
- Complex dashboards for all parties concerned with the ongoing delivery of telemetry (e.g., agency/CSP and CISA) also require periodic cross verification for accuracy and adequacy.

Alternatives

Agencies may instead determine the CLAW telemetry sharing requirements align with an existing output consumer, permitting reuse.

5 CONCLUSION

As agencies move more of their applications and services to cloud, the NCPS Program is evolving to ensure that security information for cloud-based traffic can be captured and analyzed and that CISA analysts can continue to provide situational awareness and support to the agencies. The *NCPS Cloud Interface Reference Architecture: Volume One* document introduces a framework for developing reporting patterns for how cloud logs will be collected and transferred to CLAW. This companion document (*NCPS Cloud Interface Reference Architecture: Volume Two*) provides a catalog of generic reporting patterns that match common agency cloud use cases and shows how more complex reporting patterns can be developed to describe use cases with a combination of attributes and options concurrently. Together, these two documents provide guidance for how an agency can adapt their cloud environments to allow for security data to be sent to CLAW.

Individual CSPs can use these documents to provide vendor solutions that match reporting patterns. Vendors are encouraged to develop overlays that identify how their agency customers can comply with EINSTEIN visibility requirements while using the CSP's products and services. While CISA will not provide formal authorization or approval of a vendor overlay solution, CISA may provide input to the vendor on a case-by-case basis to convey desired approaches and intent.

APPENDIX A: CLOUD TELEMETRY TIMELINESS

Different CSPs have different timeframes for log delivery. While typical values range between a few minutes to fifteen minutes of event occurrence, agencies must confirm the timeliness of a CSP's log delivery through discussions with the CSP and their own testing. In most cases, neither the service documentation nor the published Service Level Agreements make a concrete statement regarding the timeliness of log delivery. While one CSP might claim that "events are delivered within five minutes of occurrence," another might claim that "events are delivered in real-time," and another might only provide hints via screenshots. Even within a CSPs offerings, more common/popular services are likely to have better documentation around timeliness than other services.

Some generalizations may be made based on log type. Logs concerning point-in-time events (e.g., transaction logs for auditing API calls) can be delivered quickly, whereas those concerning continuous events (e.g., network flow logs or application metrics detailing resource usage) have some interval that must transpire before the event is recorded and delivered. In the latter case, tenants may be given some control over the interval, with the caveat that shorter intervals incur greater costs than the default/free interval.

CSPs may tailor their log delivery based on the destination. For example, a CSP may do hourly batching to its general purpose storage destination, within-minutes delivery to its big data service, and real-time streaming to its publish/subscribe service.[13] If agencies perform processing before delivery to CLAW, they must recognize when this behavior is present in a CSP and provide a receiving destination that allows timely receipt of raw logs. If logs are pushed directly to CLAW, then CISA will provide an appropriate receiving destination for the CSP service.

A complement to log timeliness is log *completeness*; data that never arrives is not timely at all. CSPs may not provide complete logs for several reasons, some intentional and some not.

- The methodology for generating data is based on sampling (e.g., only a sample of network packets being used for network flow logs, values for application metrics only being sampled at certain intervals, etc.).
- Events occur faster than the CSP can log them. This is more likely for data plane events (e.g., HTTP GET requests to a public storage container) than for management plane events (e.g., API calls that change the configuration of the storage container).
- Misconfiguration results in some events being seen but not logged (e.g., different audit settings based on roles) or sensors being placed such that events are not seen at all (e.g., network sensor placement relative to a gateway/firewall).
- Log comprehensiveness is related to log completeness; for some services, tenants may change the default settings to allow for more detailed or less detailed reporting.

To be timely, data must be received in a cyber-relevant timeframe. However, what is considered a cyber-relevant timeframe varies depending on threats. Recent open-source reporting[14] has measured the "breakout time" of several well-known threat actors, where breakout time is defined as the time from initial compromise to the start of lateral movement (including steps such as local network

[13] https://en.wikipedia.org/wiki/Publish%E2%80%93subscribe_pattern
[14] https://www.crowdstrike.com/resources/reports/2019-crowdstrike-global-threat-report/

reconnaissance and privilege escalation on the compromised host). Effective action within this time can stop an attack in its early stages. While most threat actors had an average breakout time of over two hours, Russian actors were found to have an average breakout time of under twenty minutes. This is significantly faster than what many organizations are prepared to handle.

