Caution Security Brief

Technical security, evidence, and framework guidance for enterprise evaluation.

CORE PROPOSITION

Replace operator assurances with verifiable evidence.

Caution helps security teams verify the software handling sensitive data, protect data all the way into a verified workload, and release secrets only to approved workload identities.

CURRENT INFRASTRUCTURE SUPPORT

Current deployments use AWS Nitro Enclaves and accept AWS Nitro attestation as an explicit trust dependency. Additional attestation backends are in active development, including Intel TDX, AMD SEV-SNP, and TPM 2.0.

GUIDE TO THE BRIEF

Contents

A concise path from the security outcome and trust boundaries to evaluation evidence, program alignment, and pilot criteria.

HOW TO USE THIS BRIEF

Start with the trust model and evidence catalogue. Use the program-alignment page to route evidence into your existing control environment, then define a bounded pilot with named owners.

EXECUTIVE SUMMARY

What security teams can verify and enforce

Caution connects an approved release to a live workload identity, then lets security teams use that identity for data protection, policy-controlled access, and independent assurance.

01

Know what software is running

Connect approved source and build inputs to an independently verifiable production identity.

02

Gate secrets on workload approval

Deny protected material when a live workload does not match the approved identity or authorization policy.

03

Protect data to the workload

Keep sensitive application payloads encrypted until they reach the verified workload.

04

Enable independent verification

Customers, auditors, and downstream systems can inspect evidence without relying only on Caution or the operator.

05

Choose the operating boundary

Use fully managed, customer-cloud, or self-hosted deployments with explicit responsibility boundaries.

HIGH-ASSURANCE USE CASES

Private AI and sensitive data processing; signing, authorization, and key systems; policy engines and trusted data feeds; and verifiable SaaS for regulated workflows.

TRUST MODEL

Architecture and trust boundaries

Infrastructure operates connectivity and lifecycle, while the verifier decides which workload identity to trust and explicitly accepts the current platform trust root.

RUNTIME AND VERIFICATION PATH

BOUNDARY 1

Verifier / customer

Selects the approved release and independently requests evidence from the running service.

BOUNDARY 2

Infrastructure boundary

Routes traffic and operates lifecycle. It controls availability and inputs but is not the source of the verification decision.

BOUNDARY 3

Confidential workload

Runs with an attested identity anchored in the currently supported platform trust root.

BOUNDARY 4

Application + EnclaveOS

Application, kernel, bootstrap, build manifest, encryption, and secret-delivery components form the measured software path.

TRUST DECISIONS

What the verifier accepts

- The currently supported AWS Nitro isolation and attestation trust root.
- A trusted verifier machine, network path, and Caution CLI build.
- The reviewed source, commits, dependencies, build inputs, and configuration.
- The security of the application code itself.

What remains outside the claim

- Application safety and output correctness.
- Availability, monitoring, response, resilience, and legal obligations.
- Infrastructure-independent plaintext protection when end-to-end application encryption is not enabled.

VERIFICATION FLOW

Independent verification from approval to production

The security team defines what it trusts, checks the live workload independently, and uses the result in release and access decisions.

01

Define the approved release

Select the reviewed source, pinned build inputs, configuration, and expected workload identity. Proprietary source can remain limited to authorized reviewers.

02

Verify the live workload

Independent tooling validates fresh platform evidence and compares the reported workload identity with the approved release.

03

Apply the decision

Use a match or mismatch in release approval, monitoring, customer assurance, and policy-controlled secret delivery.

RESULT

A verification decision your team controls

- Fresh platform evidence identifies the running workload.
- Reproduced build outputs connect that identity to reviewed source and inputs.
- A changed source, dependency, configuration, or build output produces a different identity.
- The verifier - not the operator - determines whether the evidence is acceptable.

Current implementation

Today, caution verify checks fresh platform evidence against measurements reproduced from source or an approved baseline. The verifier—not the operator—owns the identity decision.

SOFTWARE SUPPLY CHAIN

Reproducibility makes code identity meaningful

Attestation identifies a measured image. Reproducibility is what lets a verifier connect that measurement back to reviewed source and build inputs.

