National Cybersecurity Protection System (NCPS) Cloud Interface Reference Architecture

Volume 1 - General Guidance

July 24, 2020

Version 1.2

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

Revision/Change Record

Version	Date	Revision/Change Description	Sections/Pages Affected
Version 1.0	12/12/2019	Initial Release Version	All
Version 1.1	4/17/2020	Response to Comments and Feedback	Added new Section 3; moved old Section 3 to Section 4 and revised content; updated Sections 1.2, 1.4, 1.5, 2.2, Appendix A; added Figures 3-9; minor graphic and text revisions throughout.
Version 1.2	7/24/2020	Response to Comments and Feedback	Added Appendix C; updated Executive Summary and Conclusion; minor graphic and text revisions throughout.

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* (NCIRA) will be released as two individual volumes. This 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. The second volume, to be released at a later date, will provide an index of common reporting patterns and considerations for how agencies can send cloud-specific data to the NCPS cloud-based architecture. Individual cloud service providers (CSPs) can use Volumes One and Two to offer guidance on vendor solutions that align with these reporting patterns.

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 in order to:

- 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] are required to meet the requirements of the National Cybersecurity Protection System (NCPS).[2] In general, this means that the Cybersecurity and Infrastructure Security Agency[3] (CISA) monitors the flow of agency network traffic and network flow logs are forwarded to CISA. CISA analysts use this data for 24/7 situational awareness, analysis, and incident response. Traditionally, network flow data has been collected by NCPS sensors located at Trusted Internet Connections (TIC) and Managed Trusted Internet Protocol Service (MTIPS) gateways, which capture security information as traffic passes between the agency and the Internet. As agencies move their information technology (IT) infrastructure to the cloud, some of their 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. This document, the *NCPS Cloud Interface Reference Architecture* (NCIRA) explains how agencies can provide cloud-generated security information to that system.

1.1 Document Organization

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

- Section 1 provides a document overview, assumptions, and constraints.
- Section 2 presents an overview of NCPS, describes how agency adoption of cloud computing impacts the program and introduces the NCPS cloud telemetry cycle.
- Section 3 expands on the NCPS cloud telemetry cycle by introducing a staged approach to cloud sensing, agency processing, and reporting to CISA.
- Section 4 describes the cloud-based architecture that CISA is developing to collect and process NCPS-relevant data from cloud deployments of federal civilian agencies.
- Section 5 offers summary information.
- Appendix A discusses different types of cloud telemetry logs.
- Appendix B explores the various locations at which network flow information can be collected.
- Appendix C presents the implementation workflow for deploying NCPS in the cloud.

[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
[3] This document discusses programs (e.g., NCPS) that predate the creation of CISA. When discussing these programs, the term “CISA” refers to both the current agency and the predecessors who previously managed those programs.

NCPS Cloud Interface Reference Architecture: Volume Two is a companion document that provides a catalog of the most common reporting patterns. Together these volumes can be used to inform agency implementers on best practices and considerations for different deployment scenarios.

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 reporting cloud telemetry to CISA. The information provided includes the introduction of general reporting patterns. The discussion in Volume 1 is vendor-agnostic and not specific to any particular CSP.
2. Volume Two of the *NCPS Cloud Interface Reference Architecture* contains specific reporting patterns and guidance for how agencies can participate in NCPS in the cloud using various common cloud use cases.

1.3 Audience

This document is designed primarily for the federal civilian agencies, contractors, and vendors that are required to comply with the NCPS program. This document can also be leveraged by stakeholders ranging from policy, acquisition, technical, and cybersecurity personnel to agency information technology leadership (e.g., Chief Information Officers (CIOs) and/or Chief Information Security Officers (CISOs)). Non-federal organizations may also derive value from this document as programs, strategies, and approaches are considered to address cloud security needs.

1.4 Assumptions

The following assumptions were used in the development of this reference architecture:

1. CISA will expand NCPS to include cloud data sources (rather than develop a new program to accommodate this new deployment model).
2. CISA will operate its own security telemetry collection infrastructure. It will replicate CLAW within several CSPs and cloud regions.
3. Agencies will continue to seek CISA assistance in securing their operations and data by participating in NCPS.
4. Cloud computing products and services will continue to evolve and expand and their adoption by Federal Civilian Executive Branch agencies will increase.
5. Federal cybersecurity policy will permit agency security data hosted on cloud services to be accessed directly by CISA (rather than through agency on-premise infrastructure).

6. Agencies are expanding the use of encryption for all types of data and encryption is expected to become increasingly common in the future.
7. CISA's initial telemetry requirements can be satisfied without payload decryption.

1.5 Constraints

The following constraints were used in the development of this reference architecture:

1. Agencies remain as data owners for all cloud telemetry and are merely sharing a copy of that data with CISA.
2. CISA makes efforts to reduce costs to agencies for sending cloud telemetry to CISA. However, agencies may still incur financial expense to fully participate in NCPS in the cloud. This occurs most naturally when an agency operates within one cloud service provider (CSP) and CISA's collection infrastructure resides in another.
3. CISA and agencies will have a written and signed memorandum of understanding (MOU) which governs the information sharing and handling relationship between both parties.
4. CISA information collection and use shall comply with public privacy impact assessments (PIA) for the NCPS program.
5. Richness of telemetry shared with CISA is bound by the agency's encryption policy. If the agency does not perform encryption "break and inspect" functions, the agency and CISA will both be unable to observe traffic payload details.

2 BACKGROUND

NCPS is an integrated system-of-systems that delivers a range of capabilities, including intrusion detection, analytics, intrusion prevention, and information sharing capabilities that defend the civilian federal government's information technology infrastructure from cyber threats. The NCPS capabilities, operationally known as EINSTEIN, are one of several tools and capabilities that assist in federal network defense.

NCPS sensors are integrated into TIC access points. As such, agencies have traditionally been able to fulfill NCPS requirements simply by complying with the TIC program. However, in 2019, Office of Management and Budget (OMB) issued an updated TIC policy, OMB Memorandum M-19-26[4], which does not require TIC access points to be embedded in all TIC use cases. Many of these new TIC use cases describe cloud services. In these use cases, network traffic between an agency and a CSP does not pass through an NCPS sensor.

As agencies and CISA adopt cloud environments and conform to the new TIC use cases, they will continue to share telemetry and security insights. This document provides guidance on how agencies can share telemetry with CISA and fulfill the requirements of NCPS when both agencies and CISA are operating in cloud environments. It furthers the NCPS objective to support "cyber" information sharing between CISA and federal agencies in order to enable a shared situation awareness between CISA and federal networks. Under this platform, CISA and agencies gain increased security visibility and enhance existing incident response capabilities needed to tackle modern cyber threats on U.S. networks.

2.1 NCPS Overview

Traditionally, TIC access points (either MTIPS gateways[5] or agency-managed TIC Access Points[6]) contain EINSTEIN[7] sensors, so when an agency participates in the TIC program, they also automatically utilize the capabilities of the NCPS program. EINSTEIN 1 (E1) monitors the flow of network traffic (i.e., network flow records) to and from a Federal civilian executive branch agency's on-premise networks. EINSTEIN 2 (E2) is an intrusion detection service that identifies potentially malicious network activity in Federal government network traffic based on specific known signatures.[8]

Under the traditional (on-premise) TIC model, both E1 and E2 are deployed and screen all network traffic that is routed from an agency through TICs, MTIPS, and the EINSTEIN 3 Accelerated (E^3A) NEST[9] locations. For E1 and E2, the agency's telemetry, in the form of network traffic logs, are forwarded to CISA, and CISA analysts use these data for 24/7 situational awareness, analysis, and incident response. Hence, participation in TIC and ensuring all agency traffic from "inside" networks to "outside" systems traverses a TIC access point is all that is required to be in full compliance with NCPS demands for E1 and E2. The top data flow arrow in figure 1 depicts this traditional E1 and E2 telemetry pattern.

[4] https://www.whitehouse.gov/wp-content/uploads/2019/09/M-19-26.pdf
[5] https://www.dhs.gov/cisa/managed-trusted-internet-protocol-services
[6] https://www.dhs.gov/cisa/trusted-internet-connections
[7] https://www.dhs.gov/einstein
[8] https://www.dhs.gov/cisa/national-cybersecurity-protection-system-ncps
[9] https://www.gao.gov/assets/680/674829.pdf (page 48)

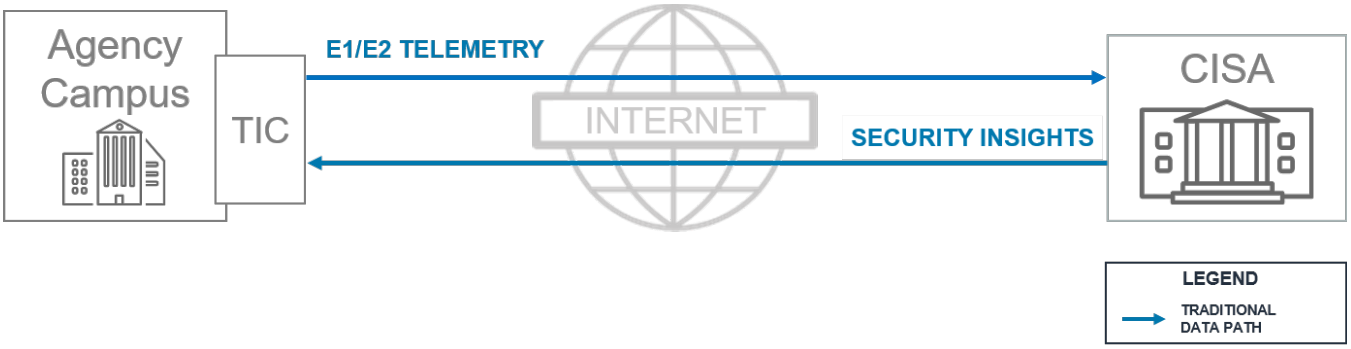


Figure 1: Current On-Premise Telemetry Configuration

Security insights are concrete intelligence data formulated for the timely identification and prevention of imminent cyber threats. Security insights may include security rules provided in a rule-based language (e.g., Snort[10] rules, Yara[11] rules, etc.), attack signatures (e.g., malware hash, malicious macros, etc.), and indicators (e.g., blacklisted IPs, email header indicators, etc.). Security insights are furnished by CISA and delivered to agencies to enable them to mitigate and counter cyber-attacks. Security insights may trigger internal processes and incident response within the agency's network to enact needed security reinforcements. Under the NCPS program, security insights can also be provided in the form of a CISA security alert to an agency concerning detected suspicious activity on the agency's network. This CISA alert may include a mitigation recommendation from CISA analysts, which will trigger an agency workflow to remediate the security threat. The bottom data flow arrow in figure 1 depicts the flow of security insights from CISA to a TIC access point.