CISA's goal is to detect, investigate, and respond to any threat before it has time to evolve and progress. Although CISA acknowledges that an agency has limited control over the timeliness of a CSP's delivery of raw logs, once the logs are received from the CSP, it is the agency that largely determines how long it takes to process the logs and deliver them to CLAW. Agencies should ensure that the time between raw logs release to the agency tenant from the CSP and the delivery of the processed logs to CLAW is within 30 minutes.

CISA Preference

When agency processing is performed, CISA expects that the time between receiving raw logs from the CSP and the delivery of processed logs to CLAW does not exceed 30 minutes.

APPENDIX B: CLOUD TELEMETRY TIMING COORDINATION

All systems that consist of multiple servers and clients (such as in cloud service delivery models) require timing synchronization. Timing synchronization is an important issue that must have its own considerations based on the level of accuracy specified by the individual D/As. This concept applies to CLAW. The need for time synchronization is dictated by the three reporting pattern stages and the individual servers and clients that are an integral part of CLAW.

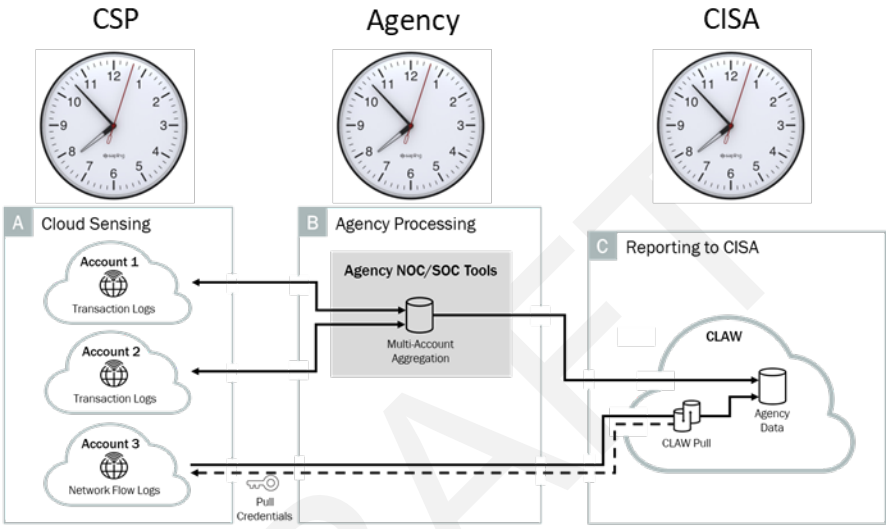


Figure 22: Typical Organizations Involved In A CLAW Reporting Transaction

As indicated above, synchronization is desired across the various geographic regions on the CSP side that participate in the generation of the cloud logs stored in CLAW. In addition, it is important to ensure that the systems of the individual agencies are also synchronized to the same (or likewise accurate) timestamps. This includes the various processing stages that are involved in the manipulation or filtering of the logs before they are stored in the CLAW. To achieve this, this document defines several standard terms that are used in the process of synchronization.

System Time

The current time and date used by computer systems to supply applications running on the system with access to accurate time. Computer systems base their system time on the current time in relation to coordinated universal time (UTC) and each time zone is designated as an offset ahead or behind by a specific number of hours.

Authoritative Time Source

A single source synchronized to UTC by which events can be time-stamped, correlated, and synchronized is required for each IT system. An authoritative time source is critical to support essential operational and analytical cybersecurity functions and processes with an accuracy determined by the functions conducted at the local site.

Time Synchronization

The coordination of the system clocks (agency, branch, and remote) and the components that comprise the systems (servers, workstations, network devices, etc.). Time synchronization is critical to support

essential operational and analytical cybersecurity functions and processes with an accuracy determined by the functions conducted at the local site.

Timestamp Standardization

Telemetry timestamping is essential for data analysis in modern networks (network troubleshooting, application performance tracking, security or threat analysis and legal compliance). Any analysis performed is dependent upon the accuracy and precision of the data being analyzed. One of the key requirements for accuracy when performing any kind of analytics is understanding precisely when a packet was captured. Modern network infrastructures may have multiple links, network tiers, or data centers between the point the data is captured and the point where the analysis is performed. The insertion of a standardized timestamp is a common method for preserving the data capture times. This method is widely used in the industry, but the implementation specifics (timestamp accuracy, format, etc.) vary based on the application.

