

ZeroTrusted.ai

Global Pricing Description — 2026 Edition

Deployment Guidance, Support Services, International Pricing, and DoD / Federal CLIN Reference

Prepared for partner and customer use. Companion to the ZTA Master Pricing Calculator (2026 v2). Final commercial terms remain subject to scope, architecture, integration count, support requirements, tax treatment, and executed commercial terms.

1. Executive Summary

ZeroTrusted.ai delivers AI governance, data protection, and runtime security for organizations deploying generative AI and agentic systems at scale. This pricing document describes how ZeroTrusted.ai is sold across three consumption meters (per-user, token, and API), across three deployment postures (on-premise, private cloud, and SaaS), and across ten commercial regions (United States, Brazil, Japan, European Union, Australia / New Zealand, GCC, India, Singapore, South Korea, South Africa).

The primary deployment models are on-premise and customer-hosted private cloud. SaaS is retained as a small-organization on-ramp but is not the recommended path for regulated, government, or enterprise customers. All deployment models require bring-your-own LLM compute — tokens or GPUs are provided by the customer through their own infrastructure or cloud tenancy; ZeroTrusted.ai is the governance and protection layer, not the inference provider.

A separate DoD / Federal scenario is included with CLIN-structured pricing covering software subscription, implementation, training, and support services for small, mid, and large-scale deployments at IL4, IL5, and IL6 impact levels.

2. Pricing Meters — When to Use Each

Customers choose a primary meter based on how ZeroTrusted.ai is consumed in their environment. Meters may be combined in a single contract when different parts of the organization use the platform differently.

Meter	Best Fit	Primary Environment	Commercial Logic
Per-User	Large enterprises and agencies where named users access ZTA directly through a hosted, private, or on-premise environment	Broad workforce adoption across business units, security teams, and analysts	Value scales with headcount; predictable licensing across a known user population
Token	Organizations protecting internal or private model workloads, or embedding ZTA into their own agents and portals with BYO LLM compute	Consumption-led environments where token volume is the clearest meter of value	Value scales with model activity; customer supplies GPUs or LLM API tokens
API	Teams integrating ZeroTrusted.ai into portals, security tools, AI monitoring platforms, or agent frameworks	Agent-heavy, automation-focused deployments with system-to-system traffic	Value scales with API call throughput; buckets from 1M to 250M calls per month

2.1 Optional Scope Modules

The following modules can be added to any meter. They stack on the core subscription and are priced as an uplift rather than a separate SKU.

- AI HealthCheck — adds 20% to the core subscription. Provides continuous posture assessment of AI systems, model drift detection, and compliance attestation.
- Shadow AI Protection — adds 40% to the core subscription. Discovers, catalogs, and governs unsanctioned AI use across the enterprise (shadow LLMs, rogue agents, unapproved model endpoints).
- AI SOAR — available on a per-user basis only. Priced against supported users plus platform users. Orchestrates AI-specific security response playbooks and integrates with existing SIEM/SOAR stacks.

2.2 Base Rate Card (USD, Pre-Tax, Pre-Distributor)

Meter	Base Price	Unit	Notes
Per-User	\$149.00	per user / year	Tiered discounts at 150, 500, 1K, 2.5K, 6K users (15% to 40%)
Token	\$149.00	per 1M tokens / year	Tiered discounts 15% to 40% from 5M to 6B tokens annually
API	\$0.00025	per API call	Monthly buckets from 1M to 250M calls

2.3 Per-User Volume Tiers

Minimum Users	Tier Discount	Price per User per Year	Price per User per Month
1	0%	\$149.00	\$12.42
150	15%	\$126.65	\$10.55
500	25%	\$111.75	\$9.31
1,000	30%	\$104.30	\$8.69
2,500	35%	\$96.85	\$8.07
6,000	40%	\$89.40	\$7.45

Volumes above 6,000 users are quoted individually. For federal, state, and local government customers, negotiated rates below the published tier floor are available under prime-contractor or schedule vehicles (GSA, SEWP, etc.).

3. Deployment Models — On-Premise and Private Cloud First

ZeroTrusted.ai is primarily deployed on-premise or into a customer-owned private cloud tenancy. SaaS is offered as an on-ramp for small organizations and proofs of concept, but is not the recommended posture for regulated, classified, or enterprise production workloads.