2.2 How Agency Cloud Adoption Impacts NCPS

As part of their IT modernization efforts, many agencies are utilizing commercial cloud products and adopting cloud email, collaboration, and software tools. Many agencies are using multiple CSPs in order to meet their mission needs and are utilizing all three cloud service models: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS).[12] When an agency creates a tenancy within a CSP, traffic between that CSP and the agency may no longer pass through a TIC access point or an NCPS sensor.

Figure 2 depicts the relationship between an agency's CSP tenancy and CISA. In this diagram, an agency still has some of its network traffic traversing the traditional TIC access point, but network traffic to or within one or more CSPs does not pass through the TIC access point. The top data flow paths show the traditional flow of E1/E2 telemetry from the agency to CISA and the flow of security insight from CISA to the agency. The bottom data flow paths show the new data flows between the agency, the CSP, and CISA. Reporting patterns for data flows and telemetry collection and sharing are discussed in section 3 with more details and specific use cases provided in *NCPS Cloud Interface Reference Architecture: Volume Two*.

[10] https://www.snort.org
[11] https://yara.readthedocs.io/en/latest/
[12] Email as a Service (EaaS) is a sub-type of SaaS.

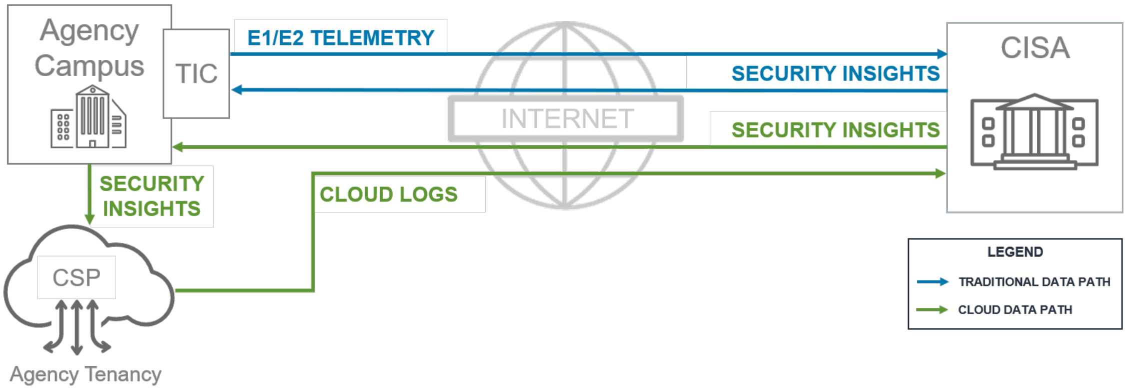


Figure 2: On-Premise and Cloud Telemetry Configuration

Because there are a wide range of CSPs and tenant-controlled security tools, there will be new data formats for telemetry (other than traditional network flows) and potential new formats for security insights for NCPS in the cloud. Data formats are discussed in Section 3 with more details in *NCPS Cloud Interface Reference Architecture: Volume Two.*

NCPS Cloud Telemetry Cycle

In order to fully realize the collection of cloud data to fulfill NCPS requirements, there is a need to define the NCPS cloud telemetry cycle, as depicted in figure 3. Each of the entities in the cycle have unique roles and responsibilities:

- CISA sends intelligence and requirements to agencies (as depicted by the blue arrow).
- An agency is responsible for protecting its data, both on-premise and in the cloud, and the agency leverages intelligence and requirements to set configurations and indicators of compromise (IOCs) in its cloud instances (as depicted by the red arrow).
- CSP monitoring and policy enforcement agents generate logs and send them to CISA as cloud telemetry (as depicted by the black arrow).
- CISA uses the cloud logs to inform situational awareness and threat discovery, resulting in new intelligence sent to agencies (as depicted by the blue arrow).

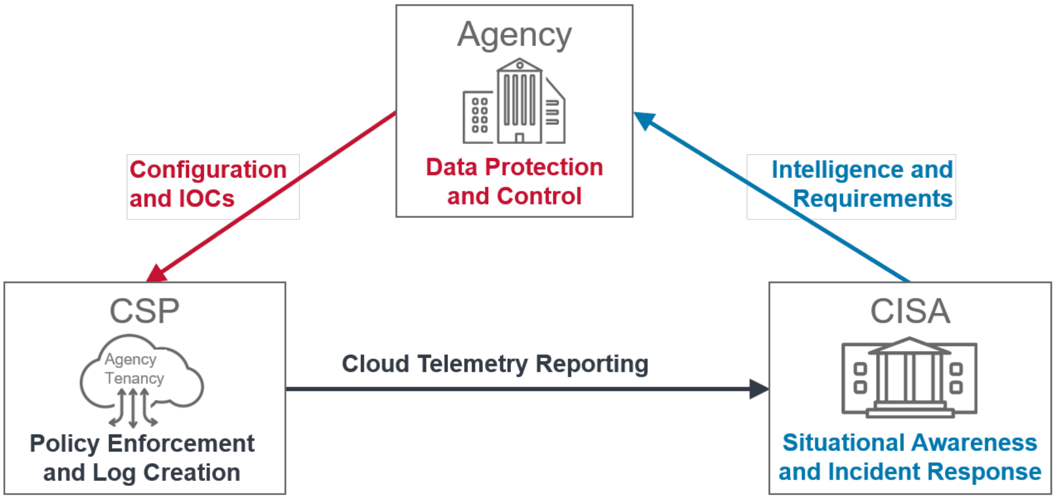

Agency
Data Protection and Control
Configuration and IOCs
Intelligence and Requirements
CSP
Agency Tenancy
Policy Enforcement and Log Creation
Cloud Telemetry Reporting
CISA
Situational Awareness and Incident Response

Figure 3: NCPS Cloud Telemetry Cycle

Benefits of Sharing Cloud Security Data With CISA

There are several benefits associated with sharing cloud security data with CISA:

1. Expanding NCPS to include agency cloud data provides CISA with the ability to gain situational awareness of threats and threat actors across the .gov domain, including on federal agencies' cloud communications. As a result, CISA can proactively respond to and mitigate cloud-based attacks against federal networks.
2. The inclusion of agency cloud telemetry extends CISA's security visibility and protection perimeter to include cloud-hosted software interactions and third-party services. This increased visibility informs and enhances incident response capabilities and federal cloud security posture. All agencies and CISA benefit from that extended visibility.
3. Additional cloud telemetry provides CISA with the ability to aggregate and correlate threat data generated and consumed in the cloud to aid in the timely discovery of security vulnerabilities and attack campaigns facing federal network cloud infrastructure.
4. Data gathered from the cloud network flow and cloud security logs provide CISA with additional intelligence and information to predict the changing security landscapes of both on-premise and cloud infrastructure, as well as to accurately plan, execute, and manage security countermeasures on the federal scale.
5. NCPS in the cloud provides a centralized model for log aggregation and analysis of a broad data set from federal cloud deployments, which result in a greater risk reduction for individual agencies as well as better availability of indicators of compromise to federal government information resources.

NCPS Roles, Responsibilities, and Cloud Operations

Transitioning to the cloud introduces new roles, actors, and procedures (e.g., an autonomous CSP, absence of TIC, third-party cloud monitoring tools, etc.) and the existing system for NCPS security insights transmission needs to be adapted. Specifically, in existing NCPS on-premise deployments, security insights in E2 are forwarded from CISA to the TIC access point (as shown in figure 1). However, when agencies utilize CSPs there is the introduction of a new telemetry exchange. E2 security insights continue to be transmitted from CISA to the TIC access points (as seen in the figure 2 blue data flow), but agencies also need to "pull" E2 security insights from CISA and transmit those security insights to their agency tenant protections hosted by CSPs (as seen in the green data flow path in figure 2).

CISA's cloud presence for collecting and analyzing NCPS information is called CLAW. It is based on a functional, module-based architecture, hosted in multiple clouds, and ingests, stores, and analyzes cloud security logs and EINSTEIN sensor data from multiple agencies using commercial CSP services. It is geared towards enabling secure and efficient methods to process cloud data in a manner that offers CISA a similar level of situational awareness provided by current EINSTEIN on-premise deployments.

