SECURE CLOUD ACCESS · ZERO TRUST · SSE

SSCA 300-740

Complete Learning Guide

Designing and Implementing Secure Cloud Access for Users and Endpoints · v2.0

UNOFFICIAL STUDY GUIDE

JOZEF BAROŠ

Designing and Implementing Secure Cloud Access for Users and Endpoints

SSCA 300-740 v2.0

Complete Learning Guide

An independent, unofficial study guide

Jozef Baroš

Disclaimer and Notices

This book is an **independent, unofficial study guide** for the Cisco examination *Designing and Implementing Secure Cloud Access for Users and Endpoints v2.0 (SSCA 300-740)*. It is not endorsed by, affiliated with, sponsored by or produced in cooperation with Cisco Systems, Inc.

Trademarks. Cisco, the Cisco logo, CCNP, Cisco Secure Access, Cisco Secure Workload, Cisco Secure Firewall, Cisco Duo, Cisco XDR, Meraki, Catalyst, ThousandEyes and all other Cisco marks are trademarks or registered trademarks of Cisco Systems, Inc. and its affiliates. Splunk is a trademark of Splunk Inc. All other trademarks are the property of their respective owners. They are used here for identification and educational purposes only.

Exam version. This guide targets **v2.0** of the 300-740 blueprint, which became the live version of the examination on 27 August 2026. Cisco may revise exam topics at any time and without notice. Always confirm the current blueprint on Cisco's own certification pages before booking, and treat any discrepancy between this book and the published blueprint as authoritative in Cisco's favour.

Configurations and examples. Every configuration, command, policy fragment and log extract in this book is **illustrative**. Product interfaces, command syntax, object names and dashboard layouts change between releases. Nothing here should be applied to a production system without testing in an environment you control, and no example should be treated as a deployment specification.

Scenarios. The situations described in the *From the Field* boxes and in the eight cases are composites drawn from patterns that recur across organisations. Names, sectors and identifying details have been changed or invented, and no real organisation is described.

No warranty. This book is provided as-is, without warranty of any kind. The author accepts no liability for any loss or damage arising from its use, including any outcome of an examination attempt.

About This Book

This guide exists because the 300-740 blueprint covers an unusually wide surface — zero-trust theory, cloud platform security, identity and multifactor authentication, data protection, artificial intelligence security, workload segmentation and security operations — and because almost nobody arrives already comfortable in all of it. A network engineer knows tunnels and struggles with SAML assertions. A security analyst knows MITRE ATT&CK and has never labelled a workload. Both need the same book, starting from different places.

Who it is for

You are preparing for the SSCA 300-740 exam as a CCNP Security concentration, or you are implementing Cisco Secure Access and want the reasoning as well as the configuration. The book assumes CCNA-level networking, basic familiarity with TLS and certificates, and general awareness of what a public cloud is. Everything else — zero-trust vocabulary, the SSE model, Duo, Secure Workload, eBPF, AI security — is built from first principles.

How each sub-section is built

The structure repeats deliberately, so that after the first few sub-sections you know where to find what you need.

- **Explanation from first principles**, with every term defined at first use.
- **A diagram** wherever the relationship is spatial, sequential or easier seen than described — 48 of them across the book.
- **Configuration**: dashboard and policy walkthroughs, and real command-line syntax wherever a command line exists.
- **Four callout types** — *Did you know?* for context worth having, *Exam Tip* for how the topic is tested, *Common Pitfall* for what goes wrong, and *From the Field* for what actually happened to somebody.
- **Summary and Key Takeaways**, written so that reading only these two boxes across the whole book is a viable final revision pass.
- **Five knowledge-check questions**, with explanations of why the right answer is right *and* why the plausible wrong ones are wrong — 260 questions in total.

How to use it

Read in order the first time. The sections build on each other: Section 2's identity plane is what Section 3's policies evaluate, and Section 4 configures what Section 3 describes. On a second pass, use the Summary and Key Takeaways boxes as the spine and return to full text only where a knowledge check

exposes a gap. The eight cases at the end are revision in narrative form — each one names the sub-sections it draws on, so a case you cannot follow points at a topic worth rereading.

> ✅ **Exam Tip — What the verbs in the blueprint tell you**
>
> Cisco's topic wording is a reliable guide to depth. *Describe* means you must be able to explain a concept and distinguish it from adjacent ones. *Determine* and *select* mean you will be given a requirement and asked which control or tool expresses it. *Implement*, *configure* and *integrate* mean the object names and their relationships are fair game. Sections 1, 3 and 5 lean on the first two; Sections 2 and 4 lean on the third.

How This Book Maps to the Exam

Every topic in the published v2.0 blueprint is covered by at least one sub-section. Domain weightings are shown as Cisco publishes them; the number of sub-sections is roughly proportional, with extra depth where the topic demands it.

Blueprint topic	Weight	Sub-sections
1.1 NIST SP 800-207 and CISA ZTMM v2.0	10%	1.1, 1.2
1.2 MITRE ATT&CK cloud frameworks		1.3
1.3 Public cloud security requirements (a–d)		1.4, 1.5
1.4 Public cloud operational requirements (a–c)		1.6
1.5 Private cloud, hypervisors, Kubernetes		1.7
1.6 eBPF tooling — Cilium, Tetragon		1.7
2.1 Identity intelligence across IdPs	20%	2.2
2.2 User and device authentication via certificates		2.3, 2.4
2.3 MFA including phishing-resistant, with Duo		2.5, 2.6
2.4 Endpoint posture policies		2.7
2.5 User and device trust via SAML, mobile and web		2.8
2.6 SSO and provisioning — SCIM and SAML		2.9
3.1 Encryption in transit and at rest	30%	3.1
3.2 IPS, DLP and malware protection		3.2, 3.3
3.3 AI Defense		3.4
3.4 AI Access and AI Guardrails		3.5
3.5 Web application firewalls and DDoS		3.6
3.6 Network security edge policy (SSE, SD-WAN)		3.7, 3.8, 3.9, 3.10
3.7 Secure Workload enforcement (a–e)		3.11–3.15
4.1 DNS security	30%	4.2

Blueprint topic	Weight	Sub-sections
4.2 Secure Web Gateway		4.3
4.3 Data Loss Protection		4.4
4.4 CASB		4.5
4.5 Private access for workloads (a–b)		4.6, 4.7, 4.8
4.6.a VPN as a service with ISE		4.9
4.6.b Digital experience monitoring		4.13
4.6.c Zero-trust access, QUIC and MASQUE		4.10, 4.11, 4.12
5.1 Tool selection for visibility and enforcement	10%	5.1
5.2 Traffic flow and telemetry interpretation		5.2, 5.3
5.3 Secure Access dashboards		5.4
5.4 XDR and Splunk integration		5.5, 5.6

Sub-sections 4.1 and 4.14, and 2.10 and 3.7, are not numbered topics in the blueprint. They cover the architecture and the ordered troubleshooting methods on which the numbered topics depend, and they are included because the surrounding material is considerably harder to apply without them.

Contents

Section 1 — Concepts

Every technology in this book — Cisco Secure Access, Duo, Secure Workload, Multicloud Defense, AI Defense — is an answer to a question somebody asked first. This section is about the questions. It covers the frameworks that define what secure cloud access is supposed to achieve, the attacker behaviour those frameworks are defending against, and the properties of public and private cloud platforms that make the defence different from anything you would have built in a data centre ten years ago.

The domain carries **10 percent** of the exam, which understates its importance. Concepts questions are cheap marks — they are almost entirely 'describe' and 'select' items with no configuration to remember — and the vocabulary established here is used in the question stems of every other domain. If you do not know what a policy enforcement point is, a question in Domain 3 that asks where to place one is unanswerable even if you know the product perfectly.

The seven sub-sections move from the abstract to the concrete. We start with the architectural definition of zero trust, then measure an organisation against it, then look at what attackers actually do to cloud identities and workloads, and finish with the platforms — public cloud, private cloud, Kubernetes — that everything else runs on.

