Unified Claude Security Best Practices

Version 1.3.6

2026-11-07

Unified Claude: Architecture and Security Changes Overview

If you have already approved Claude chat, the unified experience keeps the same data and network governance controls, identity checks, connector credential model and US data residency. What changes is that a chat can move into a cloud session in a VM where the agent can execute code in a sandboxed environment - similar to the architecture for Cowork in the Cloud. This document lays out the architectural details of unified Claude. It also provides details of where it differs from Chat and Cowork.

Introduction: Difference in security architecture

Topic	Claude chat	Cowork on the desktop	Cowork in the cloud	Chat and Cowork unified (beta)
Where the agent runs and when work starts	Anthropic's cloud. Agentic work runs only when Claude uses Code execution.	On the user's computer, inside the Claude Desktop app. By default it runs as a host process. A device configuration key can require it to run inside the local VM. Work starts when the user starts a Cowork session.	Anthropic's cloud. Work starts when the user starts a cloud session, or when one of their scheduled tasks runs.	Anthropic's cloud, as with Claude chat. When a task needs a sandbox, Anthropic's service starts using the conversation's cloud session automatically. The service makes this decision. The user does not need to do anything, and the model cannot start it.
Where code	In a Firecracker	Commands	Each session	Each session runs

runs and how it is isolated	microVM. Each conversation gets its own.	and code run in a dedicated Linux VM on the user's computer. The operating system's hypervisor isolates it. File tools work only on folders the user connects. Plugins, local MCP servers and computer use run on the computer, outside the VM.	runs in its own Firecracker microVM, the same kind Claude chat uses, on the same infrastructure. The host blocks traffic from a VM to the host and to other VMs.	in its own Firecracker microVM, the same kind Claude chat uses, on the same infrastructure. Inside the VM, Claude can use a shell, files and connectors. The same Code execution setting and connector policies apply.
What leaves the user's device	Everything the user sends. Results from tools the app declares are also sent, such as local MCP servers and phone tools for calendar, reminders and location.	Every model request, including the contents of files Claude reads.	Everything the user sends, since it is a cloud service. If a computer is linked, the output of tools run on it is also sent. This includes file contents, command output and screenshots. A granted folder's instruction and rules files are sent too.	Everything the user sends. Results from tools the app declares are also sent, such as local MCP servers and phone tools for calendar, reminders, location and health. If a computer is linked, the output of tools run on it is also sent. This includes file contents, command output and screenshots. A granted folder's instruction and rules files are sent too.
Where data is stored	One conversation record in claude.ai, kept in Anthropic storage and encrypted at	The session transcript is stored on the user's computer.	The session transcript and files are kept in Anthropic storage. The	The conversation record is kept as in Claude chat. When the cloud session is used, a

	rest.	Enterprise organisations with the Compliance API also get a server-side record.	VM's working disk is saved as encrypted checkpoints.	second copy is made in the session and on its VM disk. This copy includes recent history (with tool results, images and any memory summary), a summary of earlier history, the member's saved preferences, attached files, referenced project files and the member's skills. Local projects and scheduled task metadata *are* uploaded to the cloud
Retention and deletion	Keep within the custom data retention defined by the admin. If not defined, retained indefinitely. Activity feed events are stored for 6 years.	Data on the user's computer is not covered by your retention setting. The server-side activity feed record is kept for 6 years.	Deleting a session makes it inaccessible right away. Its transcript, files, VM and checkpoints are then deleted in the background. While a session is idle, only its disk is kept, as an encrypted checkpoint. The VM environment expires after 30 days. Custom retention does not yet apply to Cowork cloud sessions. Compliance API	Unified Claude follows the same rules as Claude chat and is governed by custom data retention. Activity feed events are stored for 6 years. The Cloud VM environment expires after 30 days. Deleting chats also deletes their cloud sessions.

			activity feed is kept for 6 years.	
Network access	Your "Allow network egress" setting controls network access from the sandbox, and Anthropic's egress proxy enforces it. Web search and web fetch are separate and not covered by this setting.	The same setting controls the local VM. Traffic leaves through the computer's own network and proxy, so your network controls can see it. Plugins and local MCP servers on the computer are not covered by the setting.	The same setting controls the cloud VM, and it is off by default for Enterprise. Anthropic's egress proxy enforces it and also blocks private and cloud-metadata addresses. The VM has no public address. Traffic leaves from Anthropic infrastructure, not from your network. As of this date, it does not come from Anthropic's published IP ranges.	The same setting controls cloud sessions. Traffic leaves from Anthropic infrastructure, so your IP allowlists and proxies do not see it. As of this date, it does not come from Anthropic's published IP ranges.
Access to the user's computer	None, apart from tools the app declares, such as local MCP servers.	Work runs on the user's computer, in folders they connect. The desktop app can also act as a linked computer for cloud sessions. It uses an outbound connection, and no inbound port is opened.	Claude can use a computer only if the member links it through the desktop app, and only while the app is online. It can then use folders the member grants, a shell in the local VM, the screen and browser, and local MCP servers. Connections are outbound only. Separately, Claude can use Claude in	Same as Cowork in the cloud. Claude can also use Claude in Chrome whether or not the conversation's cloud session is in use.