Figure 4 (below) shows a more detailed analysis of the shifting relationships for NCPS implementation in the cloud. Roles which must be implemented or coordinated by more than one party are shown within the shared space of the overlapping ovals, with the participants identified. Traditional NCPS was almost entirely implemented by CISA, with the agency only playing a role in provisioning a network tap for CISA observation and use. This two-party interaction is shown below with roles labeled "Traditional NCPS (TIC)." For cloud telemetry, the agency, CISA, and CSPs each have a responsibility to enable functionality. For information on the roles and responsibilities for implementing NCPS in the cloud, refer to Appendix C.

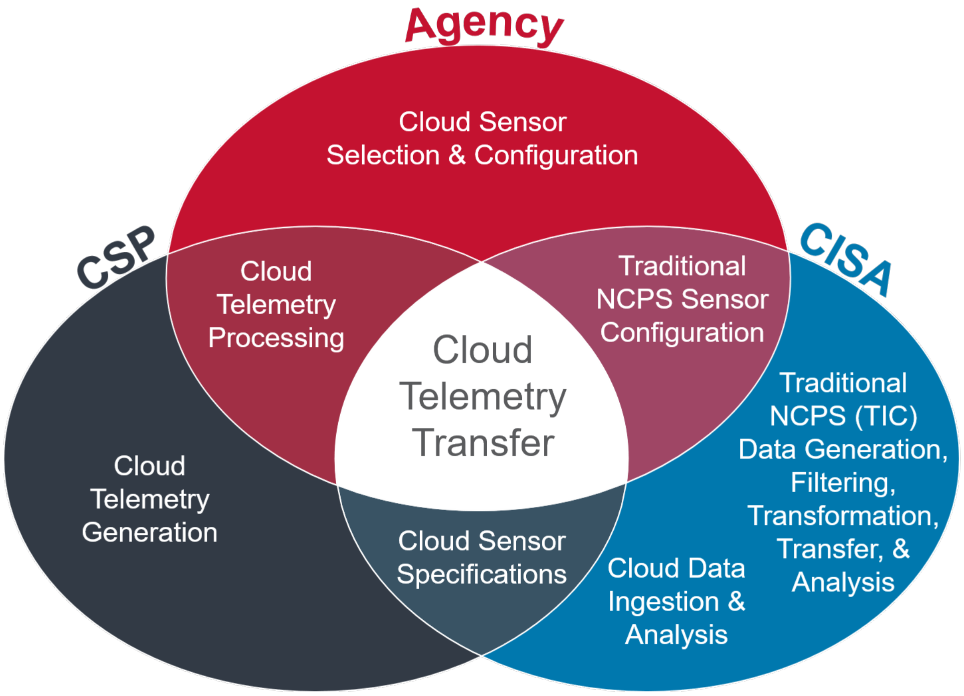

Agency
Cloud Sensor Selection & Configuration
CSP
CISA
Cloud Telemetry Processing
Traditional NCPS Sensor Configuration
Cloud Telemetry Transfer
Traditional NCPS (TIC) Data Generation, Filtering, Transformation, Transfer, & Analysis
Cloud Telemetry Generation
Cloud Sensor Specifications
Cloud Data Ingestion & Analysis

Figure 4: NCPS Roles and Responsibilities

3 AGENCY REPORTING PATTERNS

The NCPS cloud telemetry cycle was introduced in figure 3 to depict the relationship between an agency, CISA, an agency's authorized CSPs, and the information passed between parties. In this section, "cloud telemetry reporting" from the CSP to CISA (the black arrow from CSP to CISA in figure 3) will be further developed into general reporting patterns that will be used to describe unique reporting instances and possible vendor solutions in Volume Two. Figure 5 below further delineates the reporting of telemetry from agency cloud resources to CISA as taking place in three stages:

- **Stage A**: Cloud Sensing - Generates the telemetry
- **Stage B**: Agency Processing - Prepares the telemetry for communication
- **Stage C**: Reporting to CISA - Includes the transmission of information and transition from the agency to CISA infrastructure and control

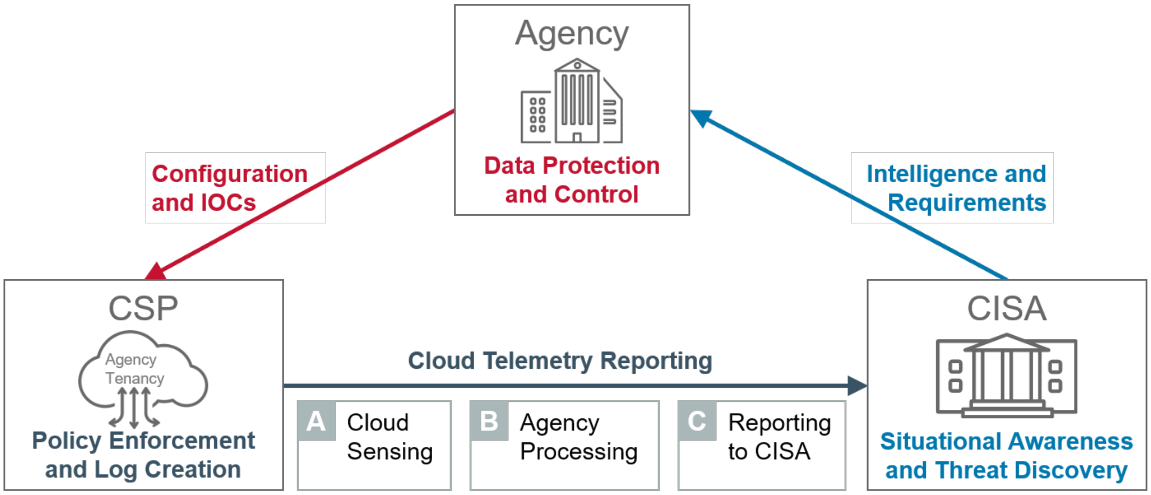

Agency
Data Protection and Control
Configuration and IOCs
Intelligence and Requirements
CSP
Agency Tenancy
Policy Enforcement and Log Creation
Cloud Telemetry Reporting
A Cloud Sensing
B Agency Processing
C Reporting to CISA
CISA
Situational Awareness and Threat Discovery

Figure 5: NCPS Cloud Telemetry Cycle Reporting Detail

Within each of the Reporting Pattern Stages there are attributes which capture the functions that take place within the stage (as shown in figure 6). Each attribute describes a specific processing element that requires the agency to select from one or more options. Within the Cloud Sensing stage (Stage A), the two attributes are "Sensor Positioning" and "Telemetry Types," which describe where and what cloud telemetry is generated. Within the Data Processing stage (Stage B), there are four attributes that describe how the cloud telemetry may be processed prior to reporting to CISA: "Data Filtering", "Data Enrichment," "Data Aggregation," and "Data Transformation." Within the Reporting to CISA stage (Stage C), there are two attributes which are used to describe how the data will be transferred to and received by CISA: "Data Transfer" and "CLAW Distribution." Details for activities within the stage as well as options for these attributes are presented in the following sections.

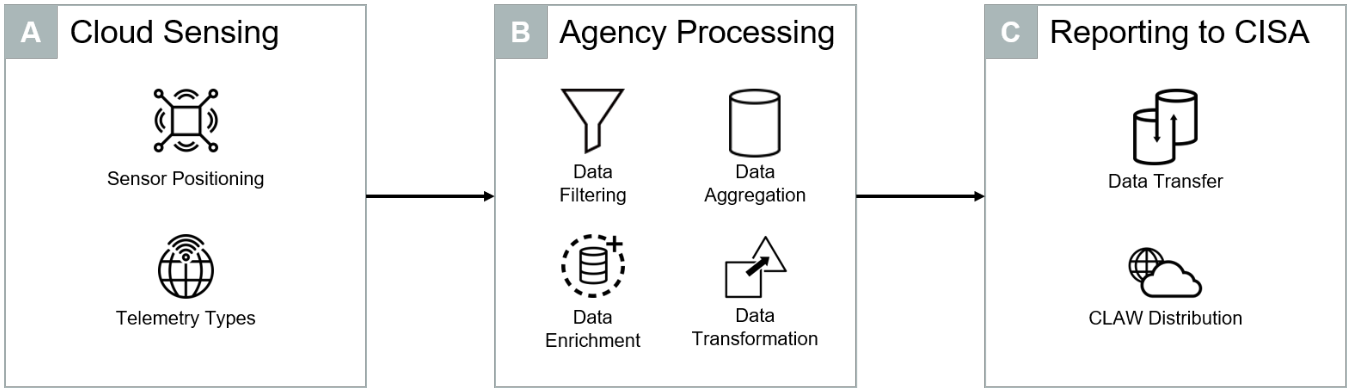


Figure 6: Agency Reporting Pattern Stages

Tools Available to Agencies

A wide range of tools are available to agencies for generating telemetry and for performing filtering, aggregation, and/or transformation, with differing functionality and costs. Options can be classified as cloud-native, agency-provided, or third-party.

Cloud Native

Cloud-native tools are provided by the CSP as services or configuration options. They are likely to offer a similar degree of trust, scalability, and interoperability as other cloud components from the same CSP. Given their native cloud support, they are also typically easier to configure and deploy. For these reasons they may be preferable to other options. However, configuration and customization of these generic capabilities will be required to align with the reporting pattern used by each agency. Services are typically priced on a tiered pay-for-what-you-use model. If data crosses CSP/region boundaries and/or is stored on agency resources, agencies also incur the associated costs.