Sub-section	Focus	What the exam asks
1.1 NIST SP 800-207	Tenets, PE / PA / PEP, deployment approaches	Which component makes the decision, which enforces it
1.2 CISA ZTMM v2.0	Five pillars, four maturity stages	Which stage a described capability represents
1.3 MITRE ATT&CK cloud	IdP, SaaS and IaaS techniques	Select the mitigation for a described technique
1.4 Cloud security requirements	IAM, shared responsibility, posture	Who is responsible for a given layer
1.5 Cloud-native controls	Security groups, NACLs, Multicloud Defense	Which construct enforces the stated requirement
1.6 Cloud operations	Telemetry, infrastructure, consumption models	Which log answers the investigative question
1.7 Private cloud and eBPF	Hypervisors, Kubernetes, Cilium, Tetragon	What eBPF observes and enforces, and why

✅ **Exam Tip — Concept questions are language questions**

In this domain the exam is testing whether you use the vocabulary the way the standards use it. 'Policy engine', 'implicit trust zone', 'phishing-resistant', 'shared responsibility', 'optimal maturity' all have precise definitions, and the distractors are usually written to be correct-sounding statements that misuse one of those words. Read the option for the word, not for the sentiment.

🌍 **From the Field — Why this domain exists at all**

The v2.0 blueprint added the concepts domain because Cisco found engineers deploying Secure Access as 'a cloud proxy' — migrating a web filtering policy into the cloud and calling the project zero trust. The result is a faster perimeter, not a different security model. Every sub-section here exists to stop you shipping that project.

1.1 Zero Trust from First Principles — NIST SP 800-207

Zero trust is the most over-marketed phrase in enterprise security, which makes it easy to forget that it has an actual, published, vendor-neutral definition. **NIST Special Publication 800-207**, *Zero Trust Architecture*, was published in August 2020 and remains the reference every other framework points back to — including Cisco's own architecture and the CISA maturity model in the next sub-section. It is thirty-odd pages long, it names no products, and it is worth reading in full at least once. This sub-section covers what it says and how Cisco's portfolio maps onto it.

The problem statement

Classical network security grants access on the basis of location. A device that is inside the building, on a corporate VLAN, or on the far end of a VPN tunnel is treated as trustworthy; a device outside is not. That model made sense when applications lived in a data centre the organisation owned and employees sat in buildings the organisation leased. It produces a large **implicit trust zone** — the region of the network where any host can reach any other host because no control sits between them.

Two changes broke the model. Applications moved to SaaS and public cloud, so the resource is no longer inside the perimeter. Users moved everywhere, so the subject is no longer inside it either. What remains is an implicit trust zone that protects assets which have mostly left it, while the workforce connects around it. Meanwhile attackers learned that once a single credential or endpoint inside that zone is compromised, lateral movement is nearly free.

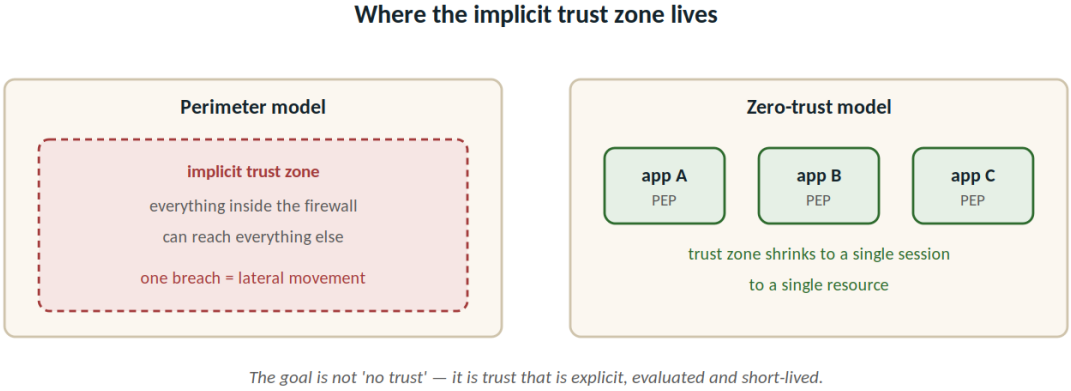

Figure 1.2 — The perimeter model creates one large implicit trust zone; zero trust shrinks it to a single session against a single resource.

NIST frames zero trust as the response: stop granting access to networks, and start granting access to resources, one request at a time, on evidence that is re-evaluated continuously. The phrase the document uses is worth memorising — zero trust is **a set of guiding principles for workflow, system design and operations**, not a product or an architecture you can buy.

The seven tenets

Section 2.1 of SP 800-207 lists seven tenets. They are abstract on purpose, and the exam draws distractors from mis-statements of them, so read each one for what it excludes as well as what it requires.

1. **All data sources and computing services are resources.** A resource is anything an organisation wants to protect — a SaaS tenant, an API, a Kubernetes pod, a printer, a personally owned device accessing corporate data.
2. **All communication is secured regardless of network location.** Being on the corporate LAN earns a session nothing. The same authentication and encryption apply inside the building as in an airport.
3. **Access to individual enterprise resources is granted on a per-session basis.** Trust is evaluated before each session and grants the least privilege needed to complete that task. Access to one resource does not imply access to another.
4. **Access is determined by dynamic policy** — client identity, application or service, the requesting asset's observable state, and behavioural and environmental attributes.
5. **The enterprise monitors and measures the integrity and security posture of all owned and associated assets.** No device is inherently trusted; posture is measured, and a device that falls out of compliance is treated differently.
6. **All resource authentication and authorisation are dynamic and strictly enforced before access is allowed.** This is a constant cycle of obtaining access, scanning and assessing threats, adapting, and continually re-evaluating trust — not a single check at login.
7. **The enterprise collects as much information as possible about the current state of assets, network infrastructure and communications, and uses it to improve its security posture.** Telemetry is not a by-product of zero trust; it is an input to it.

> ⚠️ **Common Pitfall — "Zero trust" does not mean "no trust"**
>
> The model does not remove trust; it removes *implicit* trust. Every tenet describes trust that is explicit, evidence-based, scoped to one session and constantly re-checked. An exam option that says a zero-trust architecture 'never trusts any user or device' is describing something that could not grant access at all. The accurate phrasing is that trust is never granted on the basis of network location and never assumed to persist.

The logical components

SP 800-207 splits the architecture into a control plane and a data plane, and names three logical components. They are logical: one product can implement several, and one component can be distributed across several products.

- **Policy Engine (PE)** — makes the decision. It takes the request plus input from every supporting data source, runs the trust algorithm, and produces grant, deny or revoke. The PE is the brain.
- **Policy Administrator (PA)** — executes the decision. It establishes or tears down the communication path, and issues the session credential or token the subject uses to reach the resource. The PA is the hands.
- **Policy Enforcement Point (PEP)** — sits in the data path, enables and terminates the connection, and monitors it for the life of the session. It is the only component the traffic actually passes through.

PE and PA together form the **policy decision point (PDP)**, and both live in the control plane. The PEP lives in the data plane. That distinction — who decides, who instructs, who is in the path — is the most reliably tested fact in the whole framework.

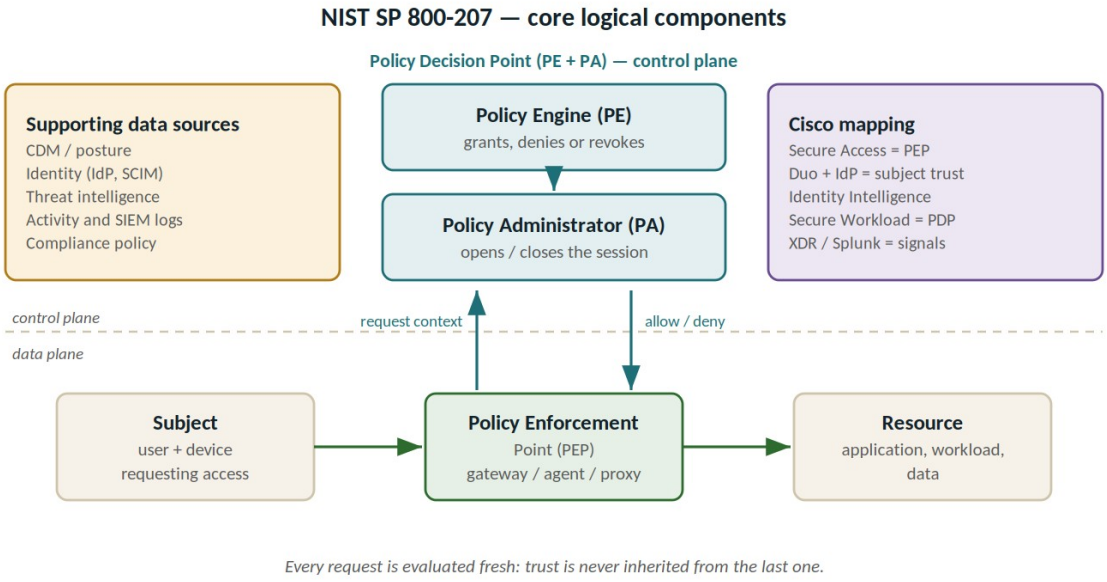

Figure 1.1 — The core logical components, the supporting data sources that feed the policy engine, and the Cisco products that play each role.