Coordinated Universal Time

UTC is a standard universal time system that does not depend on the local calendars or the geographic location of the affected systems. It is highly accurate and can be used as a standard timing source for the purposes of synchronization and other tasks that require the knowledge of time with a high level of accuracy.

UTC is the primary time standard by which the world regulates clocks and time. UTC is sometimes also referred to as Greenwich Mean Time (GMT), or simply Universal Time (UT). This is sometimes also called Zulu or Z Time.

Time is a hard subject to regulate. Science (and society) measures time with respect to the International Celestial Reference Frame (ICRF), which is computed using long baseline interferometry of distant quasars, GPS satellite orbits, and laser ranging of the moon (the local moon of the planet Earth). Irregularities in Earth's rate of rotation cause UTC to drift regularly from the time with respect to the ICRF. To address this clock drift, the International Earth Rotation and Reference Systems (IERS) occasionally introduces an extra second into UTC to keep it within 0.9 seconds of real time. This is also known as a Leap Second.

Leap seconds are known to cause application errors. This can be a concern for developers and systems administrators. This can also introduce issues with various servers across multiple geographic regions unless it is taken into consideration and accounted for in timing calculations. In some cases, timing sources smooth out leap seconds over a given period of time (commonly called "leap smearing"), which makes it easy for applications to deal with leap seconds. In all cases, and for the purposes of CLAW reporting, it is important to know how various CSPs deal with this matter and to take that into consideration whenever timing synchronization is addressed.

Timing Synchronization

In the simplest case, the source (the CSP) and the destination (the CISA CLAW) both influence timing synchronization, and discrepancies may occur between the systems. The cloud telemetry logs are timestamped when the log entries are generated. The logs are available for examination with the agency processing tools, where the original telemetry timestamps can be viewed but must not be altered. When the logs are pushed to the CISA CLAW, the originally-generated log timestamps are retained. The cloud

telemetry timestamp format must be coordinated between agencies and CISA to ensure compatibility and accurate processing.

CISA Preference

When feasible, cloud-native telemetry timestamp format, precision, and accuracy should be preserved by agency processing to ensure accurate processing and use by CLAW systems and analysts.

APPENDIX C: CLOUD TELEMETRY PROVENANCE

Provenance of an information object refers to the object's history. This involves original author authentication (i.e.., identity of the creator) plus a method to ensure integrity of the history and object itself. More advanced forms involve a complete history of all modifications to the object along with an authenticated trail of modification. Provenance in the cloud is generally more useful than for private data stores. This is due to the ease with which such information is distributed and incorporated into other data products. Generally, provenance will include author identification, modification times, and some degree of activities performed that have affected the object's content or handling.

Provenance can be applied to the major services many applications use, including object storage, databases, and messaging services. The consistency of the provenance may be a system variable. For example, AWS CloudTrail Log File Integrity file hashes are computed each hour and reflected in a list of file signatures. Finer-grain considerations might include read-after-write synchronization of underlying objects (i.e., "will a read immediately following a write on the same object return the updated object or an older one.")

The concept of provenance can also be applied to software artifacts, especially those involved in manipulating or drawing conclusions from important or sensitive data sets themselves. Likewise, the computing environments (e.g., containers) in which software artifacts execute are important components in the overall provenance picture. Generally speaking, *provenance claims* are assertions (usually with some verifier such as a cryptographically secure signature) about the origin, authorship, and modification history of a particular object.

The degree of provenance claims processing relates closely to the degree of processing applied between a collection of sensors and the resulting telemetry reported from those sensors. The degree of processing ranges from "pass-through" to "authored." In the pass-through case, sensor data may simply be forwarded from agency sensors to CISA, whereas, at the "authored" end of the spectrum, sensor data may be interpreted, edited, summarized, transformed, or otherwise manipulated or even replaced before it is reported to CISA. In this latter case, and in many intermediate cases, it is reasonable to think of the agency itself as the author of the data (or at least one of the contributing authors) as opposed to merely a pipeline for sensor data.