LAYER 1

Application

Customer source, dependencies, Containerfile, caution.hcl, and runtime configuration.

LAYER 2

Enclave image

Packaged application plus Caution runtime components and attestation manifest.

LAYER 3

EnclaveOS

Measured kernel, bootstrap, and userspace components.

LAYER 4

StageX toolchain

Deterministic, source-bootstrapped build foundations down to the kernel and toolchain.

WHAT TO REQUIRE

A reproducible application, not only a reproducible platform

Caution and EnclaveOS use StageX for deterministic, source-bootstrapped components. A customer application must also eliminate timestamps, host-specific state, unpinned inputs, and other non-determinism before source-linked verification can hold.

EVIDENCE TO INSPECT

- Attestation manifest with source URLs and commit hashes.
- Reproduced enclave image and PCR0/PCR1/PCR2 values.
- Build recipe, pinned inputs, StageX and EnclaveOS sources.
- Failure evidence after a deliberate source or configuration change.

BOUNDARY

A match establishes that the measured workload corresponds to the selected source and inputs. Use that evidence alongside code review, SAST/SCA, SBOM, dependency governance, patching, release approval, and application-specific assurance.

END-TO-END ENCRYPTED APPLICATION PATH

Protect data to the verified workload

Caution can add an application encryption layer between an approved client and the verified workload. The current implementation is STEVE; TLS remains complementary for domain and transport security.

DATA PATH

HOP 1

Approved client

The client verifies the workload key before encrypting sensitive application data.

HOP 2

Network + TLS

Standard transport security carries the additional encrypted application payload.

HOP 3

Infrastructure

Hosts and proxies transport ciphertext without needing application plaintext.

HOP 4

Verified workload

The application layer decrypts inside the approved confidential workload.

CONFIGURATION CONDITIONS

- Enable the e2e_encryption block for the HTTP service.
- Integrate the supported encryption SDK into the client path that handles sensitive data.
- Verify the workload key before encrypting application data.
- Keep the deployment outside debug mode and validate the real production path.

Ordinary TLS may terminate outside the confidential workload. Enable and validate Caution end-to-end encryption when infrastructure must not observe application plaintext.

EVALUATION TEST

Validate where plaintext exists

Send a classified test payload through the same client, proxy, host, and enclave route intended for production. Capture host and proxy traffic and inspect process memory or logs permitted by the test plan.

- Expected: infrastructure sees ciphertext, not application plaintext.
- Retain: data-flow diagram, configuration, packet captures, and test result.
- Re-test after data-path changes.

CISO DECISION

Use the end-to-end encrypted application path where the risk decision requires plaintext to remain unavailable to infrastructure operators. Document which traffic is protected, who controls the client integration, and how failures prevent sensitive data from being sent.

POLICY-CONTROLLED SECRET DELIVERY

Release secrets only after workload approval

Caution Key Services can combine workload verification with quorum approval before protected material becomes available inside the confidential workload.

APPROVAL POLICY | KEYMAKER

Create a governed quorum

- Generate a master secret and split it into Shamir shards.
- Encrypt each shard to an authorized shard holder's OpenPGP key.
- Store the public keyring, encrypted shards, threshold, and recipient public key in the quorum bundle.
- Encrypt application secrets to the quorum recipient key before committing them.

Production shard holders should use governed, protected private keys. The development helper that writes plaintext private keyrings is explicitly unsuitable for production.

VERIFIED DELIVERY | LOCKSMITH

Unlock only after workload verification

01 Shard holder requests fresh platform evidence and a workload-bound ephemeral key.

02 The CLI validates the evidence and compares the live identity with the approved baseline.

03 An approved holder sends a signed shard encrypted through ephemeral ECDH key exchange.

04 Locksmith verifies the signer and counts the configured quorum inside the enclave.

05 After quorum, the master secret is reconstructed briefly in enclave memory and derives keys used to decrypt application secrets.