Around the PDP sit the **supporting data sources** the trust algorithm consumes: a continuous diagnostics and mitigation (CDM) system reporting asset posture, an identity management system and its directory, threat intelligence feeds, activity logs and SIEM data, data access policies, an enterprise PKI, and industry compliance requirements. The quality of a zero-trust decision is capped by the quality of these inputs — an architecture with a perfect PEP and no posture data can only make identity-based decisions.

The trust algorithm

The trust algorithm is the process the policy engine uses to reach a verdict. NIST describes two axes along which implementations differ, and the vocabulary appears in exam stems.

- **Criteria-based versus score-based.** Criteria-based evaluation requires a set of qualifying attributes to be satisfied before access is granted — a binary gate. Score-based evaluation computes a confidence level from weighted inputs and compares it against a threshold, which allows graduated responses such as 'allow, but step up authentication'.
- **Singular versus contextual.** A singular algorithm judges each request in isolation. A contextual algorithm considers subject history — what this user has been doing, from where, at what pace — which catches an attacker using valid credentials but produces false positives when legitimate behaviour changes.

Cisco's implementation is contextual and score-influenced. Duo's Risk-Based Authentication raises the required factor strength when a session looks anomalous rather than simply denying it, and Secure Access evaluates identity, device posture and destination together for every connection. Expressed as policy logic, a decision reads like this:

A per-session decision, written as the policy engine evaluates it

```
IF   subject.identity        IS verified by IdP (SAML assertion, signed, unexpired)
AND  subject.mfa             IS phishing-resistant (WebAuthn / FIDO2)
AND  device.managed          IS true          # certificate present, MDM-enrolled
AND  device.posture          IS compliant     # disk encrypted, EDR running, OS
patched
AND  request.destination     IN resource_group("finance-erp")
AND  request.risk_score      <  threshold      # behavioural + threat-intel inputs
THEN grant session TO finance-erp FOR 8h WITH continuous re-evaluation
ELSE IF device.posture IS non_compliant
     THEN grant session TO remediation_portal ONLY
ELSE deny AND log AND notify
```

Read that block again with the components in mind. The policy engine evaluated the conditions. The policy administrator issued the eight-hour session credential. The policy enforcement point — in Cisco's case, the Secure Access edge or a resource connector — is what the packets traverse, and what tears the session down when continuous re-evaluation flips the posture attribute to non-compliant.

> 💡 **Did you know? — The remediation path is part of the design**
>
> Notice the middle branch. A mature zero-trust policy rarely has only two outcomes. Denying a non-compliant laptop outright generates a help-desk ticket; sending it to a remediation portal that explains which check failed converts the same event into self-service. NIST calls this out explicitly — the policy engine may grant access to a restricted resource set rather than making a binary decision.

Deployment approaches and variations

SP 800-207 describes three approaches to building the architecture. Real deployments blend them, but each has a centre of gravity.

Approach	Centre of gravity	Typical Cisco realisation
Enhanced identity governance	Identity is the primary policy input; resource access follows user and device identity attributes	Duo, an IdP integration, Identity Intelligence, SCIM provisioning
Micro-segmentation	Resources are placed behind fine-grained gateways or host agents that police east-west traffic	Secure Workload agents, Multicloud Defense, Cilium network policy
Network infrastructure and software-defined perimeters	The network overlay itself brokers every connection and hides resources until authorised	Secure Access ZTNA, resource connectors, Secure Client with MASQUE

NIST also names four deployment **variations** describing where the enforcement point physically sits. The device agent/gateway model puts an agent on the endpoint and a gateway in front of the resource. The enclave gateway model fronts a group of resources with one gateway — useful for legacy systems that cannot host an agent. The resource portal model puts a single portal in front of everything, requiring no endpoint software but yielding no device posture. Device application sandboxing runs approved applications in isolated compartments on the endpoint. You will meet all four again in Section 4: the client-based ZTA module is an agent/gateway deployment, clientless ZTA is a resource portal, and a resource connector in front of a private subnet is an enclave gateway.

⚠️ **Common Pitfall — Buying a ZTNA product does not produce a zero-trust architecture**

ZTNA replaces the VPN for the user-to-application path. That is one tenet, applied to one traffic pattern. If service accounts still hold static keys, if east-west traffic between workloads is unfiltered, and if posture is checked once at enrolment, the architecture is a perimeter with a better front door. When an exam question offers 'deploy ZTNA' as the answer to a lateral-movement or workload-to-workload problem, look for the micro-segmentation option instead.

🌍 **From the Field — What the first ninety days usually look like**

In practice almost nobody starts with micro-segmentation, because it requires an application dependency map nobody has. The common sequence is: consolidate identity into one IdP, add phishing-resistant MFA and device certificates, put ZTNA in front of the ten applications that justify the VPN, decommission the VPN for those applications, and only then start segmenting workloads with the flow data you have accumulated. Each step reduces the implicit trust zone without needing the next step to be finished.

Summary

NIST SP 800-207 defines zero trust as a set of principles rather than a product: no access is granted on the basis of network location, every session is authenticated and authorised individually against

dynamic policy, and posture is continuously re-evaluated. The architecture is expressed in three logical components — a policy engine that decides, a policy administrator that establishes or revokes the session, and a policy enforcement point in the data path — with the engine and administrator forming the control-plane policy decision point. The engine's verdict is only as good as the supporting data sources feeding it: identity, posture, threat intelligence and activity logs. Organisations reach the model through enhanced identity governance, micro-segmentation, or a software-defined perimeter, and in Cisco terms those map to Duo and identity integrations, Secure Workload, and Secure Access respectively.

Key Takeaways

- Zero trust removes **implicit** trust, not all trust — access is explicit, per-session, least-privilege and re-evaluated.
- **PE decides, PA establishes or revokes the session, PEP sits in the data path.** PE + PA = the policy decision point (control plane); the PEP is the data plane.
- The seven tenets treat everything as a resource, secure all communication regardless of location, grant per-session access, use dynamic policy, measure asset posture, enforce authentication and authorisation before access, and feed telemetry back into policy.
- Trust algorithms are **criteria-based or score-based**, and **singular or contextual**; Cisco's stack is contextual, which is what allows step-up authentication instead of a flat deny.
- The three deployment approaches are **enhanced identity governance, micro-segmentation, and network infrastructure / software-defined perimeter** — ZTNA alone covers only the user-to-application path.

Knowledge Check — 1.1 Zero Trust from First Principles

Q1. In a NIST SP 800-207 architecture, which component establishes and tears down the communication path and issues the session credential?

A. The policy engine (PE)

B. The policy enforcement point (PEP)

C. The policy administrator (PA)

D. The continuous diagnostics and mitigation (CDM) system

Correct answer: C. The policy administrator executes the policy engine's verdict by establishing, refusing or revoking the session and issuing the credential or token the subject uses. The policy engine makes the decision but does not act on the path; the PEP sits in the data plane and passes the traffic; CDM is a supporting data source that reports asset posture.

Q2. A security architect states that a zero-trust architecture must 'never trust any user or device'. Why is this characterisation inaccurate?

A. Zero trust eliminates implicit trust but still grants explicit, evidence-based trust for a single session

B. Zero trust trusts all devices that hold a valid certificate for the lifetime of that certificate

C. Zero trust applies only to external users, so internal users retain their existing trust level

D. Zero trust replaces trust decisions with encryption, so no trust evaluation is needed

Correct answer: A. SP 800-207 removes trust that is inherited from network location or a previous decision. Trust is still granted — explicitly, per session, at least privilege, on current evidence, and subject to revocation. An architecture that granted no trust could not authorise any access. Certificate lifetime, internal-versus-external scoping and encryption are all irrelevant to the definition.

Q3. Which trust-algorithm characteristic allows a policy engine to respond to a moderately risky request by requiring a stronger authentication factor rather than denying it outright?

A. Singular evaluation

B. Criteria-based evaluation

C. Stateless evaluation

D. Score-based evaluation

Correct answer: D. A score-based algorithm produces a confidence value from weighted inputs, so the engine can act on a range of outcomes — allow, step up, restrict, deny. A criteria-based algorithm gates on a fixed set of qualifying attributes and is effectively binary. Singular evaluation refers to judging a request without subject history, and 'stateless evaluation' is not a term the framework uses.