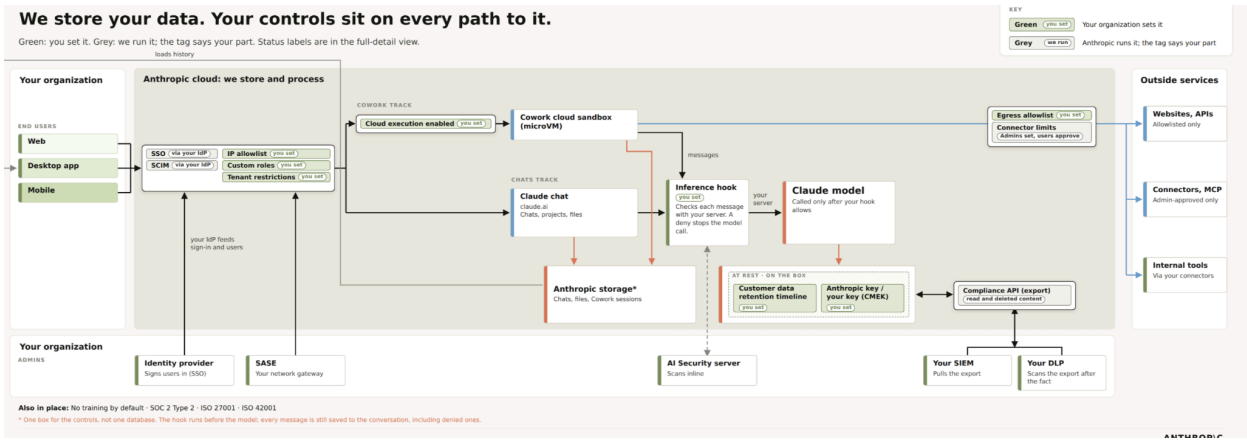
			Chrome where the member has installed and connected it. The "Remote control" setting is on by default. Turning it off stops web and phone sessions from using a linked computer.	
Connectors and credentials	Connectors are called from Anthropic's cloud. Anthropic holds connector tokens on its servers.	Remote connectors work the same as in Claude chat. Local MCP servers and desktop extensions run on the computer with the user's own credentials.	Connector tokens never enter the VM. Each connector call first passes a per-session policy check. Anthropic's connector gateway then adds the user's token. The VM holds two short-lived session tokens, which cannot call connectors directly.	Same as Cowork in the cloud. Code in the VM can read its two short-lived session tokens, but they cannot call connectors directly. Your connector policies apply, and existing connectors keep working.
Admin controls and approvals	Capability settings and Enterprise custom roles. Connector policies apply, such as Blocked and Needs approval.	The "Cowork" setting, which is on by default for Enterprise. Device (MDM) configuration keys also apply. Work in connected folders and the local VM follows the approval mode the user picks for the session. Connector	The "Cowork in the cloud" setting. It is off by default for Enterprise and on by default for Team. Remote control, Require trusted devices, network egress, connector policies and tool-approval settings also apply. Shell and file work inside the VM does not ask for approval.	It has its own setting, which is off by default. It needs Chat and Code execution turned on, but not Cowork or Cowork in the cloud. Actions outside the VM ask the user for approval. In "Auto" mode, a classifier decides on connector calls instead. If the classifier fails, the call is blocked.

		policies apply as in Claude chat.	Claude asks the user before calling a connector that has no standing approval, and before using a linked computer. In "Auto" mode, a classifier decides on those connector calls instead. If the classifier fails, the call is blocked. Turning off Auto stops new sessions from starting in Auto. A session already running in Auto stays in Auto until its user switches it.	Requests to use a linked computer always go to the member. If an admin turns off Auto for the organisation, new use stops. A conversation already in Auto switches to manual approval at the user's next message, with a few exceptions. Admins can also turn off "Always allow" for connector tools.
Audit	The Compliance API (Enterprise) covers chats, files and projects. The activity feed and audit logs also apply.	The Compliance API "local sessions" record (Enterprise) shows what Claude was asked and what it returned. It does not show what happened on the computer. OpenTelemetry export is also available.	The Compliance API "remote sessions" endpoints (Enterprise) list sessions and transcripts, read-only. Session-created and session-deleted events appear in the activity feed.	The Compliance API chats endpoints return the conversation as one record, including the cloud session's tool calls. Setting changes and session events appear in the activity feed. Ask your account team about OpenTelemetry coverage for cloud sessions.
Regulated configurations	HIPAA-ready (with a BAA), CMEK and custom retention are available on Enterprise. Not	Available to organisations with CMEK, or with ZDR on their API usage.	Not available to organisations with CMEK or ZDR. For HIPAA-ready organisations, it	As of this date, it is not available to HIPAA-ready, ZDR or CMEK organisations, and Anthropic blocks