Current implementation note: caution secret send-shard presently requires the host-toolchain CLI build. Evaluate and govern that local supply-chain dependency for production shard holders.

CONTROL OBJECTIVE

Make secret availability depend on two explicit facts: the live workload matches an approved identity, and the configured number of authorized participants approve that specific running state. Test both an identity mismatch and insufficient authorization.

SECURITY EVIDENCE

Evidence catalogue and operational use

Treat technical outputs as inputs to a control process: define who collects them, when they are reviewed, where they are retained, and what happens on failure.

EVIDENCE	WHAT YOUR TEAM CAN VERIFY	HOW YOUR TEAM CAN USE IT
Software identity and provenance	Approved source and build inputs, expected measurements, fresh verification result, reproduced output, and measured software components.	Architecture and provenance review, release baseline, change approval, periodic integrity checks, and mismatch investigation.
Protected data-path evidence	Where application encryption starts and ends, which workload identity receives the data, and which infrastructure can observe plaintext.	Confidentiality and privacy review, production-path validation, regression testing, and customer assurance.
Workload-bound secret delivery	Approved identity, authorized participants, signatures, policy threshold, and both allowed and denied release outcomes.	Access gates, separation-of-duties review, negative testing, recovery exercises, and incident investigation.

OPERATIONAL PATTERN

Collect on every approved release; run again on a defined schedule and after relevant infrastructure or configuration changes; send failures into incident and change workflows; retain the evidence with workload, source commit, verifier identity, timestamp, and decision owner.

OPERATING MODELS

Choose the deployment boundary

The verification model remains consistent, but infrastructure ownership, operational duties, and evidence collection differ by deployment model.

MODEL 1

Fully managed

Caution-managed infrastructure

CAUTION MANAGES

- Infrastructure and public ingress
- Enclave lifecycle
- Deployment orchestration

CUSTOMER MANAGES

- Application and configuration
- Source and dependency security
- Data classification and governance

MODEL 2

Bring your own compute

Customer cloud account (AWS today)

CAUTION MANAGES

- Builds and deployment orchestration
- Enclave lifecycle in the customer account

CUSTOMER MANAGES

- Cloud account and billing
- Network and region boundaries
- Cloud governance and monitoring

MODEL 3

Self-hosted

Customer-operated environment

CAUTION MANAGES

- Open-source software availability

CUSTOMER MANAGES

- Platform operations
- Infrastructure and deployments
- Upgrades, support, monitoring, and resilience

SELECTION CRITERIA

Choose the model based on the infrastructure boundary, network and region controls, operational ownership, evidence-retention design, source visibility, and the team's ability to run the platform. Customer-cloud deployments are currently available in AWS accounts. Verify commercial, support, SLA, data-residency, and incident terms during procurement; this brief does not assert them.

PROGRAM ALIGNMENT

How Caution evidence supports existing security programs

A focused review identified 21 credible technical evidence relationships across four security programs: 2 direct, 4 configuration-dependent, and 15 supporting. These relationships show where Caution can contribute useful evidence to an existing control environment.

PROGRAM	RELEVANT SCOPE	EVIDENCE RELATIONSHIP	CUSTOMER / ASSESSOR
NIST CSF 2.0	6 relevant outcomes	2 direct / 1 configuration-dependent / 3 supporting	Select expected identity; operate configuration, monitoring, response, and recovery
ISO/IEC 27001:2022 + Amd 1:2024	7 relevant Annex A controls	1 configuration-dependent / 6 supporting	ISMS, risk treatment, SoA, control operation, audit and certification
SOC 2 TSC (2017; points of focus revised 2022)	4 relevant criteria	1 configuration-dependent / 3 supporting	System description, complete control design, period evidence, auditor evaluation
NIS2 Directive	4 relevant Article 21 measures	1 configuration-dependent / 3 supporting	Applicability, governance, proportionality, reporting, resilience, accountability
ITAR	Secondary context; not included in 21	Potentially relevant protected-path and access evidence	Export counsel, authorization, persons, countries, cryptography, and records
NIST SP 800-171 Rev. 3	Secondary context; not included in 21	Potentially relevant integrity, configuration, and supply-chain evidence	CUI scope, SSP, requirement implementation, assessment, and remediation
CMMC	Secondary context; not included in 21	Potential technical evidence adjunct only	CMMC scope, assessment, Plan of Action rules, and certification decision
SLSA v1.2	Secondary context; not included in 21	Reproducible comparison and source/build manifest may support evaluation	Formal provenance, signing, builder/source controls, and level assessment