Q4. An organisation must restrict east-west traffic between workloads in a data centre so a compromised application server cannot reach the database tier directly. Which SP 800-207 deployment approach does this requirement describe?

A. Enhanced identity governance

B. Micro-segmentation

C. Network infrastructure and software-defined perimeter

D. Device application sandboxing

Correct answer: B. Micro-segmentation places fine-grained enforcement — host agents or gateways — between workloads to police east-west traffic, which is exactly the stated requirement. Enhanced identity governance drives access from user and device identity attributes. A software-defined perimeter brokers user-to-resource connections. Device application sandboxing is a deployment variation that isolates applications on an endpoint, not a data-centre segmentation approach.

Q5. Which of the following is a supporting data source that feeds the policy engine, rather than a core logical component of the architecture?

A. The threat intelligence feed

B. The policy enforcement point

C. The policy administrator

D. The policy engine

Correct answer: A. SP 800-207 names three core logical components — policy engine, policy administrator and policy enforcement point. Threat intelligence sits among the supporting data sources alongside CDM, identity management, activity logs, SIEM data, data access policy, PKI and compliance requirements, all of which inform the trust algorithm without being part of the decision or enforcement chain themselves.

1.2 Measuring Maturity — CISA Zero Trust Maturity Model v2.0

NIST tells you what a zero-trust architecture is. It does not tell you how far along you are, which is the question every steering committee actually asks. The **Cybersecurity and Infrastructure Security Agency (CISA)** filled that gap with the *Zero Trust Maturity Model*, whose version 2.0 was published in April 2023 and is the version named in the exam blueprint. It is a measurement instrument: five pillars, three cross-cutting capabilities, four maturity stages, and a set of descriptions that let an organisation place itself honestly on a grid.

Five pillars and three cross-cutting capabilities

The five pillars divide the estate into areas that can be assessed separately, because organisations are rarely uniform — identity is often years ahead of data, and networks ahead of applications.

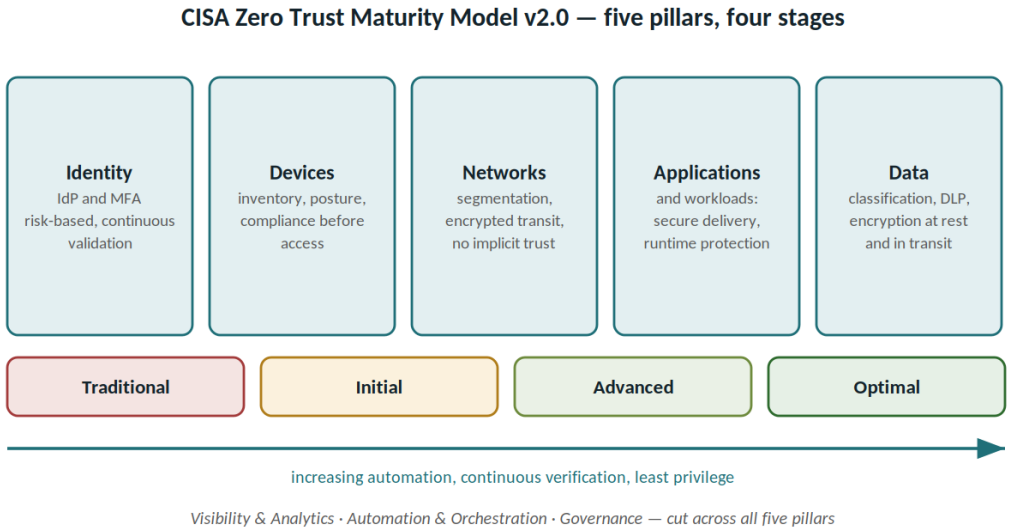

Figure 1.3 — The five pillars are assessed independently against four maturity stages; three capabilities cut across all of them.

- **Identity** — how the organisation authenticates and authorises people and machines: authentication strength, identity stores, risk assessment, access management.
- **Devices** — inventory completeness, device posture, threat protection on the endpoint, and whether device state gates access.
- **Networks** — segmentation, traffic encryption, network resilience, and whether network location still grants anything.
- **Applications and workloads** — secure application delivery and access, application security testing, threat protection at runtime, and how application state is monitored.
- **Data** — inventory and categorisation, availability, access control, and encryption at rest and in transit.

Three **cross-cutting capabilities** are assessed within every pillar rather than as a sixth pillar: **visibility and analytics**, **automation and orchestration**, and **governance**. This is the part organisations skip, and it is the part the model treats as the difference between a set of good products and an architecture. A pillar cannot reach optimal maturity if its decisions are not driven by telemetry, executed by automation and constrained by policy.

The four stages

Each pillar is scored against four stages. The definitions are deliberately behavioural — they describe what the organisation does, not what it owns.

Stage	What it looks like	Identity example
Traditional	Manually configured, static policy, siloed pillars, manual response, attributes assigned at provisioning	Passwords only, on-premises directory, access reviewed annually by spreadsheet
Initial	Some automation of configuration and policy, initial cross-pillar integration, manual response to most events	MFA added for remote access; identities consolidated into one IdP; some SSO
Advanced	Automated controls and lifecycle management, policy coordinated across pillars, centralised visibility, response increasingly automated	Phishing-resistant MFA for privileged users, risk signals feed access decisions, automated joiner / mover / leaver
Optimal	Fully automated just-in-time and just-enough access, dynamic policy across pillars, continuous validation, enterprise-wide visibility	Continuous validation with real-time risk scoring, phishing-resistant MFA everywhere, access granted just in time

Two words carry most of the weight in those definitions. **Automated** separates initial from advanced: if a human has to act for the control to work, the pillar is not advanced. **Continuous** separates advanced from optimal: if the decision is made once per session rather than re-evaluated during it, the pillar is not optimal.

> ✅ **Exam Tip — Read the stem for 'manual', 'automated' and 'continuous'**
>
> Stage-identification questions almost always turn on one adjective. 'MFA is required at login' is initial or advanced depending on the factor. 'Sessions are re-evaluated when device posture changes' is optimal, because it is continuous. 'The security team reviews the report weekly and disables stale accounts' is traditional to initial, because the control is a person.

The same capability across four stages

The clearest way to internalise the model is to take one capability and walk it up. Device posture is a good example because it appears in every Cisco deployment in this book.

Device posture, expressed as the four maturity stages

```
Traditional  Devices are inventoried in a spreadsheet. Posture is not checked.
             Any device with a valid password reaches any internal application.

Initial      An MDM enrols corporate laptops. A monthly report lists devices
             missing patches. Access is not affected by the report.

Advanced     Duo Device Health and the Secure Client posture module check disk
             encryption, EDR and OS version at authentication. Non-compliant
             devices are blocked automatically and sent to a remediation page.

Optimal      Posture is re-evaluated during the session. A device whose EDR
             agent stops mid-session has its ZTNA session revoked within
             minutes; access is restored automatically when compliance returns.
```

Notice that the product does not change between advanced and optimal — the same Duo and Secure Client deployment supports both. What changes is whether the check happens at the door or for the whole visit, and whether the response is automatic in both directions. Maturity is mostly a configuration and process property, which is why the model is useful as a roadmap rather than a shopping list.

> 💡 **Did you know? — Version 2.0 changed the scoring**
>
> The original 2021 model had three stages: traditional, advanced and optimal. Version 2.0 inserted **initial** between traditional and advanced, precisely because agencies reported that the jump was too large to plan against and organisations were stuck claiming 'traditional' for years. If an exam item lists three stages, it is describing the superseded version — the current model, and the blueprint, name four.

Using the model against a Cisco estate

The practical use of the model is to produce a current-state and target-state assessment per pillar, then identify which capability moves the needle furthest. The table below maps each pillar to the technology in this book that most directly advances it.

Pillar	What optimal requires	Where it is built in this book
Identity	Phishing-resistant MFA, continuous validation, just-in-time access	Duo, IdP connections, SCIM — Section 2

Pillar	What optimal requires	Where it is built in this book
Devices	Complete inventory, posture enforced at and during access	Duo Device Health, Secure Client posture — 2.7
Networks	Micro-segmentation, all traffic encrypted, no location-based trust	Secure Access, Multicloud Defense — Sections 3 and 4
Applications and workloads	Per-application authorisation, runtime protection, continuous monitoring	ZTNA, Secure Workload, AI Defense — Sections 3 and 4
Data	Categorised inventory, dynamic access control, encryption everywhere	DLP, CASB, encryption design — 3.1, 3.3, 4.4, 4.5
Cross-cutting	Telemetry-driven decisions, automated response, governed policy	Dashboards, XDR and Splunk integration — Section 5

⚠️ **Common Pitfall — Do not average the pillars into one number**

Executives want a single maturity score, and producing one destroys the model's value. An organisation at optimal on identity and traditional on data is not 'advanced overall' — it is one well-defended front door in front of an unclassified data estate, and the average hides exactly the gap an attacker will use. Report per pillar, with the cross-cutting capabilities scored inside each.