Examples of CSP-provided capabilities include Amazon Web Services (AWS) Virtual Private Cloud (VPC) Flow Logs, Azure Network Security Groups Flow Logging, and Google Cloud Platform (GCP) VPC Flow Logs for generating network flow telemetry; AWS CloudWatch with Athena, Azure Monitor, GCP Cloud Logging with BigQuery for filtering; AWS Elasticsearch, Azure Monitor and Event Hubs, and GCP Pub/Sub for data aggregation; AWS Lambda, Azure Functions Consumption Plan, and GCP Cloud Functions for data transformation; and more sophisticated pipelines such as AWS Glue, Azure Data Factory, and GCP Cloud Data Fusion and Data Flow, which may include support for filtering, aggregation, and transformation.

Agency

Agency tools are capabilities that are developed by the agency and provided within the cloud (e.g., in virtual machines or “serverless”). Agencies using virtual machines (VMs) under their own control will be responsible for the proper scaling and load balancing between the VM instances; they will also be responsible for the associated operational costs. Software licensing costs may also apply.

Third Party

Third-party tools are provided by an external entity in the form of cloud SaaS services, cloud-based virtual machines (deployed by agencies), and/or remotely accessible web services. Agencies must interface with third-party services (e.g., through application programming interface (API) calls or publish/subscribe channels) to retrieve raw or processed data; using such services external to the agency and CLAW CSPs will entail a dependency on the provider for continuous operations and security of the service. It may also involve additional procurement and purchasing arrangements, including vendor vetting. Software licensing costs are typical when using virtual appliances available from a CSP's marketplace (although they also come with support, upgrade, and training offerings).

3.1 Stage A - Cloud Sensing

Cloud data creation is the first step of the reporting pattern (as shown in figure 6). The success of the NCPS program is directly impacted by the type of data or logs made available for analysis. Different types of security logs are available for different cloud service models (IaaS, PaaS, or SaaS) and different CSPs. The selection of the cloud log types used as E1/E2-equivalent cloud telemetry will impact whether CISA is able to attain efficient and high-fidelity threat correlation and may affect the processing performance and costs incurred by the agencies providing the telemetry.

Network flow logs will initially be considered as the primary source of data to satisfy NCPS in the cloud visibility objectives. Later, additional types of cloud logs may be considered as a data source. Appendix A elaborates on network flow and other log types.

CISA and the participating agency will determine the specific collection location for network flow records based on the agency's requirements. In all cases, the following guiding principles help scope the generation of network flow records to be sent to CISA and those retained by the agency:

1. CISA is primarily interested in the network flow records which describe agency interactions with systems or components beyond the agency visibility, control, and administration (as opposed to interactions between internal agency components).
2. Network flow record collection is enabled for all data sensitivity designations of agency information hosted in the cloud (i.e., low and high sensitivity data will both require observation).
3. When agencies have more robust information collection needs for their internal purposes, as well as for post-collection processing or filtering of logs, network flow records can be used to align data sharing with agency and CISA MOU requirements.

NCPS Cloud Interface Reference Architecture: Volume Two details specific reporting patterns.

In the Cloud Sensing stage (Stage A), agencies configure one or more telemetry sources to send raw data to the Agency Processing stage (Stage B), or, in the case of no processing, directly to the Reporting to CISA stage (Stage C).

3.1.1 Attributes and Options

As shown in table 1, the two attributes for consideration in the Cloud Sensing stage are Sensor Positioning and Telemetry Types.

Table 1: Cloud Sensing Options

Stage A – Cloud Sensing	
Attribute	**Options**
Sensor Positioning	Gateway
	Subnet
	Interface
	Service
	Application
Telemetry Types	Network Flow Logs
	Packet Captures
	Application Logs
	Transaction Logs

Sensor Positioning: Options are based on where the telemetry is generated. Sensors for network flow logs may be placed at the gateway, subnet, or interface level. Other logs may be generated on application servers or with the various CSP services used by the agency. Appendix B discusses potential network flow data collection locations in more detail.

- **Gateway**: Network sensors are placed at the gateway between an agency's cloud tenancy and the internet[13], allowing monitoring to and from all agency cloud resources. When network address translation (NAT) is used, the agency must ensure that the records report public IP addresses. A suitable example would be an agency hosting a public website – and no private resources – on its cloud tenancy. NOTE: Traffic at the gateway location may include agency "private/internal" sources not typically monitored by NCPS. Processing may be required to exclude these records prior to sharing.
- **Subnet**: Network sensors are placed at individual subnets within an agency's cloud tenancy, allowing monitoring to and from agency cloud resources in each subnet. Private and public data flows can be separated so that only the latter is shared with CISA. A suitable example would be a cloud tenancy cohosting a public website and internal HR applications on different subnets. Only the subnet with the public website is provisioned to share records with CISA.
- **Interface**: Network sensors are placed at the individual network interfaces used by cloud virtual machines, allowing monitoring to and from each configured interface. Private and public flows can be separated, but with finer granularity than the subnet option and potentially greater insights for event correlation and analysis. A suitable example would be a server cohosting a public website and internal HR applications on different interfaces. Only the interface with the public website is provisioned to share records with CISA.
- **Service**: Telemetry is generated from CSP services that provide key functions, such as load balancing, network/application firewalls, DNS, identity/authentication, key management, and more. In this way, services can double as sensors; however, they differ in their ability to be used

[13] As well as other networks peered with the agency's cloud tenancy.

as a data feed. Some can be configured to periodically deliver telemetry to agency cloud storage resources or to the CSP's monitoring service. Others may only make telemetry available through API calls or a manual export process.

- **Application**: Telemetry is generated from application servers, such as web servers and mail servers. Similar to services, application servers double as sensors, generating logs. They also vary in the level of visibility offered and log access mechanisms.

CISA Preference

CISA prefers that agencies place network flow sensors at each public subnet. This is the simplest way to ensure coverage of all public data flows while excluding many private data flows.

Telemetry Types: Options are based on what kind of telemetry is generated and include network flow logs, packet captures, application logs, and transaction logs. Appendix A describes each type in greater detail.

- **Network Flow Logs**: Network flow logs provide basic information about the data flows to and from agency publicly accessible cloud resources.
- **Packet Captures**: Partial packet captures consist of network packets that provide context to one or more security events, such as an intrusion detection system (IDS) alert.
- **Application Logs**: Application and event security logs for individual applications (such as web, email, and DNS) provide information about client access and use of these services.
- **Transaction Logs**: Transaction (or audit) logs for individual servers and/or their cloud tenancy provide information about administrative access and changes to their systems.

CISA Preference

CISA prefers that agencies generate and share network flow logs as a first step and will coordinate with agencies on additional telemetry types.

3.1.2 Caveats and Considerations

Agencies can use the following caveats and considerations for evaluating options in the Cloud Sensing stage:

- **Processing Requirements**: The telemetry at this stage is raw and unfiltered and agencies must provision the storage, network and compute resources necessary to process the full volume of the data they share; agencies should consider these requirements when choosing what sources to share and how to configure them.
- **Output Formatting**: Many sources can be configured with settings affecting the formatting and fields of the telemetry they provide. Agencies can use these settings to eliminate some work that

would otherwise be done in the Agency Processing stage but should do so cautiously. For example, fields excluded from the original telemetry cannot be recovered later, whereas they would still be available to the agency if they were filtered out during the processing stage.

- **Visibility**: Agencies should provide as much visibility as possible about "public" interactions between their cloud systems and external networks and should minimize sharing on "private" interactions between internal components. Agencies should also consider whether each additional source they share increases visibility or is merely redundant.
- **Break-and-Inspect:**[14] Sensors which generate telemetry by inspecting traffic payloads should be able to "break-and-inspect" encrypted traffic. They should be positioned to minimize the associated risks (embedding a certificate authority, decrypting potentially sensitive traffic, etc.) and to ensure that break-and-inspect is not performed redundantly.
- **Encryption**: Telemetry data should be protected in transit. This includes when data from cloud-native and third-party sources are transferred to the agency and when data is transferred to later stages. Parameters for key length, key rotation, and cipher suites should be restricted to those that provide sufficient protection; specific details are given in Volume Two. Telemetry data should also be protected at rest (e.g., when it resides in cloud storage or on an agency-provisioned sensor).
- **Source Costs**: To the extent that agencies are sharing telemetry that they already generate and use for themselves, the costs of generating telemetry to share with CISA are minimal; however, an agency may need to add additional sensors. CSP services providing cloud-native telemetry are typically provided at little to no cost,[15] whereas agency-provided telemetry typically involves the cost of operating the sensors (and potentially the cost of licenses) and third-party telemetry typically involves subscription costs.

3.2 Stage B – Agency Processing

In the Agency Processing stage (as shown in figure 6), data collected from the cloud are filtered, enriched, aggregated, and transformed into appropriate data formats that can be ingested by CLAW. The simplest option is where no processing takes place: essentially raw log data are copied or moved directly from agency cloud sources to the reporting stage via a push or pull operation. More sophisticated processing may involve aggregation, enrichment, filtering, and data format transformation (called "data wrangling" or "data munging"). An agency may have multiple sensors and data streams distributed across multiple cloud subscriptions, regional cloud instances, or even CSPs. Combining these together implies some method for ordering and interleaving (e.g., by time) and filtering out certain information, such as internal data transfers or logs containing sensitive information not required by CISA for threat analysis. Filtering of data may also be necessary to reduce the volume of information delivered to CLAW.

Data wrangling may be performed by an agency or by another party chosen by the agency for this task. Many CSPs offer such cloud-based capabilities either in their native ingestion pipelines and services or

[14] This consideration does not apply to network flow logs and may be ignored by agencies which are in the early processes of sharing telemetry with CISA. It is, however, applicable as a general best practice.