Method: direct means the current capability produces evidence closely aligned to the technical outcome; configuration-dependent requires a named capability or workflow; supporting contributes to a broader customer-operated control. The focused review considered 25 exact references and excluded four that depend on organizational monitoring or information-classification processes. This is not framework coverage, certification, or controls satisfied.

PRIMARY MAPPING DETAIL

NIST CSF 2.0: where Caution helps

Six outcomes have a credible relationship to current Caution evidence: two direct, one configuration-dependent, and three supporting. Each mapping below shows how a security team can use the evidence and what it must operationalize.

REFERENCE	RELATIONSHIP	CAUTION EVIDENCE AND BOUNDARY	CUSTOMER / ASSESSOR
GV.SC-06	Supporting	Independent verification and source-linked build evidence can inform supplier due diligence before use.	Own the complete supplier assessment, contracts, risk decision, and approval.
GV.SC-07	Supporting	Retained verification evidence can inform ongoing assessment of a supplier product or service.	Monitor the supplier relationship, reassess risk, and respond to changes over time.
ID.RA-09	Direct	Fresh signed attestation and comparison with a verifier-selected identity directly assess software authenticity and integrity before use.	Select or approve the expected source, commit, or measurements; assess whether the software is safe.
PR.DS-02	Configuration-dependent	End-to-end encryption can protect request and response data into the workload when enabled and integrated by the client.	Classify data, validate the complete path, and protect traffic not using the encrypted application path.
PR.DS-10	Direct	Enclave isolation directly supports confidentiality of data in use from the infrastructure operator.	Own application integrity, availability, correctness, vulnerability management, and data handling.
PR.PS-01	Supporting	Measured release identity and mismatch evidence support configuration baselines and unexpected-change detection.	Establish configuration policy, approvals, inventory, drift handling, and remediation.

Interpret each relationship together with its condition and customer-owned responsibility. A mapping is not a claim that the referenced control, criterion, measure, certification, or regulatory obligation is satisfied.

PRIMARY MAPPING DETAIL

ISO/IEC 27001: where Caution helps

Seven Annex A references have a credible evidence relationship: one configuration-dependent and six supporting. Caution evidence can strengthen the technical record used within a customer-operated ISMS.

REFERENCE	RELATIONSHIP	CAUTION EVIDENCE AND BOUNDARY	CUSTOMER / ASSESSOR
A.5.21	Supporting	Source-linked build and workload-identity evidence can support ICT supply-chain assurance activities.	Operate supplier governance, dependency review, contractual controls, and vulnerability remediation.
A.5.23	Supporting	The documented trust boundary and deployment evidence can support cloud-service risk evaluation.	Govern cloud acquisition, use, change, exit, IAM, monitoring, resilience, and contractual requirements.
A.8.2	Supporting	Enclave isolation reduces infrastructure privilege over workload memory; Key Services can further restrict secret access.	Operate IAM, PAM, access approval, review, segregation of duties, and break-glass controls.
A.8.9	Supporting	The manifest, build inputs, measurements, and mismatch result provide reviewable configuration-integrity evidence.	Define approved baselines, change procedures, monitoring frequency, and remediation.
A.8.24	Configuration-dependent	End-to-end encryption and Key Services provide evidence only when the relevant capability is configured and validated.	Own cryptographic policy, algorithm approval, key lifecycle, rotation, recovery, and legal requirements.
A.8.25	Supporting	Reproducible verification strengthens build and release integrity within a secure development lifecycle.	Own secure design, code review, testing, dependency governance, vulnerability handling, and release approval.
A.8.32	Supporting	Workload identity comparison can detect deployed configuration changes and tie a release to reviewed inputs.	Authorize, test, document, approve, and monitor normal and emergency changes.