🌍 **From the Field — The assessment that changed a budget**

A retailer scored themselves advanced on networks because every site had segmentation and every tunnel was encrypted. The assessment of the data pillar came back traditional: no inventory, no categorisation, and card data discovered in three file shares nobody owned. The segmentation project had been protecting a perimeter around data whose location was unknown. The next year's budget moved from network refresh to data discovery and DLP — a decision the model made obvious and a product comparison never would have.

Summary

The CISA Zero Trust Maturity Model v2.0 measures progress rather than defining architecture. It assesses five pillars — identity, devices, networks, applications and workloads, and data — independently, and scores three cross-cutting capabilities, visibility and analytics, automation and orchestration, and governance, within each. Four stages describe behaviour: traditional is manual and siloed, initial adds some automation and cross-pillar integration, advanced brings automated controls and coordinated policy with centralised visibility, and optimal delivers fully automated, just-in-time, continuously validated access driven by dynamic policy. The words that distinguish the stages are 'manual', 'automated' and 'continuous', and pillars should be reported separately rather than averaged into a single score.

Key Takeaways

- Five pillars: **Identity, Devices, Networks, Applications and Workloads, Data**.
- Three cross-cutting capabilities assessed inside every pillar: **visibility and analytics, automation and orchestration, governance**.
- Four stages: **Traditional → Initial → Advanced → Optimal**; v2.0 added 'initial' to the original three.
- **Automated** marks the step to advanced; **continuous** marks the step to optimal.
- The model measures maturity, not compliance — it prescribes no products and no completion date, and pillars are scored independently.

Knowledge Check — 1.2 CISA Zero Trust Maturity Model v2.0

Q1. An organisation enforces phishing-resistant MFA at login and automatically revokes a user's application session when the endpoint's EDR agent stops running mid-session. Which CISA maturity stage does this describe?

A. Advanced

B. Optimal

C. Initial

D. Traditional

Correct answer: B. Re-evaluating and revoking access during a session is continuous validation, which is the defining property of optimal maturity. Advanced would enforce the same checks automatically but at access time only. Initial implies partial automation with manual response, and traditional implies manual, static policy with no posture influence on access.

Q2. Which three capabilities does CISA describe as cutting across all five pillars rather than forming pillars of their own?

A. Identity, devices and data

B. Encryption, segmentation and inventory

C. Detection, response and recovery

D. Visibility and analytics, automation and orchestration, governance

Correct answer: D. Version 2.0 defines visibility and analytics, automation and orchestration, and governance as cross-cutting capabilities assessed within each pillar. Identity, devices and data are three of the five pillars themselves. Encryption, segmentation and inventory are individual capabilities inside pillars, and detection, response and recovery are functions from the NIST Cybersecurity Framework, a different document.

Q3. What distinguishes the initial stage from the advanced stage in ZTMM v2.0?

A. At advanced, controls and lifecycle management are automated and policy is coordinated across pillars; at initial, automation is partial and most response is manual

B. At advanced, the organisation has purchased a zero-trust platform; at initial it has not

C. At advanced, all traffic is encrypted; at initial, no traffic is encrypted

D. At advanced, MFA is deployed; at initial, only passwords are used

Correct answer: A. The stages describe behaviour. Advanced requires automated controls, automated lifecycle management, policy coordinated across pillars and centralised visibility, while initial has some automation of configuration and policy but leaves most response manual. Product purchase is never a stage criterion, and encryption and MFA are individual capabilities that appear in several stages at differing strengths.

Q4. A CISO asks for a single overall zero-trust maturity score for board reporting. What is the strongest objection to producing one?

A. The model forbids reporting maturity to executives

B. Maturity can only be scored by an external assessor

C. Averaging hides a pillar at traditional maturity behind pillars that are more advanced, concealing the gap an attacker would use

D. The four stages cannot be represented numerically

Correct answer: C. The model assesses pillars independently precisely because organisations progress unevenly. An identity pillar at optimal averaged with a data pillar at traditional produces a comfortable middling number that conceals the weakest link. Nothing in the model restricts executive reporting or mandates an external assessor, and the stages can be represented numerically — the objection is to collapsing them into one figure.

Q5. Which statement about the relationship between NIST SP 800-207 and the CISA Zero Trust Maturity Model is correct?

A. The CISA model supersedes SP 800-207 and replaces its component definitions

B. SP 800-207 applies to federal agencies only; the CISA model applies to private industry

C. Both documents mandate the same specific set of security products

D. SP 800-207 defines the architecture and its components; the CISA model measures how far an organisation has progressed toward it

Correct answer: D. The two are complementary: NIST defines the architecture, its tenets and its logical components, while CISA provides a measurement instrument with pillars and stages for tracking progress. Neither supersedes the other, neither mandates products, and although CISA's model was written for federal agencies it is widely adopted in private industry — as is SP 800-207.

1.3 MITRE ATT&CK for the Cloud — Techniques and Mitigations

Frameworks describe the defence. **MITRE ATT&CK** describes the offence — a curated, public knowledge base of the tactics and techniques adversaries have actually been observed using, built from incident reporting rather than theory. The blueprint asks you to *select mitigation methods* against attacker tactics and techniques using the cloud matrices, which means this sub-section is not about memorising technique numbers. It is about reading a described behaviour, recognising which technique it is, and choosing the control that stops it.

How ATT&CK is structured

ATT&CK is organised in three layers. **Tactics** are the adversary's objectives — the *why* — and run roughly in the order of an intrusion: reconnaissance, initial access, execution, persistence, privilege escalation, defence evasion, credential access, discovery, lateral movement, collection, exfiltration, impact. **Techniques** are the *how*, each with an identifier such as T1078. **Sub-techniques** refine them: T1078.004 is *Valid Accounts: Cloud Accounts*. Separately, **mitigations** carry M-numbers, such as M1032 *Multi-factor Authentication*, and each technique page lists the mitigations that apply to it.

The Enterprise matrix contains platform-specific views. The ones that matter here are **IaaS**, **SaaS**, **Identity Provider** and **Office Suite** — the last two added as cloud attacks moved decisively toward identity. A technique that appears in the Identity Provider matrix is one that works against Entra ID, Okta or Ping regardless of which cloud the workloads run in.

> 💡 **Did you know? — The matrices were re-cut around identity**
>
> Earlier versions of ATT&CK had a single 'Azure AD' platform inside the cloud matrix. As incidents increasingly began at the identity provider and only later touched infrastructure, MITRE separated Identity Provider into its own platform. That re-cut mirrors what the SSCA blueprint assumes throughout: in cloud environments, identity is the primary attack surface and the network is downstream of it.

The techniques that define cloud intrusions

A small number of techniques account for most real cloud compromises. Learn these by behaviour rather than by number — the exam describes the behaviour.

Technique	What the adversary does	Primary mitigation
Valid Accounts: Cloud Accounts (T1078.004)	Signs in with legitimate credentials obtained by phishing, purchase or reuse	Phishing-resistant MFA (M1032), conditional access, account use policies

Technique	What the adversary does	Primary mitigation
Steal Application Access Token (T1528)	Tricks a user into consenting to a malicious OAuth application that then reads mail or files without a password	Restrict application consent to administrators, application vetting, audit granted scopes
Multi-Factor Authentication Request Generation (T1621)	Floods the user with push prompts until one is approved — 'MFA fatigue'	Verified push with number matching, FIDO2, rate limiting, user training (M1017)
Modify Authentication Process: MFA (T1556.006)	Registers an additional factor or disables MFA for an account after gaining access	Privileged account management (M1026), alerting on factor enrolment changes
Additional Cloud Credentials (T1098.001)	Adds an access key, service-principal secret or SSH key to maintain access after the password is rotated	Audit credential creation (M1047), short-lived credentials, workload identity federation
Cloud Infrastructure Discovery (T1580)	Enumerates instances, buckets, roles and network configuration to find the next target	Least-privilege read permissions, control-plane logging and detection
Data from Cloud Storage (T1530)	Reads objects from misconfigured or over-permissioned buckets and blob containers	Block public access, encryption with managed keys, access control audits, DLP
Resource Hijacking (T1496)	Spins up compute for crypto-mining using stolen credentials	Budget alerts, egress control, privileged account management