For pass-through cases, most CSPs provide annotations regarding which sensors provided logging information, and the connection from CSP to agency to CISA is generally carried over an encrypted and integrity-protected channel (e.g., Transport Layer Security). CSPs also often provide an additional integrity checking mechanism for log information that involves providing periodic checksums or hashes on data written to log files. These integrity checking mechanisms may be invoked to provide an end-to-end assessment as to the veracity of the CSP-provided log data. These are applied straightforwardly to IaaS cloud tenants, as the nature and format of the logging is largely determined by the tenant (agency).

When the CSP is assumed to be offering information regarding PaaS services, the types of information being reported and corresponding integrity mechanisms may be somewhat different than conventional logging (such as flow logging). As the PaaS service may have access to higher layer information, it may be possible to have annotations regarding the individual user or account responsible for an action. In any case, the integrity protection and author identity information will be provided by the CSP's existing

logging facilities. Some minimal additional processing may be required by CISA to filter information provided by a CSP's PaaS services that is not yet implemented or require a name translation or mapping.

When an agency inhabits multiple tenancies and reports information to CISA in a push form, log information may require a form of fusing and editing. Assuming a common log type and format across each data source in each tenant, the agency is able to aggregate the sources either by interleaving or combining them in some other fashion (e.g., data from one tenancy might precede that from another). In this case, provenance claims are likely to be made by the agency above and beyond each data source. In particular, although multiple streams may arrive at the agency labeled and integrity-protected, the process of interleaving would create a new stream that itself requires provenance metadata. In short, the agency would be responsible for asserting that it provided the aggregation of the multiple streams, and constituent streams may retain sufficient provenance information to be checked end-to-end by CISA when no agency filtration is performed.

A multi-tenant agency responding to CISA pull requests may be enable custom-tailored responses. For example, the type of provenance information needed by CISA may be specified in the telemetry request and the agency could respond appropriately. In addition, the agency is not necessarily guaranteed to receive incoming telemetry requests at a predetermined rate, so the agency may need to decide which data to retain or discard. Should it be necessary for the agency to discard data, this fact should be noted and integrity protected as part of ordinary provenance processing.

As an agency performs additional levels of processing, data removal and addition may occur. In this case, the agency is an author of log information, as it is providing enrichment and editing. Provenance claims in this context are (at least) three-fold: (1) the origin of the information from the SaaS, IaaS, or SECaaS service, (2) the origin of the information used in performing the enrichment, and (3) the resulting stream provided to CISA by the agency. Agency processing should be arranged to convey both the nature of the modifications (e.g., enrichment) performed, the type of information removed, and the processing mechanisms (e.g., software artifacts) used in performing the processing.

Moving to the highest level of data processing, data from sensors or other additional services may be combined, processed, and exported to a CISA-acceptable format, along with arbitrary data transformations, filtration, and enrichment decided by the agency. In this case, the agency is the primary author of log information. Provenance claims in this context are multiple (depending on the complexity of the agency processing performed) but include: the origin of the information from the sensors and services, the origin of the information used in performing the enrichment, and information regarding the resulting stream provided to CISA and authored by the agency. The stream is being freshly authored based on information provided by services and sensors and may involve nearly arbitrary data processing. As this provides such a large degree of freedom for the agency, an indicator of the agreement between the agency and CISA demonstrating how the stream provided to CISA is sufficient for NCPS operations should be included or referenced from the provenance claims.

In many cases, a cloud tenant will have multiple sensors, services, and analytics running simultaneously to achieve multiple objectives, such as security, reliability, and performance. Consequently, an individual scenario may involve provenance from different types of systems (e.g., SaaS, IaaS) and locations, or administrative controls may be interleaved, each with different levels of abstraction and granularity or reporting capabilities (e.g., time, identity). In cases where multiple different log types can be aggregated and processed, a timestamp or transaction identifier is commonly used to provide

temporal ordering and correlation, but note that measured quantities (such as time) should be of sufficient precision and accuracy to make such log aggregation possible.[15]

CISA Preference

Provenance of cloud telemetry must be conveyed by agencies to CISA at sharing initiation and on an ongoing basis.

[15] See, for example, minimum requirements for 1msec granularity in the financial industry (Consolidated Audit Trail NMS; available at https://www.sec.gov/rules/proposed/2010/34-62174.pdf).