	eligible for ZDR.	Cowork itself is not a ZDR service. For HIPAA-ready organisations, an admin must opt in. Not covered by the BAA.	is off by default. It cannot be turned on while Anthropic's HIPAA-ready restrictions on Claude Code and Cowork apply to your organisation. Not covered by the BAA.	access on its servers. Organisations with custom retention can use it.
Identity and data residency	SSO, SCIM, IP allowlisting and tenant restrictions apply. Data is processed and stored in US regions.	The desktop app signs in to your organisation and keeps its tokens in the operating system's keychain. Anthropic-side processing runs in US regions.	Each session belongs to the signed-in user in your organisation. Data is processed and stored in US regions.	No change from Claude chat. The same per-request checks apply, and the session runs as the end user, not as a service account. Data is processed and stored in US regions.

Part 1 — Solution architecture for unified Claude

Unified Claude architecture overview



In September 2026, Anthropic announced its intent to merge chats and Cowork into a single user interface, Unified Claude, for enterprise customers. From an architecture perspective, Unified Claude is a chat session that can upgrade itself to require a VM when a task needs agentic work, and downgrade back to no-VM mode when it doesn't, without the user switching modes. When a conversation elevates to agentic execution , the chat history, files and referenced project files are copied into the cloud hosted VM.

Unified Claude is built on a combination of Claude chat and Cowork in the Cloud (see "Part 4- the architecture of Cowork in the Cloud") but these distinctions will not be apparent to the end user. Admins control who gets Unified Claude with role-based access control. Security teams can approve Unified Claude by reviewing the merged architecture.

At the moment,it is not currently available to organizations with CMEK, ZDR or HIPAA requirements.

Why Anthropic Cloud hosted (remote) execution exists

Benefits to end users

Cloud-hosted execution improves the user experience in three ways:

1. **The laptop does not have to be on** — long-running and scheduled tasks keep going with the lid closed, and a session can be picked up from the browser or the phone;
2. **Moving to a new laptop is painless** — session history and state live with the user's account; and
3. **Unified Claude works on VDI (such as Citrix) and locked-down builds** that block the virtualization a local sandbox needs. If the local VM cannot start, unified Claude runs without it. See "What runs in the VM on the device" for what changes.

Benefits to security admins

Cloud execution also helps security: controls are applied once in the cloud instead of on every endpoint. The blast radius is contained since every session runs in an isolated cloud sandbox.

1. Contained blast radius: each session runs in its own locked, attested microVM that is destroyed at teardown, with no inbound traffic and no direct network route to other sessions or private networks.
2. Controls are enforced outside the agent, centrally: egress allowlists, connector tokens and tool approvals are all enforced server-side and configured once in the admin console, rather than per device through MDM
3. Less endpoint to defend: session history, outputs and code execution don't sit on the laptop, where retention and admin deletion can't reach them. Instead, they are

stored centrally and can be retrieved through the Compliance API. Security teams guard one central environment instead of thousands of endpoints.

What runs in the local VM on the device - why is virtualization required

Why is the VM needed? Unified Claude runs the code it writes and the commands it executes inside a dedicated virtual machine on the device, not directly on the user's operating system. That VM is hypervisor-isolated and runs a signed, integrity-checked image that Anthropic builds and ships; sessions can't alter it. Inside the VM, every session runs as its own unprivileged user in its own sandbox: it sees only the folders the user has explicitly connected to, and it can reach the network only through the destinations your administrators allow. The VM exists both for the local and cloud architecture with Claude Desktop. Examples of flows that require writing scripts:

1. Searching text in large folders.
2. Cleaning up a batch of files — resizing a folder of images.

However, reading and editing files in the connected folders happens outside the local VM, on the device itself, and is limited to those folders by unified Claude's permission rules.

If the VM is disabled or cannot start, for example on VDI or locked-down builds, tasks like the ones above still complete successfully: Claude does not run scripts on the device and does the work through file operations instead, so it takes longer. Based on recent usage data, only about 5% of Cowork sessions run commands in the VM, so most sessions are unaffected.

Why use the Claude Desktop App for unified Claude

For users — analysts, operations and finance teams, anyone whose day lives in files, browsers and internal apps — the desktop app is the simplest way to use Claude.

1. **One app for both shapes.** The same app runs sessions on the device and steers sessions in the cloud, so people switch shape without switching tools.
2. **Works on the files people already have.** Connect a folder and unified Claude reads and writes the documents, spreadsheets and decks there — no uploading, no copying out.
3. **Automates the browser people already use.** With Claude in Chrome and a built-in browser, unified Claude fills forms, pulls data and drives internal web apps using the logins the user already has, instead of asking IT for an integration.