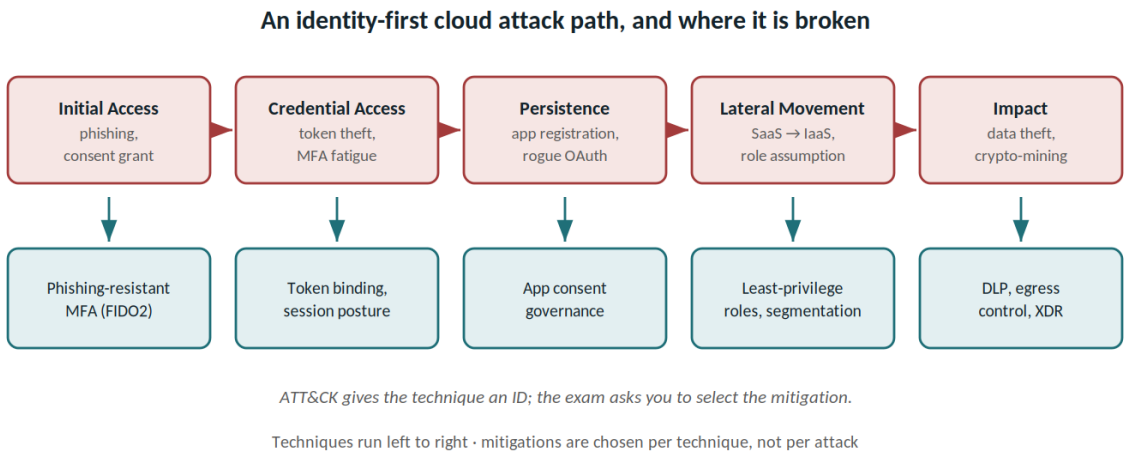

Figure 1.4 — A typical identity-first cloud attack path, with the mitigation selected for each step rather than for the attack as a whole.

Selecting a mitigation, the way the exam asks

The blueprint verb is *select*, and the skill is narrower than it looks. A mitigation must address the specific technique described, not the general area. If the stem describes push-notification flooding, the answer is number matching or a phishing-resistant factor — not 'deploy MFA', because MFA is already deployed; that is the premise of the attack. If the stem describes a malicious OAuth consent grant, the answer is consent governance — not MFA, because the user authenticated legitimately and the token was granted afterwards.

> ✅ **Exam Tip — Match the control to the step, not to the story**
>
> Cloud attack narratives in exam questions run several steps. The question then asks what would have prevented one of them. Locate the sentence containing the technique being asked about, and choose the control that intercepts that sentence. A control that would have stopped an earlier step is a distractor, because by the premise of the question it did not.

Detection is the other half of the answer. Most cloud techniques leave a distinctive control-plane event, and the mitigation for a persistence technique is often 'alert on the event' rather than 'block the action', because the action is legitimate when an administrator does it. The query below is the kind of detection an organisation writes after reading the T1098.001 technique page — you will build the Splunk side of this in Section 5.

Detecting additional cloud credentials being attached to a principal (Splunk, AWS CloudTrail data)

```
index=cloudtrail sourcetype=aws:cloudtrail
  (eventName=CreateAccessKey OR eventName=CreateLoginProfile
   OR eventName=UpdateAccessKey OR eventName=CreateServiceSpecificCredential)
| eval target=coalesce(requestParameters.userName, userIdentity.userName)
| stats count min(_time) as first max(_time) as last
        values(sourceIPAddress) as src values(eventName) as actions
        by userIdentity.arn, target
| where userIdentity.arn != target                /* key created for another
principal */
| eval age_days=round((now()-first)/86400,1)
| sort - count
```

Reviewing which OAuth applications hold consent in a tenant (Microsoft Graph PowerShell)

```
# Every delegated permission grant in the tenant, newest first
Get-MgOauth2PermissionGrant -All |
  Select-Object ClientId, ConsentType, PrincipalId, Scope |
  Sort-Object ClientId | Format-Table -AutoSize

# Applications holding high-value mail or file scopes
Get-MgOauth2PermissionGrant -All |
  Where-Object { $_.Scope -match "Mail.Read|Files.ReadWrite.All|Directory.Read.All"
}
```

⚠️ **Common Pitfall — Mitigating the technique you recognise instead of the one described**

The most common wrong answer in this area is a correct control applied to the wrong step. 'Enable MFA' is right for credential stuffing and useless against a stolen session token, because the token was minted after MFA succeeded. Token theft is mitigated by shorter token lifetimes, token binding, continuous access evaluation and device-bound sessions — controls that make a replayed token invalid somewhere other than the login page.

🌍 **From the Field — Three prompts at 02:10**

A finance user approved a push notification at ten past two in the morning after the fourth prompt, because it was the only way to make her phone stop. The attacker had the password from an unrelated breach; MFA was deployed and working exactly as configured. The remediation was not more MFA — it was number matching, which replaces 'approve or deny' with 'type the digits shown on the screen you are looking at', and a policy that locks the account after repeated denials. Both are configuration changes in Duo that take an afternoon.

Summary

MITRE ATT&CK catalogues observed adversary behaviour as tactics, techniques and sub-techniques, each with mitigations identified by M-number. Its IaaS, SaaS, Identity Provider and Office Suite platforms describe cloud intrusions, which today almost always begin at identity: valid cloud accounts, stolen application tokens, MFA request generation and modification of the authentication process, followed by additional credentials for persistence, infrastructure discovery, access to cloud storage and resource hijacking for impact. Selecting a mitigation means matching a control to the specific step described — number matching and FIDO2 against push fatigue, consent governance against malicious OAuth grants, short-lived credentials and control-plane alerting against added keys — rather than applying a general control to the story as a whole.

Key Takeaways

- ATT&CK structure: **tactics** (objectives) → **techniques** (T-numbers) → **sub-techniques**, with **mitigations** as M-numbers.
- Cloud-relevant platforms: **IaaS, SaaS, Identity Provider, Office Suite** — identity was separated out because it is where cloud intrusions start.
- **T1078.004** valid cloud accounts, **T1528** stolen application access token, **T1621** MFA request generation, **T1556.006** MFA modification, **T1098.001** additional cloud credentials, **T1530** data from cloud storage, **T1496** resource hijacking.
- **MFA does not stop token theft** — the token is issued after authentication; use short lifetimes, token binding and continuous evaluation.
- For persistence techniques the practical mitigation is often **detection and alerting on the control-plane event**, because the action itself is legitimate for administrators.

Knowledge Check — 1.3 MITRE ATT&CK for the Cloud

Q1. A user receives repeated authentication push notifications late at night and eventually approves one, giving the attacker access. Which mitigation most directly addresses the technique used?

A. Enabling multi-factor authentication for the affected user

B. Enforcing a longer minimum password length

C. Blocking logins from countries where the organisation has no offices

D. Number matching or a FIDO2 factor, so approval requires interaction with the screen initiating the login

Correct answer: D. The technique is MFA request generation (T1621), and its premise is that MFA is already enabled — so enabling MFA changes nothing. Number matching forces the user to read a value from the authenticating screen, and a phishing-resistant FIDO2 factor removes the approve-or-deny prompt entirely. Password length does not affect an attacker who already has the password, and geo-blocking is trivially bypassed with a proxy.

Q2. An adversary convinces a user to grant consent to a third-party application, which then reads the user's mailbox using the issued OAuth token. Which control best mitigates this technique?

A. Requiring MFA at every sign-in

B. Restricting application consent to administrators and vetting requested scopes

C. Rotating the user's password every 60 days

D. Deploying a secure web gateway for all outbound traffic

Correct answer: B. This is Steal Application Access Token (T1528). The user authenticated legitimately and then consented, so MFA and password rotation are irrelevant — the token is independent of the password. Restricting consent to administrators, combined with scope review and periodic audit of granted permissions, prevents the grant from being made. A web gateway does not see API calls made from the attacker's own infrastructure to the SaaS provider.

Q3. After compromising an account, an attacker creates an additional access key for a different IAM principal so access survives a password reset. Which technique is this, and what is the most practical mitigation?