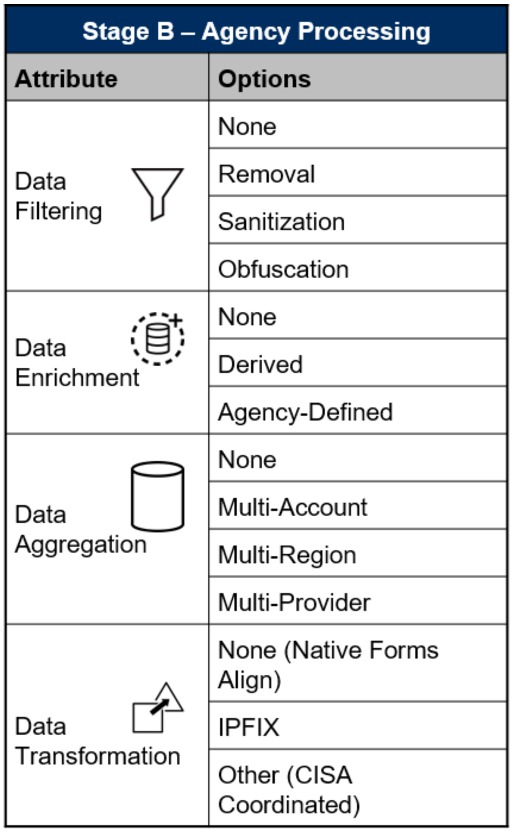
[15] For telemetry generated by one CSP and delivered to another (i.e., the agency is using multiple CSPs), traffic egress and ingress costs also apply. See Section 3.2.1, Data Aggregation.

using third-party capabilities from within their respective marketplaces. Note that a very broad set of implementation options are available for the processing stage; the details vary among different CSPs and continue to change as their offerings improve.

3.2.1 Attributes and Options

As shown in table 2, the four attributes for consideration in the Agency Processing stage are Data Filtering, Data Enrichment, Data Aggregation, and Data Transformation.

Table 2: Agency Processing Options



Stage B – Agency Processing	
Attribute	**Options**
Data Filtering	None
	Removal
	Sanitization
	Obfuscation
Data Enrichment	None
	Derived
	Agency-Defined
Data Aggregation	None
	Multi-Account
	Multi-Region
	Multi-Provider
Data Transformation	None (Native Forms Align)
	IPFIX
	Other (CISA Coordinated)

Data Filtering: Agencies providing data to CLAW will be required to filter logs to only provide the material the agency wishes to be analyzed by NCPS, both to satisfy privacy requirements and to improve the efficiency of analysis. Data selectors may include sensitivity markings (e.g., FOUO), network flow information (e.g., IP addresses, port numbers, protocols), or other information (e.g., domain names, user or system credentials, time of activity). Records containing unwanted information may be handled through a variety of mechanisms, including removal, sanitization, and obfuscation.

- **None**: Unfiltered logs are sent to CISA. This option is only appropriate when the agency is confident that the raw logs will not contain any information that they do not wish to share.
- **Removal**: Records containing unwanted information are discarded. As they otherwise contain information of interest, this option is only appropriate when the agency is confident that discarding these records will not result in a gap in visibility.

- **Sanitization**: Records containing unwanted information are sanitized, such that the information of interest is retained, and the undesirable information is completely erased. This can range from systematically removing a field from every record to blanking fields in individual records on a case-by-case basis.
- **Obfuscation**: Records containing unwanted information are obfuscated, such that the information of interest is retained. The undesirable information undergoes a transformation that preserves its usefulness for analytics while making it impossible to derive the original values. For example, real names may be substituted with a number, permitting the agency to recover the original content using a lookup table.

CISA Preference

CISA prefers that agencies sanitize records containing unwanted information. Agencies may also perform obfuscation; however, CISA analytics will not be dependent on data that agencies would only share in an obfuscated form.

Data Enrichment: As opposed to data filtering, in which agencies subtract unwanted information, with data enrichment agencies add desirable information to the records they share with CISA. Enrichment, if performed, may consist of either derived data and/or agency-defined data.

- **None**: Unenriched logs are sent to CISA. This option is acceptable when the logs already contain all the data fields that are expected by CISA.
- **Derived**: Agencies use existing information within records to derive and insert additional information of interest. Derivation can be used to provide required fields. For example, an agency has a cloud telemetry feed that provides a destination IP address/URL but omits the destination port; the agency derives the missing port based on the service offered at the IP address/URL.
- **Agency-Defined**: Agencies supplement records with additional contextual information that would otherwise only be known to the agency. Examples include identifying endpoints as either client or server, distinguishing between administrator, user, and guest entities, and mapping IP addresses to names of subnets. Agencies should coordinate with CISA when providing this additional information so that CISA is able to make the best use of it.

CISA Preference

Agencies may choose to perform any or no data enrichment, as long as CISA receives records with all the desired fields.

Data Aggregation: Agencies may wish to aggregate multiple sensor data sources. They may possess sources in multiple CSPs, multiple regions/cloud types, and/or multiple accounts/tenancies within the same CSP. Likewise, there are multiple instances of CLAW in different locations acting as target(s). As moving data between CSPs or regions is comparatively expensive versus remaining “local,” aggregation functions will likely be located based on the relative position of the sources and CLAW(s). Aggregation options include none, multi-account, multi-region, and multi-provider; more details are provided in Volume Two.

- **None**: Each telemetry stream is sent separately to CISA; alternatively, the agency only has a single telemetry stream.
- **Multi-Account**: Cloud logs from multiple accounts or tenancies within the same CSP are aggregated.
- **Multi-Region**: Cloud logs from multiple regions within the same CSP are aggregated.
- **Multi-Provider**: Cloud logs from more than one CSP are aggregated.

CISA Preference

Agencies may choose to use data aggregation at any level – account, region, and/or provider – as long as CISA receives the desired log fidelity and delivery is timely.

Data Transformation: Agencies will need to transform their data into a format known to CLAW; target formats include the native log format, IPFIX, and other CISA-coordinated formats. Additional details and guidelines, including supported formats and expected fields, may be found in Volume Two.

- **None (Native Forms Align)**: The native log format generated by the sensors is already known to CLAW.
- **IPFIX**: The native log format generated by the sensors is unknown to CLAW and logs are transformed into IPFIX.
- **Other (CISA Coordinated)**: Agency-managed, third-party, and/or proprietary formats may be provided to CLAW if they provide a mechanism for CLAW ingestion processing and use. Approval for such log formats will be given on a case-by-case basis as coordinated with CISA.

CISA Preference

CISA prefers no data transformation (assuming format is compatible with CLAW).

3.2.2 Caveats and Considerations

Agencies can use the following caveats and considerations for evaluating options in the Agency Processing stage:

- **Content**: The selection of which data to provide to CLAW and NCPS may be driven by several factors, including privacy, data rates, network or storage costs, and formats. Every agency is responsible to provide all the required log data with appropriate security controls at a rate that does not overwhelm CLAW's ingestion system.[16] The data must also contain enough detail that NCPS analysts and analytic processes can produce useful threat analyses and alerting. CLAW guidelines for expected fields within various log formats and log categories will be further discussed in this document as well as in Volume Two and other guidance provided by CISA.

- **Aggregating Across CSPs**: Data aggregated from several sources that span CSP, region, or account may require careful account access control configurations and may incur additional communication costs. Security protections and usage tracking (billing) also tend to be tailored for use within a single CSP, so spanning providers may imply the need to create a set of compatible contracts and procurement processes across multiple vendors.
- **Combining Streams**: Several log streams may be combined into a smaller number to reduce the volume of data ingested by CLAW. This is an important factor, given CLAW's task of handling the volume of data created by all participating federal civilian executive branch departments and agencies. When combining data streams, data are generally ordered by some field. Most commonly for network metadata, this is a timestamp provided by sensors during collection. Issues regarding timing include time synchronization, precision, and accuracy of the timestamp. Different sensors, if not synchronized in a uniform manner, or with insufficient accuracy, will likely result in an unwanted stream of interleaved data record order, making subsequent analysis more difficult. An insufficient precision may result in the erroneous appearance of multiple simultaneous events. This can also frustrate subsequent processing and analysis.
- **Formats**: Data format conversion may be required if sensors produce logs in formats that are unknown to CLAW. Formats such as NetFlow and IPFIX are similar and may have relatively straightforward transformations. Logs from other components (e.g., IDS systems, web proxies) tend to be of a more proprietary nature and more complicated data transformations may be necessary. Complicated data transformations may consume significant processing time, thereby increasing the overall end-to-end ingestion processing time.
- **Data Rates**: Although it can use cloud scaling to handle additional load, CLAW (and the environment in which it resides) ultimately has a limit to the rate at which it can ingest data. This may be limited by one or more bottlenecks in processing, networking, or storage. In performing data wrangling involving the aggregation of multiple flows, assuming de-duplication has already been performed, some method for adjusting the incoming flow rate to CLAW may be required. Options include flow control back to the sensor sources (e.g., employing the underlying network protocol flow control), buffering data for a limited period of time (if the sensor data is bursty rather than persistently exceeding CLAW ingestion rate), or sampling the incoming data to reduce its rate. Care must be taken in sampling, as periodic sampling can under- or over-emphasize certain periodic phenomena.
- **Encryption**: Data arriving for processing is likely to be encrypted. In sophisticated processing scenarios (e.g., that involve de-duplication or merging by timestamp) the processing must have access to the contents of the sensor data streams supplied. Consequently, the processing agents must have access to keys to decrypt incoming data. Likewise, the outgoing (to CLAW) data or connections are also encrypted. The keys or certificates used to support this encryption must also

[16] In addition, given that multiple agencies may start providing log data to CLAW simultaneously, a throttling and/or load shedding mechanism between the agencies and CLAW for the log feeds may be required (but is not yet specified). Without such a mechanism, some ingestion data loss may be unavoidable.

be of sufficient strength and maintained appropriately in order to ensure the security of the most sensitive data handled between all the sensors and CLAW.