4. **Reaches the tools on the machine.** Local connectors and desktop extensions bring the user's own tools into a session, within the limits IT sets.
5. **Long tasks keep running.** Long tasks and scheduled runs continue in the cloud with the laptop closed, and the user picks them up again in the same app.

Part 2 — Cowork and Unified Claude best practices

This section is relevant for customers currently using Cowork and migrating to unified Claude. These controls will get centralized over time under a single setting.

Settings that exist for Cowork and Unified Claude specifically: who may use it and where it runs, connectivity between the agent and your systems, autonomy, Cowork's own telemetry and records, cloud execution, and third-party inference. Where a connectivity control is an organization-wide setting shared with Claude chat (network egress, connectors, web tools), the text says so.

This is the configuration we recommend you review and start from. Deviations should be deliberate and written down.

Controls are set in three places. The "Set where" column tells you which, and names the console path or device key.

1. **The admin console (claude.ai organization settings; 1P only).** Most controls live here. They follow the org and apply wherever a user signs in.
2. **MDM device policy** (a managed configuration file pushed by Jamf, Intune, or Group Policy). Note that this is a separate payload from Claude Code's managed settings (though on a shared machine the two interact; see the note under the Cowork baseline table). Setup: Enterprise configuration for Claude Desktop App, with deployment steps for macOS and Windows. Applies to both 1P and 3P customers.
3. **Your network** (your web proxy or SASE). Only one control lives here: tenant restrictions, which limits which Claude organizations can be signed into from your network. You configure it in your proxy, not in Claude.

This table covers the 16 Cowork-specific settings; the 13 organization-level settings are in "Part 3". Each control name links to the section of the full guide that explains it.

Control	Baseline	In one line	Set where	Applies to
Cowork access and where it runs				
Cowork access	Explicit role grant	On by default for Enterprise. Turn the org default off and grant by role. Default: on for Enterprise.	Organization settings → Cowork → Enable for your organization; Organization settings →	Both Device and Cloud hosted

			Roles Docs: Use Claude Cowork on Team and Enterprise plans	
Unified Claude (chat and Cowork unified)	Explicit role grant	Off by default for Enterprise. Turn it on for the organization, or grant it by role. Default: off for Enterprise.	Organization settings → Capabilities → Chat and Cowork unified Organization settings → Roles Docs: Use Claude Cowork on Team and Enterprise plans	Cloud hosted
Cloud execution	Decide explicitly	See "Where the work runs" section. Default: off for Enterprise (on for Team). Beta feature.	Organization settings → Cowork → Run Cowork in the cloud Docs: Claude Cowork remote sessions	Cowork in the Cloud
Remote control (browser or mobile sessions reaching laptops)	Off, or on with trusted devices	Lets a cloud session driven from claude.ai or the mobile app reach files, tools, and apps on the user's computer via the desktop app. On by default once cloud execution is on. Default: on once cloud execution is on; Require trusted devices off by default; we recommend turning it on (one org-wide setting that also covers Claude Code Remote Control and unified Claude's links to a member's computer).	Organization settings → Cowork → Security → Remote control Docs: Use Claude Cowork on web, desktop, and mobile	Cowork in the Cloud
Scheduled tasks	Decide per populatio n	Lets Cowork tasks run on a schedule in the cloud. When off, scheduled tasks that execute remotely are disabled for the organization, and users can't create new scheduled tasks in the cloud. Default: on for Enterprise; off for HIPAA-ready organizations until an admin turns it on.	Admin settings Docs: Schedule recurring tasks in Claude Cowork	Cowork in the Cloud
Disabling Cowork via Claude Desktop App	Allowed for all populatio ns	Turns Cowork off entirely on that device — every Cowork session type in the Claude Desktop App — and keeps Chat. Cowork stays available to the user on claude.ai in the browser. Default: key not set — Cowork	MDM device policy: `secureVmFeaturesEnabled=false` Docs: Enterprise configuration for Claude Desktop	Cowork on the device