A. Cloud Infrastructure Discovery (T1580); disable the cloud provider's API

B. Resource Hijacking (T1496); apply a compute budget alert

C. Additional Cloud Credentials (T1098.001); alert on credential-creation events in the control-plane log and enforce least privilege for IAM write actions

D. Data from Cloud Storage (T1530); block public access on all buckets

Correct answer: C. Creating additional credentials for persistence is T1098.001. Creating an access key is a legitimate administrative action, so the practical defence is detection — alerting on CreateAccessKey and similar events, especially when the creating principal differs from the target — combined with least privilege on IAM write permissions. The other options name different techniques with unrelated mitigations.

Q4. In ATT&CK terminology, what is the difference between a tactic and a technique?

A. A tactic is the adversary's objective; a technique is the specific method used to achieve it

B. A tactic is used by nation states; a technique is used by commodity malware

C. A tactic applies to on-premises systems; a technique applies to cloud systems

D. A tactic is a mitigation; a technique is an attack

Correct answer: A. Tactics answer why — initial access, persistence, credential access, impact — and techniques answer how, each carrying a T-number and often refined by sub-techniques. Attribution, platform and mitigation are all separate dimensions of the knowledge base and none of them define the tactic-versus-technique distinction.

Q5. A session token is stolen from a compromised endpoint and replayed by the attacker from a different device. Why does existing multi-factor authentication fail to stop this?

A. Multi-factor authentication does not apply to cloud applications

B. The token was issued after authentication succeeded, so replaying it does not return the attacker to the login flow

C. The token is encrypted, so the identity provider cannot validate it

D. Multi-factor authentication is only evaluated for administrative accounts

Correct answer: B. A bearer token represents an authentication that has already happened; presenting it bypasses the login flow entirely, so no factor is requested. Mitigations work by invalidating the replay — short token lifetimes, binding tokens to a device or client certificate, continuous access evaluation that reacts to a changed IP or posture, and revoking sessions on risk signals. The remaining options misstate how MFA and tokens work.

1.4 Public Cloud Security Requirements — Identity, Responsibility and Posture

The blueprint asks you to describe security requirements for AWS, Azure and Google Cloud. That phrasing hides the single most important idea in cloud security: the requirements are not a longer version of the data-centre requirements. They are a different set, because the provider has taken over some of the work, the control plane has become internet-facing, and the primary access mechanism is an API call authorised by a token rather than a packet arriving on a port.

The Shared Responsibility Model

Every provider publishes a version of the same picture, usually summarised as *security of the cloud* versus *security in the cloud*. The provider secures the physical facilities, the hardware, the hypervisor and the managed service itself. The customer secures what they put on top. The boundary is not fixed — it moves with the service tier, and the exam tests exactly that movement.

Shared Responsibility Model — the line moves with the service tier

	On-prem	IaaS	PaaS	SaaS
Data and classification	customer	customer	customer	customer
Identity and access	customer	customer	customer	customer
Application	customer	customer	customer	provider
Operating system	customer	customer	provider	provider
Network controls	customer	customer	shared	provider
Hypervisor / host	customer	provider	provider	provider
Physical infrastructure	customer	provider	provider	provider

You can outsource the operation of a layer. You can never outsource accountability for your data.

Figure 1.5 — Responsibility for each layer by service tier. The line moves; accountability for the data never does.

Read the figure column by column. With **IaaS** you rent a virtual machine, so you patch the operating system, configure the firewall and manage the network controls. With **PaaS** — a managed database, a serverless function, an app service — the provider runs the operating system and the runtime, leaving you identity, data, application code and configuration. With **SaaS** you hold only identity, data, and whatever policy knobs the vendor exposes. In all three, **data classification and identity remain yours**, which is why every Cisco control in this book attaches to one or the other.

> ✅ **Exam Tip — Two questions resolve most responsibility items**

First, which tier is described — IaaS, PaaS or SaaS? Second, which layer does the requirement name? Patching an operating system is the customer's job in IaaS and the provider's in PaaS and SaaS. Configuring identity, access control and data protection is always the customer's. If an option makes the provider responsible for customer data or access policy, it is wrong in every tier.

⚠️ **Common Pitfall — Responsibility can be shared; accountability cannot be delegated**

A provider running your database does not make a breach of that database their regulatory problem. Under GDPR, PCI DSS and most contracts, the organisation that determined the purpose of the processing answers for it. 'The provider is responsible for that layer' is a statement about who operates a control, never about who is answerable for the outcome.

Identity and access control across the three clouds

Cloud identity is where zero trust becomes concrete, because in a public cloud every action — creating a VM, reading an object, changing a firewall rule — is an API call that must be authorised. The three providers use different words for near-identical concepts, and the exam uses all three vocabularies.

Concept	AWS	Azure	Google Cloud
Directory of human identities	IAM users, or federated identities via IAM Identity Center	Microsoft Entra ID	Cloud Identity / Workspace
Permission container	IAM policy (JSON) attached to user, group or role	RBAC role definition assigned at a scope	IAM role bound to a principal on a resource
Assumable identity	IAM role assumed via STS, returning temporary credentials	Managed identity (system- or user-assigned)	Service account, impersonated or attached
Organisation-wide guardrail	Service control policy (SCP) on an organisational unit	Azure Policy and management-group scope	Organisation policy constraints
Machine credential without secrets	IAM Roles Anywhere, OIDC federation	Workload identity federation	Workload identity federation

Three requirements apply in all three clouds. **Eliminate long-lived static credentials** — an access key in a repository is the single most common root cause of cloud compromise, and every provider now offers a federation mechanism that issues short-lived credentials instead. **Enforce least privilege with explicit conditions** — a permission is not least privilege because it names one service; it is least privilege when it

names the action, the resource and the circumstances. **Protect the root or global administrator identity** with phishing-resistant MFA, no access keys, and use only for tasks that genuinely require it.

A least-privilege IAM policy: one action, one bucket prefix, only from a managed session with MFA

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "ReadFinanceReportsOnly",
      "Effect": "Allow",
      "Action": ["s3:GetObject", "s3:ListBucket"],
      "Resource": [
        "arn:aws:s3:::acme-finance-reports",
        "arn:aws:s3:::acme-finance-reports/quarterly/*"
      ],
      "Condition": {
        "Bool":        { "aws:MultiFactorAuthPresent": "true" },
        "NumericLessThan": { "aws:MultiFactorAuthAge": "3600" },
        "StringEquals":{ "aws:PrincipalTag/department": "finance" },
        "IpAddress":   { "aws:SourceIp": ["198.51.100.0/24"] }
      }
    },
    {
      "Sid": "DenyUnencryptedTransport",
      "Effect": "Deny",
      "Action": "s3:*",
      "Resource": "arn:aws:s3:::acme-finance-reports/*",
      "Condition": { "Bool": { "aws:SecureTransport": "false" } }
    }
  ]
}
```

The second statement is worth noticing. In AWS policy evaluation an explicit **Deny** always wins over any Allow, which makes deny statements the reliable way to express a guardrail that no other policy can override. The same pattern in Azure is a deny assignment or an Azure Policy with a deny effect; in Google Cloud it is an organisation policy constraint or a deny policy.

Equivalent checks from the CLI — the three commands worth knowing per cloud

```
# AWS  — who has console access without MFA, and how old are the keys?
aws iam generate-credential-report
aws iam get-credential-report --query Content --output text | base64 -d | \
    awk -F, 'NR==1 || ($4=="true" && $8=="false")' | cut -d, -f1,4,8,9,10

# Azure — which principals hold Owner or Contributor at subscription scope?
az role assignment list --include-inherited --all \
   --query "[?roleDefinitionName=='Owner'||roleDefinitionName=='Contributor'].
```

```
            {p:principalName, role:roleDefinitionName, scope:scope}" -o table