Interpret each relationship together with its condition and customer-owned responsibility. A mapping is not a claim that the referenced control, criterion, measure, certification, or regulatory obligation is satisfied.

PRIMARY MAPPING DETAIL

SOC 2 and NIS2: where Caution helps

Four SOC 2 criteria and four NIS2 measures have a credible evidence relationship: two configuration-dependent and six supporting. The mappings focus on technical evidence that can support broader organization-operated controls.

REFERENCE	RELATIONSHIP	CAUTION EVIDENCE AND BOUNDARY	CUSTOMER / ASSESSOR
CC6.1	Supporting	Enclave isolation and attestation-aware secret delivery can support logical-access boundaries.	Own IAM lifecycle, authorization, PAM, physical access, and periodic access reviews.
CC6.6	Supporting	The workload boundary and protected data path can support restrictions on access across system boundaries.	Own network architecture, authentication, firewalling, endpoint security, and boundary monitoring.
CC6.7	Configuration-dependent	Evidence is relevant when end-to-end encryption or Key Services protects the applicable transmission or secret flow.	Classify information and operate handling, transmission, removal, key-management, and disposal controls.
CC8.1	Supporting	Verification supports deployed configuration integrity and detection of unexpected changes only.	Authorize, design, test, approve, implement, and document the complete change lifecycle.
21(2)(d)	Supporting	Provenance and workload verification can support technical supply-chain assurance.	Own supplier governance, contracts, risk assessment, monitoring, and response.
21(2)(e)	Supporting	Build and deployment integrity evidence can support secure acquisition, development, and maintenance.	Own secure development, vulnerability handling and disclosure, patching, and maintenance.
21(2)(f)	Supporting	Repeatable positive and negative verification tests can contribute to effectiveness assessment.	Define measures, testing frequency, governance review, remediation, and evidence retention.
21(2)(h)	Configuration-dependent	Encryption evidence is relevant only when the required end-to-end or key capability is configured and validated.	Own cryptographic policies, proportionality, key governance, lifecycle, and regulatory conclusions.

Interpret each relationship together with its condition and customer-owned responsibility. A mapping is not a claim that the referenced control, criterion, measure, certification, or regulatory obligation is satisfied.

COMMON EVALUATION QUESTIONS

Technical answers for security review

These answers describe current documented functionality. Commercial commitments, service levels, audit reports, support terms, and roadmap capabilities must be confirmed separately.

Which infrastructure is supported today?

Current deployments use AWS Nitro Enclaves. Additional attestation backends are in active development, including Intel TDX, AMD SEV-SNP, and TPM 2.0.

How can a workload be verified?

Reproduce expected measurements from source and build inputs, or compare fresh live evidence with a known approved baseline.

Does verification require public source?

No. Public source enables open verification; authorized customers or auditors can verify private repositories or local source without publishing proprietary code.

How is freshness and integrity checked?

The verifier challenges the running service with a fresh nonce, validates signed platform evidence and its certificate chain, and compares measured workload identity.

What is required for production verification?

Use a production, non-debug deployment and a customer-trusted verifier with approved source or known measurements. Debug identities are not meaningful production evidence.

Must the application build be reproducible?

Yes for source-linked verification. Caution platform components are reproducible; each customer must make its application build deterministic.

How is sensitive application data protected?

An optional end-to-end application encryption layer can keep payloads encrypted until the verified workload. Client integration is required; TLS remains complementary.

How are secrets released?

Policy-controlled delivery can require an approved workload identity and quorum authorization before protected material becomes available inside the workload.

What evidence can customers retain?

Software-identity baselines, verification results, signed platform evidence, reproduced build artifacts, protected-path tests, and secret-release decisions.

What remains customer-owned?

Application security, IAM, governance, monitoring, vulnerability remediation, incident response, resilience, evidence retention, and legal or regulatory conclusions.