		allowed on the device.		
Connectivity				
VM network egress	Keep default, add narrowly	Default: off for Enterprise — code in the sandbox reaches only Anthropic's own hosts. Turn on "Allow network egress", keep "Package managers only" or list specific domains, and add your collector and named internal hosts.	Capabilities → Code execution → Allow network egress Docs: Use Claude Cowork on Team and Enterprise plans	Both Device and Cloud hosted
Workspace folders	Explicit allowlist	No limit on attachable folders until set. List specific working folders, read-only where possible. Also bounds what a cloud session reaches through the desktop app. Default: no allowlist — any folder can be attached.	MDM device policy: allowedWorkspaceFolders; or Organization settings → Capabilities → Allow workspace folders Docs: Enterprise configuration for Claude Desktop	Cowork on the device
Web search	On	Server-side. On or off org-wide, with domain allow- or block-lists; off if untrusted web content is unacceptable. Default: off for Enterprise until an Owner turns it on.	Organization settings → Capabilities Docs: Enabling and using web search	Both Device and Cloud hosted
Web fetch	On	Server-side. Only URLs from search results or the user. Default: on.	Organization settings → Capabilities; Also removable via MDM device policy: `disabledBuiltinTools` Docs: Use Claude Cowork on Team and Enterprise plans	Both Device and Cloud hosted
Claude in Chrome	Off	Acts inside the user's authenticated browser. Enable per role and per site only on a named case. Default: on for Team and Enterprise, unless an Enterprise organization turned it off before 10 September 2026; off for HIPAA-ready organizations until an Owner turns it on.	Claude in Chrome toggle plus your Chrome policy Docs: Claude in Chrome admin controls	Both Device and Cloud hosted
Connectors	Reads allowed, writes reviewed	Only what the work needs. Unreviewed tools default Blocked; each write tiered by consequence and granted per role. Default: no connector is	Organization settings → Connectors; custom roles Docs: Use Claude Cowork on Team and	Both Device and Cloud hosted

		available until an admin enables it; tool permissions are set per connector.	Enterprise plans	
MCP tunnels (internal MCP servers)	On, if internal MCP servers are needed	Connects Claude to MCP servers inside your network that aren't reachable from the internet, over an outbound-only connection with no inbound firewall rules. Tunneled servers are added as org connectors and follow the same connector policy. Works from web, desktop and mobile. Research preview, enabled per organization on request. Default: off.	Through your account team; then Organization settings → Connectors. Docs: MCP tunnels	Both Device and Cloud hosted
Locally configured MCP servers (set up on the device via Claude Desktop App; reachable by cloud sessions through the Desktop app)	Off	A developer path to connectors the console never sees: servers configured in the Claude Desktop App on the device, which cloud sessions used from mobile and web, can also reach through its bridge. MDM device policy is the only off switch and removes them from a cloud session's reach too. Default: allowed (key not set).	MDM device policy: isLocalDevMcpEnabled=false; or Organization settings → Plugins → Connectors → Allow local MCP servers. Docs: Enterprise configuration for Claude Desktop	Both Device and Cloud hosted
Desktop extension allowlist (extensions installed on the device via Claude Desktop App)	On	One-click local MCP servers, installed in the Claude Desktop App. Turn the allowlist on. Default: allowlist off — any extension can be installed.	To block all extensions: Organization settings → Plugins → Connectors → Block all desktop extensions Docs: Enterprise configuration for Claude Desktop	Both Device and Cloud hosted
Autonomy and capabilities				
Automatic approval (auto mode)	Off for most populations	Per-tool caps hold regardless of mode. Off for roles handling mostly inbound external content. Default: available to members (on).	Cowork → Permissions; custom roles Docs: Use Claude Cowork on Team and Enterprise plans	Both Device and Cloud hosted
Inline visualizations	On	Off means no visualization tool in the cloud session.	Capabilities	Cloud hosted
Sharing chats that have a cloud session	Decide explicitly	A chat that has a cloud session can be shared only when both "Share chats" and "Share cloud sessions" are on.	Data & privacy (Share chats); Claude Code settings (Share cloud sessions)	Cloud hosted

Cowork telemetry and records				
Compliance API	On	The server-side record of Cowork and Unified Claude session content and activity, feeding your SIEM and DLP. Default: available to Enterprise organizations; requires setup. The API supports reading sessions and deletion.	Data and privacy → Compliance API Docs: Access the Compliance API See security partner integrations.	Both Device and Cloud hosted
Open Telemetry export (OTEL)	On	Live per-user action and cost stream to your collector. Prompt and response text is not exported by default: on the local VM it is sent only if an admin opts in, and the cloud-hosted VM sends metadata only. Default: off until an endpoint is configured. Compliance API should be the preferred path and has the relevant OTEL data (see comparison).	Organization settings → Cowork → OTLP endpoint Docs: Monitor Claude Cowork activity with OpenTelemetry	Both Device and Cloud hosted

If the same machines also run Claude Code, test one real session before fleet-wide push. Two interactions are known. Two Cowork device keys also apply to the Code tab inside the desktop app: the workspace-folder allowlist gates its working directory too, and `disabledBuiltinTools` removes the named tools from Code sessions as well. That surprises teams that meant to restrict Cowork and leave developer tooling alone. (The console egress allowlist does not reach the Code tab on first-party.) In the other direction, a Claude Code managed-settings file on the device reaches Cowork: a shell deny rule there also blocks Cowork's shell tool, and "managed permission rules only" strips Cowork's built-in permission to write to the attached folder, so every write fails. (Claude Code's console-delivered server-managed settings do not reach Cowork at all.) Validate with both payloads present.

On third-party-inference deployments the console rows become device-policy keys.