# GCP   — which members hold primitive roles on the project?
gcloud projects get-iam-policy acme-prod --format=json | \
   jq -r '.bindings[] | select(.role|test("roles/(owner|editor)")) |
          .role as $r | .members[] | "\($r)\t\(.)"'
```

💡 **Did you know? — The control plane is a public endpoint**

There is no cable to unplug. The AWS, Azure and Google APIs are reachable from anywhere on the internet, and a valid credential from a coffee shop is indistinguishable from the same credential in your data centre unless you make it distinguishable — with conditions on source IP or private endpoint, with device-bound sessions, and with MFA age checks like the one in the policy above. This is why identity, not network placement, is the primary cloud control.

Posture and compliance

Posture management answers a different question from access control: not *who may act*, but *is the resulting configuration safe right now*. **Cloud Security Posture Management (CSPM)** tools continuously read the control plane, compare the configuration against a benchmark, and report drift. The benchmarks that matter in practice are the CIS Benchmarks per provider, the provider's own well-architected security pillar, and whatever regulatory framework applies — PCI DSS, HIPAA, ISO 27001, the NIS2 baseline in the EU.

A small set of misconfiguration classes accounts for most findings, and they recur in exam scenarios: storage exposed publicly, security groups allowing 0.0.0.0/0 to administrative ports, unencrypted volumes or buckets, disabled or unaggregated logging, over-permissive IAM roles, unrotated keys, and default network configurations left in place. Posture drift matters as much as the initial state — the environment is safe at deployment and unsafe six weeks later because somebody opened a port to debug something at midnight.

🌍 **From the Field — The rule that came back**

A platform team fixed a finding for an RDP port open to the internet, closed the ticket, and saw the same finding three days later. The security group was managed by Terraform; somebody had fixed it in the console, and the next pipeline run restored the committed state. Posture management without a path back into the infrastructure-as-code repository produces an endless loop of true findings and honest remediation. The fix is to treat a CSPM finding as a code defect, not a console task.

⚠️ **Common Pitfall — Compliant is not the same as secure**

A benchmark is a floor written for a general audience. An account can pass every CIS control and still allow a developer role to read the production customer database, because no benchmark knows

which of your data matters. Use posture management to eliminate the known-bad configurations quickly, then spend the time it saves on the controls that are specific to your data — which is the work Sections 3 and 4 are about.

Summary

Public cloud security starts with the Shared Responsibility Model, whose boundary moves with the service tier: in IaaS the customer owns everything from the guest operating system upward, in PaaS the provider takes the OS and runtime, and in SaaS only identity, data and exposed policy remain the customer's — but data classification, identity and access control are the customer's in every tier, and accountability is never transferred. Because every cloud action is an authorised API call, identity is the primary control: eliminate long-lived static credentials in favour of federated short-lived ones, express least privilege with explicit actions, resources and conditions, use explicit deny statements as guardrails, and protect the root identity with phishing-resistant MFA. Posture management then continuously compares the live configuration against benchmarks such as the CIS Benchmarks, catching the recurring misconfiguration classes — public storage, open administrative ports, missing encryption, disabled logging, over-permissive roles — and the drift that reintroduces them.

Key Takeaways

- The responsibility line moves with the tier; **data and identity are always the customer's**, in IaaS, PaaS and SaaS alike.
- **Responsibility is shared; accountability is not delegable** — regulators hold the data controller answerable regardless of who operates the layer.
- Replace static access keys with **federated, short-lived credentials**: IAM roles and STS, Entra managed identities, GCP service-account impersonation, workload identity federation.
- In AWS policy evaluation an **explicit Deny always wins**; use deny statements, SCPs, Azure Policy deny effects or GCP organisation constraints as guardrails.
- **CSPM measures configuration, not permission** — and passing a benchmark is a floor, not evidence that your data is protected.

Knowledge Check — 1.4 Public Cloud Security Requirements

Q1. An organisation runs a managed PaaS database service. Under the Shared Responsibility Model, which task remains the customer's responsibility?

A. Classifying the data stored in the database and controlling who may access it

B. Patching the underlying operating system that hosts the database engine

C. Maintaining the physical security of the data centre

D. Replacing failed storage hardware in the provider's region

Correct answer: A. In PaaS the provider operates the host operating system, the runtime and the physical infrastructure. Data classification and access control stay with the customer in every service

tier, which is precisely why identity and data controls dominate cloud security design. The other three tasks sit on the provider's side of the line for a managed service.

Q2. Which approach best eliminates the risk of long-lived cloud credentials being leaked from a CI/CD pipeline?

A. Store the access key in an encrypted environment variable in the pipeline configuration

B. Rotate the access key manually every ninety days

C. Configure workload identity federation so the pipeline exchanges its OIDC token for short-lived cloud credentials

D. Restrict the access key to a single region

Correct answer: C. Federation removes the static secret entirely: the pipeline presents a signed OIDC token proving its identity and receives temporary credentials scoped to a role. Encrypting or rotating a key still leaves a long-lived secret that can be exfiltrated and used until rotation, and a regional restriction limits the blast radius without addressing the leak.

Q3. In AWS, an IAM policy attached to a role allows s3:GetObject on a bucket, while a service control policy on the account's organisational unit denies all S3 actions. What is the effective result?

A. Access is allowed, because the identity-based policy is more specific

B. Access is denied, because an explicit deny overrides any allow

C. Access is allowed, because service control policies apply only to root users

D. The evaluation fails and the request returns an error to the administrator

Correct answer: B. AWS policy evaluation is deny-by-default with explicit-deny precedence: any applicable deny, whether from an SCP, a resource policy or an identity policy, overrides every allow. Specificity does not break the tie, SCPs apply to all principals in the account other than the management account's root in specific cases, and the request is simply denied rather than producing an evaluation error.

Q4. A CSPM tool reports that a security group permits inbound TCP 3389 from 0.0.0.0/0. The team fixes it in the console, but the finding reappears after the next infrastructure pipeline run. What is the correct remediation?

A. Suppress the finding in the CSPM tool, as it is a false positive

B. Disable the pipeline until the security review is complete

C. Move the instance to a different subnet

D. Correct the rule in the infrastructure-as-code repository so the deployed state matches the intended state

Correct answer: D. The declared state in code is authoritative, so any console fix is overwritten at the next run. Treating the finding as a code defect fixes it permanently and keeps the repository truthful. Suppressing a true finding hides a real exposure, disabling the pipeline blocks unrelated work, and relocating the instance does not change the rule that exposes it.

Q5. Which statement best describes the difference between cloud security posture management and identity access management?

A. Posture management issues credentials; access management scans for vulnerabilities in workloads

B. Posture management applies only to SaaS; access management applies only to IaaS

C. Posture management evaluates whether the resulting configuration is safe; access management governs who may perform an action

D. They are two names for the same control and differ only by vendor

Correct answer: C. The two answer different questions. Access management authorises actions on the control plane, while posture management continuously reads the configuration those actions produced and compares it against a benchmark or policy. Posture management issues no credentials, is not tier-specific, and workload vulnerability scanning is a third, separate capability covered in Section 3.

A Preview — Eight Cases from the Field

The book closes with eight incident narratives, because real failures do not respect domain boundaries: they begin in one section of the blueprint, are missed because of a decision made in another, and are resolved by somebody who understood a third. Each case names the sub-sections it draws on, so a case that is hard to follow points straight at a topic worth rereading.

Case	What it turns on
The SaaS tenant nobody owned	Four hundred unfederated accounts in a tenant security had never heard of
The MFA prompt that was not the user's	Push fatigue, and a factor that asks a human to decide with no information
The branch that failed open	A temporary steering exception that protected nobody for nine months
The DLP rule that stopped payroll	A correct classifier with no destination scoping and no threshold
The private app that worked only for the pilot group	A resource definition missing the port the pilot reached another way
The AI assistant that leaked source code	Four months of prompts nobody could see, and credentials inside them
The segmentation rollout that broke backups	A monthly flow that never appeared in a four-week discovery window
The alert that XDR closed too early	Containment scoped to the device when the incident belonged to the identity

A short chapter then draws out what all eight have in common — among other things, that in every case the technology did exactly what it had been configured to do, and the failure lived in the gap between the configuration and somebody's assumption about it.

Continue Reading

This sample contains the book's front matter and the first four sub-sections of Section 1. The complete guide runs to **324 pages** and covers every topic in the published v2.0 blueprint.

The full book contains	
Sections	5, one per blueprint domain
Sub-sections	52, each with the same structure as those in this sample
Knowledge-check questions	260, with explanations of the wrong answers as well as the right one
Diagrams	48
Cases from the field	8 cross-domain incident narratives, plus what they have in common
Reference	A full glossary and a blueprint-to-sub-section map

The structure you have just read repeats for all 52 sub-sections: explanation from first principles, a diagram where the relationship is easier seen than described, real configuration with verification steps, four kinds of callout, a summary and key takeaways written to work as a standalone revision pass, and five knowledge-check questions whose explanations cover the plausible wrong answers as well as the right one.

Section 2 covers identity: certificates, Duo, phishing-resistant multifactor, posture, SAML and SCIM. Section 3 covers policy, from encryption and data loss protection through AI Defense and AI Guardrails to Secure Workload microsegmentation. Section 4 configures Cisco Secure Access end to end — DNS security, the secure web gateway, CASB, connectors, tunnels, VPN as a service and zero-trust access with QUIC and MASQUE. Section 5 covers operations, dashboards and the XDR and Splunk integrations.

Get the complete guide

leanpub.com/ssca

Thank you for reading.