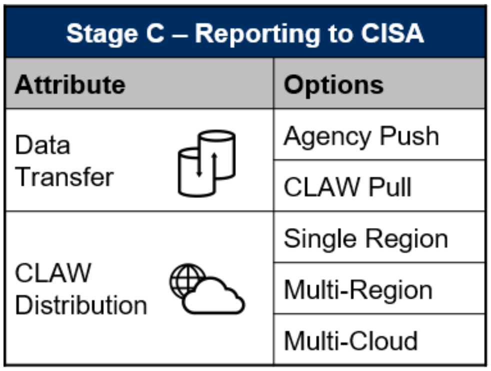
3.3 Stage C – Reporting to CISA

Once cloud telemetry has been sourced and processed in the Cloud Sensing and Agency Processing stages (respectively), the results are reported to CISA. The Reporting to CISA stage (as shown in figure 6) consists of the data transfer of cloud telemetry to one or more CISA CLAW repositories. This represents a transition from data handling protections being handled by the agency to being handled primarily by CISA. As stated above, agencies retain authoritative data ownership and are only relaying a copy of their security telemetry to CISA for use in situational awareness and incident response.

3.3.1 Attributes and Options

As shown in table 3, the two attributes for consideration in the Reporting to CISA stage are Data Transfer and CLAW Distribution.

Table 3: Reporting to CISA Options



Stage C – Reporting to CISA	
Attribute	**Options**
Data Transfer	Agency Push
	CLAW Pull
CLAW Distribution	Single Region
	Multi-Region
	Multi-Cloud

Data Transfer: Data transfer involves the mechanism by which agency cloud data is transferred to CISA after the data is collected and processed. Based on the party initiating the communication request, data transfer to CLAW can be classified as an agency push or a CLAW pull.

- **Agency Push:** The agency initiates a data transfer from their infrastructure (either at the Cloud Sensing or from Agency Processing stage) to CLAW. CISA hosts the receiving end (CLAW) in a listening fashion and issues credentials for agency use. The agency utilizes these credentials to authenticate, establishes a secure transfer means, then transfers the telemetry to CLAW. The agency may use the same credentials to transfer more than one data type to CLAW. In other words, CLAW will host a unique repository for each protected entity and agencies may populate their repository with multiple data types comingled in the same data store.
- **CLAW Pull:** The agency establishes a repository of interesting telemetry with reachability from CISA systems, issues access credentials for CISA use, and then listens for CISA pull requests. Agencies must negotiate polling interval, buffer duration during connection disruptions, link capacity, multiplexing, error correction, and other technical details with CISA. Although some of

these details also apply with push option, they are more relevant here, as the agency is no longer in direct control of the data transfer.

CISA Preference

CISA prefers the agency push option to enable agencies to more fully execute their role as data owner and more tightly control the volume, rate, and content of shared telemetry. CISA hosts the receiving end in a listening fashion and issues credentials for agency use.

CLAW Distribution: To reduce cost (both latency and monetary), CLAW infrastructure is hosted in multiple locations. The intent is to position CLAW infrastructure in such a way that agencies can transfer their cloud data to a "local" repository that is in the same CSP and region.

- **Single Region:** Agency data transfers are to a single CLAW location. This option is especially low-cost if the agency telemetry is hosted within the same CSP infrastructure (such as AWS GovCloud West) as a CLAW instantiation (this may not always be possible).
- **Multi-Region:** Agency data transfers occur to more than one regionally-located CLAW within the same CSP infrastructure, such as AWS GovCloud East and AWS GovCloud West.
- **Multi-Cloud:** Agency data transfers occur to more than one CLAW hosted on more than one CSP infrastructure (such as AWS GovCloud and Microsoft Azure).

CISA Preference

CISA prefers an agency to use the CLAW distribution option that matches its cloud deployment in order to reduce cost (both latency and monetary).

3.3.2 Caveats and Considerations

Agencies can use the following caveats and considerations for evaluating options in the reporting to CISA stage:

- **Encryption**: The mechanisms employed to provide protection of data in transit must be mutually agreed upon by the agency and CISA. Parameters such as accepted encryption ciphers, key lengths, key lifetimes, authentication factors, key/credential distribution, and others must be established.
- **Initiation Costs**: Typically, the party that initiates the data transfer incurs additional costs. However, regardless of the initiating party, data is always outbound from the agency; if the data leaves the CSP, the agency incurs outbound data transfer costs, which can be greater than inbound costs.
- **Transfer Frequency**: The frequency or timeliness of the transfer can be based on time differentials (e.g., polling every five minutes) or based on log size (e.g., after every 20MB) of accumulated new

data. In addition to the batching mechanism, transfers may also be triggered by noteworthy events (such as an unusually high volume of traffic).

- **High Availability/Durability**: CLAW infrastructure will provide high availability and data durability at each instantiation, ensuring resilient service offerings and increasing agency confidence.
- **"Local" Transfers**: Agency telemetry sharing with multi-region and multi-cloud CLAW infrastructure can reduce agency data transfer costs. This is due to multiple "local" data transfers potentially being more cost efficient than a single transfer to a "remote" location. However, this may increase technical complexity and administrative overhead.
- **Finite Deployments**: While attempts will be made to host CLAW resources as close to agency tenancies as possible, there will only be a finite quantity of CLAW instantiations. These locations may not fully accommodate the breadth of agency service locations. This will require the transfer of agency cloud telemetry to a "nearby" CLAW that may not be co-hosted on the same CSP.
- **Data Retention**: After confirmed transfer to CISA of telemetry copies, agencies must determine the duration to retain transferred data prior to deletion or removal.
- **Least Privilege**: CLAW must only be able to obtain data intended to be shared with CISA. Access permissions granted to the principal which pushes data to CLAW, or which pulls data from the agency on behalf of CLAW, should follow the principle of least privilege.

4 CISA CLOUD DATA AGGREGATION

CISA seeks to improve performance, reduce costs, and enhance threat discovery and incident responsiveness for agencies. CLAW is designed to support these goals by supporting agency adoption of cloud technologies. This section explains CLAW in more detail.

4.1 Cloud Log Aggregation Warehouse Overview

CLAW is a CISA-deployed architecture for the collection and aggregation of NCPS data from agencies using commercial CSP services. While agency NCPS data is currently aggregated on-premise at CISA, CLAW is deployed in the cloud to aggregate agency security logs that originate in the cloud. CLAW presents a functional, module-based architecture to ingest, store, and analyze security logs and sensor data from agencies. It is geared towards enabling secure and efficient methods to process cloud data in a manner that offers CISA a similar level of situational awareness provided by current EINSTEIN on-premise deployments.

4.1.1 CLAW Distribution

The CLAW architecture supports log aggregation at multiple locations (optimized for performance, cost, and efficiency) utilizing centralized threat discovery with distributed analytics. Figure 7 depicts how agencies in different cloud regions can each transfer their data to CLAW without requiring each to push their data to a single region.

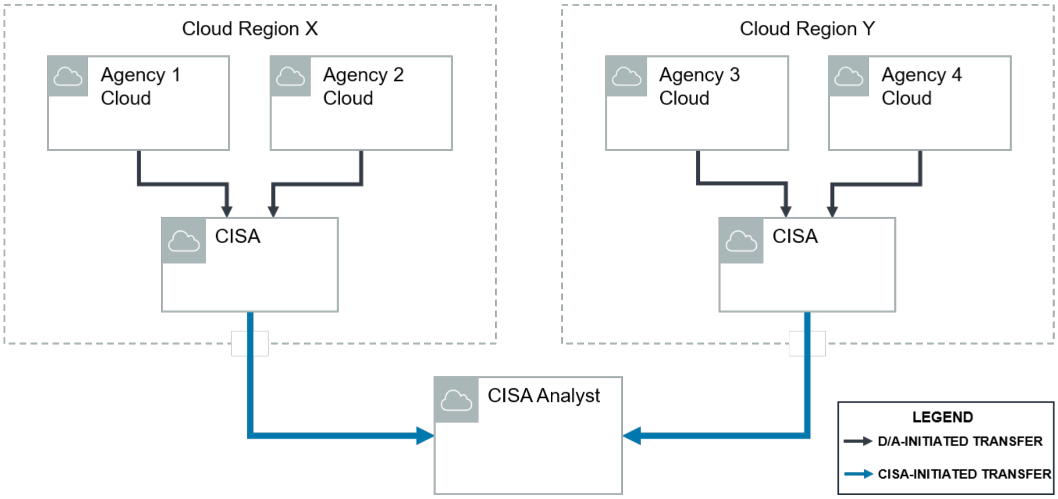


Figure 7: Responsibility for Transferring Security Data (Agency vs. CISA)

Agencies can transfer their data to the closest CISA CLAW location based on their CSP and region, reducing data transfer costs, technical complexity, and transmission latency. The CLAW architecture supports collocating CLAW aggregation points with agency tenants on major CSPs. Any additional data aggregation or consolidation required will occur within CISA's purview. Further details about issues and caveats related to CLAW distribution can be found in Section 3.3.3.