Part 3 — Organization-level best practices

Settings that belong to the Claude Enterprise organization as a whole and that unified Claude inherits: identity (who can sign in, from where, and with what role), the content the agent runs on (plugins, scanning, organization instructions), the organization-level detection layers (the Compliance API's wiring and inference hooks), and data handling. Almost everything here is a precondition for Part 2.

The recommended baseline: organization-level settings

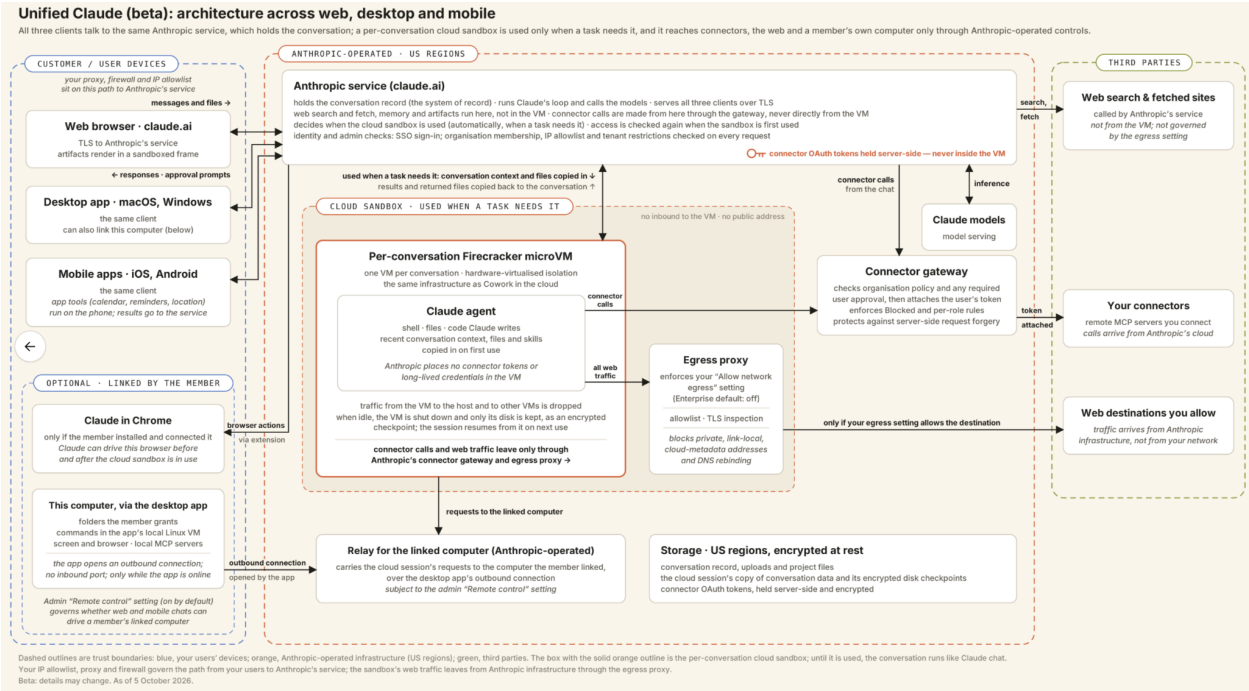
Each control links to the section of the full guide that explains it.

Control	Baseline	In one line	Set where	Applies to
Identity				
SSO, SCIM, verified domains	On	Provision and deprovision through your IdP; the basis for every other row. Default: not configured.	Identity and access Docs: Setting up single sign-on (SSO)	Both Device and Cloud hosted
Domain capture	On	Pulls personal accounts on your verified domains into the org.	Organization admin → Security Docs: Claim and migrate accounts on your domain	Both Device and Cloud hosted
Enterprise Domain Guard	On	Stops company systems being connected to personal Claude accounts. Off by default, so turn it on. Default: off.	Organization admin → Connector domains Docs: Restrict verified-domain connectors to your Enterprise	Both Device and Cloud hosted
Block new organization creation	On	Stops shadow orgs on your verified domains.	Organization admin → Security Docs: Claim and migrate accounts on your domain	Both Device and Cloud hosted
Session length	7 days	Default is 28. One day is the high-assurance option but breaks scheduled tasks daily. Default: 28 days.	Identity and access → Session duration	Both Device and Cloud hosted
Tenant restrictions	On, at your proxy	Blocks use of any org but yours from your network, on every request, not only at sign-in	Your SWG / SASE: allowed-org header Docs: Tenant restrictions (help-center article)	Both Device and Cloud hosted
Required organization on managed devices	Set	Same job as tenant restrictions, enforced in the desktop app on or off your network. MDM blocks outside accounts before anyone signs in. The Organization settings option needs no MDM. Default: any organization can sign in.	Organization settings → Organization and access → Security → Limit Claude Desktop to allowed organizations (shadow IT); MDM device policy: `forceLoginOrgUUID`; Docs: Enterprise configuration for Claude Desktop	Both Device and Cloud hosted
Content the agent runs on				