RECOMMENDED PROCUREMENT FOLLOW-UP

Request a workload-specific architecture, shared-responsibility matrix, evidence-retention design, support and incident terms, and a pilot that includes negative verification, plaintext-path, and denied secret-release tests.

EVALUATION PLAN

Pilot criteria and required assumptions

Use a bounded workload to establish whether the trust model, evidence, and operational responsibilities are useful in your environment before production adoption.

PILOT ACCEPTANCE CRITERIA

- Select one representative sensitive workload with bounded blast radius and named business and security owners.
- Make the application build reproducible and verify the live workload identity from a customer-controlled machine.
- Confirm a changed source or configuration produces a mismatch and blocks the intended trust decision.
- If end-to-end application encryption is required, prove infrastructure components cannot observe application plaintext.
- If policy-controlled secret delivery is required, test an identity mismatch, unauthorized participant, and insufficient authorization before a successful release.
- Export evidence into the chosen SIEM, GRC, or audit repository and assign review and retention owners.
- Record residual-risk acceptance before production use.

DECISION OWNERS

Name the people who can accept the evidence and the residual risk

- Application owner: release scope, source baseline, dependency risk, and remediation.
- Security owner: threat model, verification policy, protected-data path, and incident response.
- Platform owner: deployment boundary, verifier operation, monitoring, and resilience.
- Risk or compliance owner: evidence retention, program mapping, assessor engagement, and residual-risk acceptance.

REQUIRED ASSUMPTIONS

Reviewed source or an approved workload identity; reproducible application build; production configuration; trusted verifier; the current platform trust root; end-to-end application encryption where infrastructure must not observe plaintext; and customer-owned application security and governance.

PRODUCTION DECISION

Proceed only when the evidence changes a real control decision

Define what a match enables, what a mismatch blocks, who investigates, and how evidence is retained. If verification is merely observed but cannot affect release, data handling, or secret delivery, the pilot has not yet operationalized the trust model.

SOURCES AND CONTACT

References and next steps

Product claims and program mappings in this brief should be validated against the current Caution documentation, official standards, your architecture, and qualified assessor or legal guidance.

Caution: What is Caution?
https://docs.caution.co/what-is-caution/

Caution: Attestations
https://docs.caution.co/concepts/attestation/

Caution: Security assumptions
https://docs.caution.co/concepts/security-assumptions/

Caution: Reproducibility
https://docs.caution.co/concepts/reproducibility/

Caution: Encryption
https://docs.caution.co/concepts/encryption/

Caution: Key Services
https://docs.caution.co/concepts/key-services/

Caution: Deployment models
https://docs.caution.co/reference/deployment-models/

EnclaveOS technical overview
https://distrust.co/blog/enclaveos.html

AWS Nitro Enclaves: Root of trust
https://docs.aws.amazon.com/enclaves/latest/user/verify-root.html

NIST CSF 2.0
https://csrc.nist.gov/pubs/cswp/29/final

ISO/IEC 27001:2022
https://www.iso.org/standard/27001

ISO/IEC 27001:2022/Amd 1:2024
https://www.iso.org/standard/88435.html

AICPA SOC and Trust Services Criteria
https://www.aicpa-cima.com/soc

NIS2 Directive
https://eur-lex.europa.eu/eli/dir/2022/2555/oj/eng

NIS2 Implementing Regulation 2024/2690
https://eur-lex.europa.eu/eli/reg_impl/2024/2690/oj/eng

ITAR 22 CFR 120.54, 120.55 and 120.56
https://www.ecfr.gov/current/title-22/part-120

NIST SP 800-171 Rev. 3
https://csrc.nist.gov/pubs/sp/800/171/r3/final

CMMC
https://dodcio.defense.gov/CMMC/About/-DoD/

SLSA v1.2
https://slsa.dev/spec/v1.2/tracks

REVIEW A CANDIDATE WORKLOAD

Bring the workload, threat model, deployment boundary, and evidence requirements to a Caution security engineer.

caution.co/contact.html