4.1.2 CISA Analysis of Agency Data

Using CLAW, CISA provides the environment and tools to correlate and discover threats from application and network data that has been shared by agencies. Current analytics approaches involve signature-based (pattern recognition) and non-signature-based (heuristic and statistical) analytics for identification of IOCs and for identification of anomalous activities. Analysis will also bring in enrichment data to enhance the analysis results.

Figure 8 shows how cloud data from individual agency cloud tenancies is collected and analyzed at CISA cloud sites while preserving agency data isolation. Each agency's data is separated to prevent data comingling and corruption (using means such as independent data indexes and data stores). Analysis results obtained will subsequently be sent to CISA analysts for processing and assimilation (i.e., for threat detection and correlation, and for synthesis of security indicators).

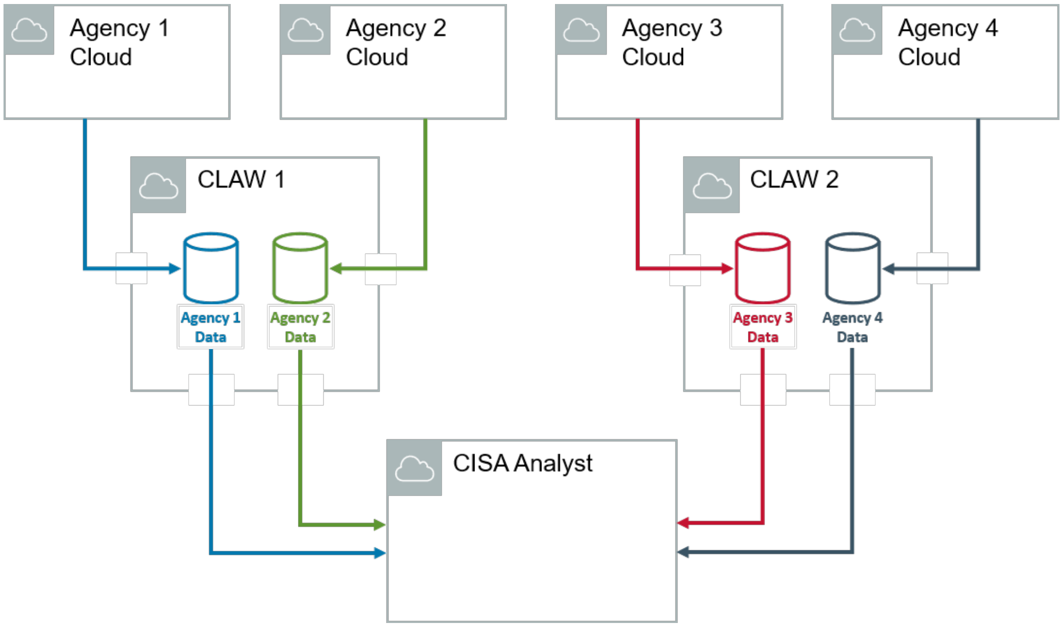


Figure 8: Agency Log Ingestion (Autonomy Preserved with Log Isolation)

Analysis will be coordinated from a central location with standardized tools. Those centralized tools will be able to interact with the sensor data distributed across CLAW locations. The data will be ingested and processed at a "local" CLAW location relevant to their CSP and region; in other words, the agency data will not be backhauled to a central repository. This will provide analysts with global situational awareness without requiring a corresponding centralized data store or requiring multiple copies of the same tools at each of the distributed data stores.

The data will be protected to ensure confidentiality and integrity using encryption for both data in transit and at rest that is compliant with Federal Information Processing Standard (FIPS) Publication 140-2. In addition, data retention compliance requirements and data recall capability for long-term forensic discovery will be met until data is destroyed or removed. CISA sustainment operations (NOC/SOC) will have oversight of the CLAW.

5 CONCLUSION

As agencies move more of their applications and services to CSPs, the NCPS program is evolving to ensure that security information about cloud-based traffic can be captured and analyzed and that CISA analysts can continue to provide situational awareness and support to the agencies. This document introduces general reporting patterns for how cloud logs will be collected and transferred to CLAW. Appendix C describes the implementation workflow for how agencies, CSPs, and CISA will partner to deploy NCPS in the cloud. The companion document (*NCPS Cloud Interface Reference Architecture: Volume Two*) provides additional details for reporting patterns and options that match common agency cloud use cases. Together, these two documents provide guidance for how an agency can adapt their cloud environments to allow for security data to be sent to NCPS. Individual CSPs (e.g., AWS, Microsoft, etc.) can use these documents to provide vendor solutions for agencies and CISA to utilize.

APPENDIX A: CLOUD TELEMETRY TYPES

Network Flow Logs

IP network flow logs describe the communication that takes place between endpoints and enables modern network management and security. Network flow log protocols[17] specify how information about such communication is to be generated and formatted. Participating network devices implement these protocols by generating, compiling, and organizing network flow records as traffic traverses them. When collected and sent to CISA, network flow logs enable CISA to have situational awareness of an agency's cloud network activities.

Cloud activities deployed in multiple service models (IaaS and PaaS specifically) may generate different types of network flow logs. For example, when a node in the cloud initiates communication with another node, an intermediate sensor device with network flow record generation capability creates a flow record for that communication. Subsequent packets with the same network flow attributes update previously created flow records, which are continuously monitored and updated until the communication concludes. Flow records are then sent to a collector, where data logs are stored and further analyzed.

Packet Captures

Packet captures, often referred to as PCAP, consist of network packets that are intercepted in real-time and stored for analysis. As this includes the full header and payload of each individual packet, they have the potential to offer much greater insights than network flow logs alone. Due to the sheer volume of data that can pass through the wire on any given day, full captures are retained for a shorter period than other telemetry types (if at all). Typically, initial analysis determines if any packets are suspected as being part of an adversary transaction and worthy enough to retain long-term and the rest of the packets are discarded. Related approaches are to only capture packet headers and/or to apply filters restricting what packets are initially captured. Partial packet captures, such as those consisting of the packets that triggered an IDS alert, may be sent to CISA under certain circumstances to enable CISA to have situational awareness comparable with E2.

The ability to conduct packet captures is only available in the IaaS service model and only recently have many CSP vendors made a cloud-native service available on a general basis. Cloud-native capabilities have an immense advantage over agency or third-party capabilities, which require forwarding traffic to a central sensor or distributing sensors to each node. In some cases, cloud-native packet capture can target network interfaces provisioned by other CSP services (on which the agency would not have been otherwise able to conduct a packet capture).

Application & Event Security Logs

Cloud application and event security logs are actively generated and stored to collect security and statistical information on observable cloud activities. The types of logs that are available are dependent on the type of cloud service model (IaaS, PaaS, or SaaS) and the specific CSP.

[17] IPFIX is a prominent example and is related to the E1 and E1E formats.

When application and event security logs are collected and stored, they may be sent to CISA under certain circumstances (as defined in Volume Two) to enable CISA to have situational awareness of additional agency cloud activities (beyond network flow). The applications that are deployed at the IaaS, PaaS, and SaaS domains in the cloud actively generate logs for analysis by CISA and the individual agency. Examples of common application and event security logs that can provide required visibility to CISA include:

- Web (HTTP)
- Email (SMTP)
- Naming (DNS)
- Identity and Authentication services (Active Directory and Certificates)

Typically, these logs are first generated by the application server when clients request server resources. Subsequent interactions with the same resource update previously created records, which are continuously monitored and updated until the communication ends. When the communication is over, the records are sent to a collector, where data logs are stored and further analyzed.

The format and fields of application and event security logs are defined by the underlying application and not necessarily by cloud providers. For example, a web application record will include fields such as HTTP headers, client user agent, content type, number of bytes, TCP port numbers, TLS session information, and timestamp.

If stored application and event security logs are collected and sent to CISA, these logs will provide CISA with visibility and situational awareness into an agency's cloud application activities. Moreover, these records can be used during post-event analysis, incident response, and potential root cause analysis on known and perceived threats.

More details about application and event security logs will be discussed in Volume Two.

Transaction Logs

Transaction logs, also known as audit logs, document the sequence of changes made to a system (such as updates to configurations and administrative actions which require elevated privileges). Authentication events are typically included (which, depending on the entity, enable subsequent changes to the system), as are events corresponding to attempted but unauthorized/unsuccessful changes. Read-only events are sometimes included. As with application and event security logs, transaction logs may be sent to CISA under certain circumstances to enable better situational awareness.

In the IaaS service model, transaction logs are generated by individual servers. However, cloud tenancies of all service models (IaaS, PaaS, and SaaS) are themselves systems that offer transaction logging capabilities. Given that CSPs follow a multi-user and self-service paradigm for each tenant, sharing these transaction logs with CISA enables visibility and situational awareness beyond individual applications (i.e., into the cloud tenancy as a whole).

APPENDIX B: FLOW RECORD COLLECTION LOCATION

The flow record collection guidelines discussed in this document apply differently depending on where in the agency's cloud system the demarcation point occurs. While CISA provided guidance on demarcation points, found in Section 3.1, it is the agency's responsibility to identify the relevant demarcation points within their cloud environments. Three typical deployment locations for network flow generation in IaaS deployments are shown below in figure 9. In addition, this appendix enumerates some conditions under which each is suitable.