Plugins and marketplaces	Org-managed sources only	Admins stock the shelf, scanning gates additions, install policy per group. Default: plugins follow the Cowork toggle; no separate on/off. Source restrictions not set.	Organization settings → Plugins Docs: Manage Cowork plugins for your organization	Both Device and Cloud hosted
Skill and plugin security scanning	On	Scans every skill and plugin on upload or edit; a fail blocks it org-wide. Opt-in. Default: off (opt-in).	Organization settings → Skill and plugin scanning	Both Device and Cloud hosted
Organization instructions	Set, org-specific	Classification labels, never-search terms, regulatory identity, where to get help. Not general safety rules. Default: none set.	Organization admin → Organization instructions	Both Device and Cloud hosted
Detection				
Inference hooks (inline DLP)	Shadow mode, then enforce	Your DLP allows or denies every prompt and tool call before it executes. Default: none configured.	Security settings → Inference hook endpoint Docs: Inference hooks overview	Both Device and Cloud hosted
Data handling				
Retention window	Set to your policy	One org-wide window, commonly set to 30 or 90 days. Does not reach Cowork data on the device, which stays until deleted; does not yet govern cloud hosted Cowork (see Data and lifecycle in the cloud appendix). below). Default: no custom window set.	Data and privacy → Retention Docs: Configure custom data retention controls for Enterprise plans	Both Device and Cloud hosted
Customer-managed keys	No, unless required	Cowork on Device supported, not on the Cloud.	Through your account team Docs: What are customer-managed encryption keys (CMEK)?	Both Device and Cloud hosted

Part 4 — the architecture of Cowork in the Cloud

Unified Claude architecture - Detailed



Security properties of Cloud hosted (remote) sessions for unified Claude and Cloud based Cowork

Isolation

1. **Per-session microVM, destroyed at teardown.** Each Remote Cloud session runs in its own Firecracker microVM (KVM) with its own guest kernel; resource limits are pre-set and not adjustable from inside, and the VM is destroyed when the session ends.
2. **Designed to contain a compromised agent.** The controls are designed to contain a prompt-injected or otherwise compromised agent, which is the framing of this whole posture.
3. **Compute is per session; storage is multi-tenant, as it is for chat.** The microVM isolates only the running session. Session data is stored on Anthropic's shared multi-tenant infrastructure, the same systems that hold your organization's Claude chat data, and is protected by the same access controls.
4. **Supply chain: attested images, version-pinned agent.** Sandbox images are attestation-verified at launch and the agent binary runs at a version Anthropic selects; runtime packages come only from registries the org's network policy allows.

Identity and credentials

1. **Connector credentials are kept out of the sandbox.** OAuth tokens are attached at Anthropic's connector proxy, not inside the VM, so the agent never has broader third-party access than the user.
2. **Session tokens expire on their own within hours rather than being revoked individually;** deleting a session stops new ones from being issued.
3. **Fixed built-in tool set per surface.** The agent's built-in tools are a fixed set Anthropic configures for each surface; users and admins cannot add to it; on a device, MDM device policy can remove named tools.
4. **The VM holds two short-lived, session-scoped tokens, for model inference and file upload.** Code running in the VM can read them. They are scoped to the session and to those two purposes. Connector calls from the VM pass through a per-session proxy that checks organization policy and any required approval.

Network and egress

1. **No inbound traffic to sandboxes.** Sandboxes accept no inbound connections; all traffic is outbound through the governed egress path.
2. **Private-range and SSRF protections on egress.** Private address ranges, loopback, link-local, and cloud metadata endpoints are blocked at multiple layers, with DNS-rebinding defenses.
3. **Egress logs are internal to Anthropic.** Anthropic logs network egress from cloud sandboxes for internal monitoring. Customers cannot access these logs, and no export or API exists for them. For visibility into what a session did, use the Compliance API and OpenTelemetry export.

Device bridge and local tools

1. **Local tools bridge: a Remote Cloud session can use local MCP servers and connected folders through the Claude Desktop app.** When the user drives a Remote session from the Claude Desktop app (or from a web browser or a mobile phone with Remote control on), the desktop keeps an outbound WebSocket connection to an Anthropic-operated bridge, and through it the remote agent can call user-approved tools on that device — locally configured MCP servers (tool by tool, under the allow/ask/deny policy) and the connected folders — while the agent itself keeps running in Anthropic's sandbox; tool calls are policy-checked server-side before they reach the desktop. Within granted folders, file transfers and shell commands in the local VM don't prompt again. The organization bounds this with the workspace-folder allowlist and the `isLocalDevMcpEnabled=false` device key.
2. **Local file mount modes: read-only, read-write-no-delete (default), read-write with delete.** A folder connected to a session is mounted read-only, read-write without

delete (the default; deletes are blocked at the filesystem layer), or read-write with delete on an explicit per-session grant. This applies to folders attached to a local session and to folders a Remote session reaches through the desktop app.