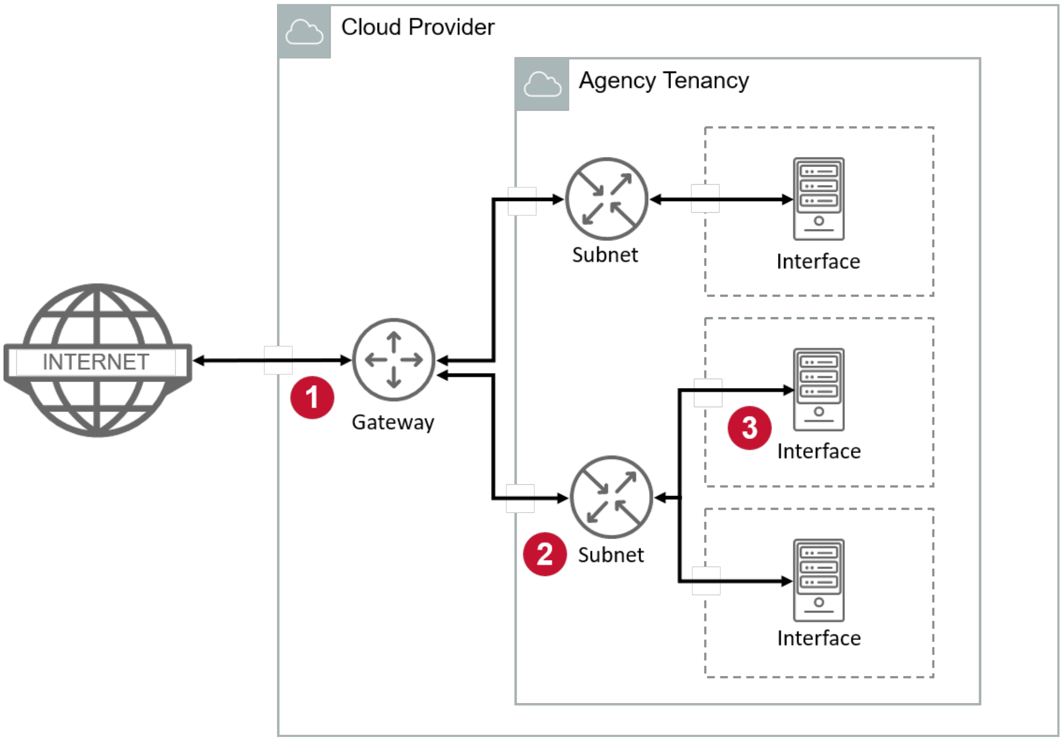


Figure 9: Network Flow Log Generation Positions for IaaS

There are at least three potential collection locations for agency tenancy network flow records. Each of these collection locations has unique visibility scope and detail.

❶ Internet Gateway

The first potential collection location is the gateway at the internet to agency cloud tenancy interface(s). Collection of network flow records at the gateway allows monitoring of all traffic to and from all agency cloud resources. When NAT is used, the agency must ensure that the records gathered reports the public IP addressing. The traffic monitored at this location may include agency "private/internal" sources not typically monitored by the NCPS sensors. The records gathered here may require processing to exclude those records prior to being sent to CISA. An example would be a publicly accessible web site hosting publicly available information, where the agency is not cohosting any additional resources on same cloud tenancy.

❷ Subnet

The second potential collection location(s) are the subnet(s) utilized by the agency tenancy. Collection of network flow records at the subnet level allows for monitoring of all traffic to and from cloud server(s) on each individual subnet. The "private/internal" and "public" data flows can be separated, thereby constraining the sharing of data flow information with CISA to only the "public" resources and reducing post-collection processing requirements. An example would be a publicly accessible web site hosting publicly available information and internal human resources applications in the cloud with "public" and "private/internal" data flows (respectively) destined for resources on independent subnets. The subnet with the publicly available information is provisioned to share network flow records with CISA.

❸ Interface

The third potential collection location(s) are the interface(s) utilized by the agency tenancy to provide access to cloud virtual machines. Collection of network flow records at the interface level allows for monitoring of all traffic to and from the individual interfaces on each of the cloud server(s) that has been properly configured to provide this capability. The "private/internal" and "public" data flows are separated by the individual virtual interfaces. This type of collection location provides the finest granularity for the network flow records, minimizes the post-collection processing requirements, and permits greater insights for event correlation and analysis. An example would be a publicly accessible web site hosting publicly available information in the cloud with the public data flows destined for resources on independent interfaces. The interfaces with the publicly available information are provisioned to share network flow records with CISA.

APPENDIX C: NCPS IN THE CLOUD IMPLEMENTATION WORKFLOW

In order to implement NCPS in the cloud, CISA, Agencies and CSPs will need to take separate, coordinated actions. Figure 10 shows a workflow sequence of the actions that must come together in order to successfully implement NCPS in the cloud. The sections below give details about the actions so that Agencies can plan accordingly.

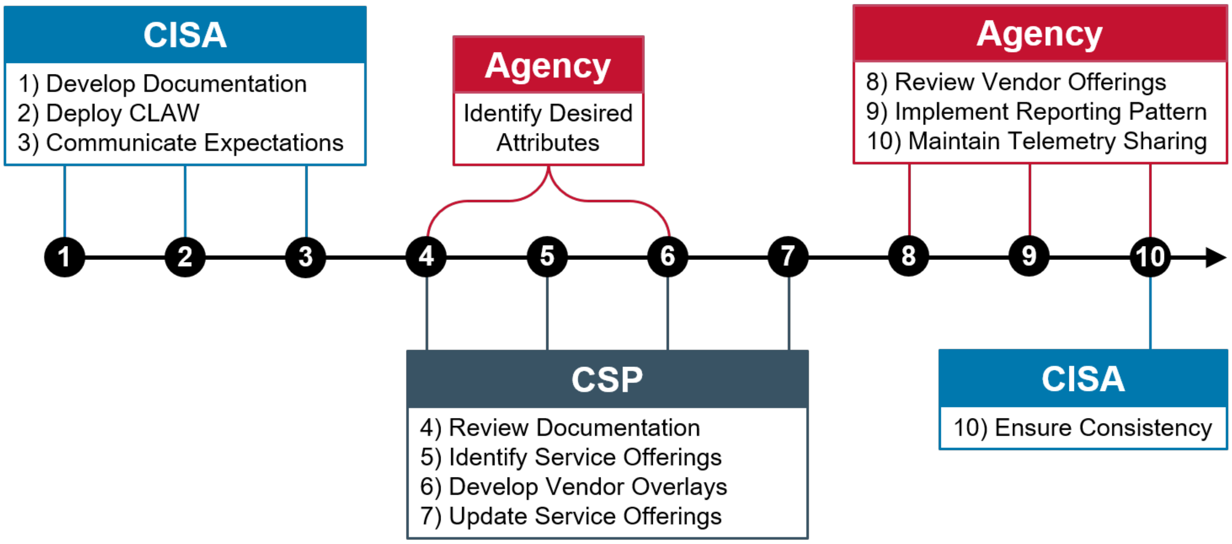

CISA
1) Develop Documentation
2) Deploy CLAW
3) Communicate Expectations
Agency
Identify Desired Attributes
Agency
8) Review Vendor Offerings
9) Implement Reporting Pattern
10) Maintain Telemetry Sharing
1
2
3
4
5
6
7
8
9
10
CSP
4) Review Documentation
5) Identify Service Offerings
6) Develop Vendor Overlays
7) Update Service Offerings
CISA
10) Ensure Consistency

Figure 10: Implementation Workflow for NCPS in the Cloud

CISA

As NCPS evolves to accommodate cloud services, CISA will have multiple roles and responsibilities in order to implement NCPS in the cloud. CISA's roles and responsibilities are as follows:

- **Develop Documentation**: CISA will enumerate generic reporting patterns and components within Volumes One and Two of this reference architecture.
- **Deploy CLAW**: CISA will deploy CLAW across a number of CSPs and regions, giving agencies options for where to deliver cloud telemetry.
- **Communicate Expectations**: CISA will communicate reporting expectations to CSPs and Agencies in a number of ways including released documentation, outreach activities and one-on-one interactions between an Agency and CISA.
- **Ensure Consistency**: CISA will work to ensure the consistency of agency cloud telemetry inputs. This will be a continuous improvement process.

CSP

As NCPS evolves to accommodate cloud services, CSPs will have multiple roles and responsibilities in order to implement NCPS in the cloud. CSP roles and responsibilities are as follows:

- **Review Documentation**: CSPs will review the cloud telemetry reporting pattern documentation in Volumes One and Two of this reference architecture.
- **Identify Service Offerings**: CSP will identify which service offerings they provide that could satisfy cloud telemetry reporting pattern options.
- **Develop Vendor Overlays**: CSP will author and publish agency guidance for utilizing service offerings in alignment with NCPS cloud telemetry reporting patterns.
- **Update Service Offerings**: If desired, the CSP vendor may modify their product offerings to align with NCPS in the cloud.

Agency

As NCPS evolves to accommodate cloud services, Agencies will have multiple roles and responsibilities in order to implement NCPS in the cloud. The Agency roles and responsibilities are as follows:

- **Identify Desired Attributes**: Agency identifies their desired options for each of the attributes for reporting cloud telemetry in a reporting pattern that matches their use case.
- **Review Vendor Offerings**: Agency will review vendor reporting pattern documentation and identify cloud service offerings that can be used to satisfy options.
- **Implement Reporting Pattern**: Agency will select, configure, and verify vendor or agency-created services to instantiate reporting patterns for sharing cloud telemetry with CISA.
- **Maintain Telemetry Sharing**: Agency will maintain telemetry sharing with CISA.