Data and lifecycle

1. **Session data is stored under the same Anthropic data-security policy as chat data.** Remote-session conversation history, files and artifacts are held in Anthropic systems, encrypted at rest and in transit, with the same access controls and enterprise terms as chat.
2. **The Compliance API** returns sessions, attachments and generated files. Data is more extensive in Unified Claude and Cowork in Cloud (comparison here).
3. **Idle snapshot after about 5 minutes, encrypted.** After idle time, the VM is snapshotted (memory and disk), encrypted, and torn down; a new message restores the files into a fresh VM and restarts the session
4. **Sandbox state expires 30 days after creation;** Sandbox state and snapshots expire 30 days after the session was created.
5. **Custom retention window does not yet apply.** Unlike chat, cloud session history is not currently subject to your organization's custom retention window. Support is planned for the Unified Claude public beta. See more data here.
6. **Data residency.** Data is stored in US regions. There is no EU data-residency option today.
7. **Data from Claude is not used to train Anthropic's models**, and cloud-session prompts, responses, uploaded files and artifacts are handled under the same commercial data-use commitments (Teams and Enterprise).

Appendix: Stored data on Anthropic’s Cloud for Cowork on Device

Most of a Cowork-on-the-device session stays on the user's computer. Conversation history, the connected folders, the files Claude reads and writes, the files a user attaches and the outputs Claude produces all live on the device, where your organization's retention window does not reach them; they remain until the user deletes them.

Every model request, including the contents of files Claude reads and screenshots it takes, is sent to Anthropic's API. Web search, web fetch and remote connectors run from Anthropic's cloud.

Examples of what is stored on Anthropic's servers: the session's model traffic (prompts, responses, tool calls and results), recorded as the session transcript for the Compliance API when it is enabled, including for trust and safety review; skills a user saves from Cowork, which go to the organization's library in claude.ai; and, if your organization enables Cowork Scheduled Tasks, their definitions, schedules and run history — even when a routine is bound to a user's computer and works on its folders.

Appendix: first-party vs. third-party inference compared

First-party (1P) = Claude Enterprise with Anthropic running the model; **third-party (3P)** = the same Claude Desktop app with the model served from your Amazon Bedrock, Google Vertex AI, or Microsoft Foundry account.

Recommendation: first-party (1P), unless one of the following rules it out: a requirement for in-region inference outside the US, a policy forbidding any processing on Anthropic-hosted infrastructure, or an air-gapped environment.

Control	First-party (1P)	Third-party (3P)
Cloud execution, mobile, Claude in Chrome, user-to-user sharing	Available (cloud execution off by default for Enterprise; Claude in Chrome on by default)	Not available
Device hosted execution	Available	Available
Where inference runs	Anthropic, US-only available	Your Bedrock, Vertex AI, or Microsoft Foundry account, in the region you choose
Admin console	Admin console plus MDM device policy	MDM device policy instead of the admin console
Connectors	Console-managed, per-tool and per-role policy, hosted by Anthropic	Pushed by device policy and run locally on the device; per-server tool policy in the same payload
Data controls by Anthropic	Compliance API, inference hooks and OpenTelemetry export	OpenTelemetry export only
Other Claude products	The full Claude product stack — Claude Tag, Claude	N/A

	Security and the admin console	

Appendix: Additional FAQs

1. Zero data retention (ZDR) - Anthropic's ZDR configuration does not apply to Cowork, because the product is stateful and ZDR would break the user experience.
2. Cowork cannot be disabled for mobile alone; to remove it from mobile, block the Claude mobile app.
3. An LLM gateway is supported only with 3P inference, configured as the inference provider.
4. Compliance API storage can't exclude specific inputs - everything that reaches Anthropic is stored with the session. To keep content out, block it upstream with inference hooks that deny matching prompts and tool calls. Legal hold is not directly supported; customers can pull session data through the Compliance API. Sandbox state and snapshots expire after 30 days and cannot be exported.
5. Proxy, VPN and firewall rules - your on-prem and device network controls do not apply to cloud sessions, because their traffic leaves from Anthropic infrastructure, not your network. Use the "Allow network egress" setting to control it, and set an IP allowlist in your Claude organization settings to limit where users can connect from. To limit where users can connect from, Owners can set up and manage an IP allowlist themselves.
6. EDR integration - On the desktop, the VM is isolated from host-based security tools by design: your host EDR sees the agent process and its host file activity, not the inside of the VM. With the EDR integration (public beta), a supported EDR sensor can now run inside the local VM on the user's device, giving you visibility into that activity. Compliance API and hooks provide additional visibility and control.
7. Unified Claude has been rolling out gradually to personal accounts since 09/2026.

This guide is provided for information only, to help you configure Claude. Features and defaults, including beta features, may change. Your organization is responsible for its own configuration and for deciding whether its use of Claude meets its legal and regulatory obligations.