Deployment Model	Implementation Fee	Primary Use Case	Architecture Notes
On-Premise / Special Architecture	20% of Year-1 subscription	Primary model for Government, DoD, and regulated industries. Classified enclaves, air-gapped networks, SIPR/NIPR, and data-sovereignty-bound deployments.	Customer datacenter, isolated enclave, or dedicated hardware. ZTA runs fully within customer-controlled infrastructure. Sharper terms available for large opportunities — the cost profile is concentrated in software plus maintenance and operations.
Private Cloud (customer-hosted)	15% of Year-1 subscription	Primary model for enterprise and commercial customers who want isolation without operating their own datacenter.	ZTA deployed into customer AWS, Azure, GCP, Salesforce, or Red Hat OpenShift instance. Customer retains full control over VPC/tenant, encryption keys, identity, and network egress.
SaaS / Hosted	10% of Year-1 subscription	Small organizations, proofs of concept, and pilot deployments. Not recommended for regulated or government workloads.	ZTA-managed environment aligned for customer onboarding and tool connection setup.

3.1 Bring-Your-Own Compute

Across all three deployment models, LLM inference compute is not included in the ZeroTrusted.ai subscription. Customers supply their own tokens (through commercial LLM API agreements), GPUs (through cloud or on-premise hardware), or both. This reflects the nature of ZeroTrusted.ai as a governance and protection layer rather than a model provider, and it keeps subscription pricing independent of volatile inference-compute markets.

For government customers, this separation also enables the use of authorized government-furnished compute infrastructure (GovCloud, IL5/IL6 Azure Government, AWS GovCloud, or on-premise GPU clusters) without disrupting the software subscription commercial structure.

4. International Pricing, Distributor Treatment, and VAT

The companion calculator now includes dedicated regional sheets for the United States (direct list), Brazil, Japan, the European Union, Australia and New Zealand, the Gulf Cooperation Council (UAE and Saudi Arabia), India, Singapore, South Korea, and South Africa. Each region has three tabs — per-user, token, and API — and presents both private cloud and on-premise options side by side in a single view.

All international (non-US) regional sheets apply a 20% distributor discount on subscription value and overlay local VAT or GST. Implementation fees are calculated on pre-distributor Year-1 contract values, with VAT then applied. A centralized Assumptions and FX sheet holds every rate and FX pair; changing a value there propagates to all downstream sheets.

4.1 Regional VAT / GST Summary

Region	Currency	VAT / GST	Notes
United States	USD	Pre-tax	State sales tax applies per jurisdiction; federal contracts generally pre-tax
European Union	EUR	21% (avg)	EU average; range 17% (Luxembourg) to 27% (Hungary). Adjust per member state at quote time
Brazil	BRL	18%	VAT-inclusive subscription with 20% distributor discount

Region	Currency	VAT / GST	Notes
Japan	JPY	10% (verify)	Consumption tax 10%. Calculator uses 18% historically — verify at quote
Australia	AUD	10%	GST on digital services
New Zealand	NZD	15%	GST on digital services (separate from AUD)
UAE	AED	5%	VAT on digital services
Saudi Arabia	SAR	15%	VAT raised from 5% in 2020
India	INR	18%	GST on cross-border digital services (OIDAR regime)
Singapore	SGD	9%	GST effective January 2024
South Korea	KRW	10%	VAT on digital services; B2B reverse-charge typically applies
South Africa	ZAR	15%	VAT on digital services

Rates valid as of April 2026. The EU representative rate of 21% is a modeling convenience; always charge the customer-country rate at invoice. Japan consumption tax in the legacy calculator sheets was modeled at 18% — the actual rate is 10% and should be verified before quoting in that region.

4.2 International Distributor Treatment

- All non-US regional sheets apply a 20% distributor discount to subscription value before local VAT / GST.
- Implementation fees are calculated against pre-distributor Year-1 contract values, then VAT / GST is applied on top.
- Local-currency amounts are computed from USD using the FX rates on the central Assumptions and FX sheet; update in one place for all downstream sheets.
- Final commercial terms should always be confirmed against the customer's tax jurisdiction, reseller structure, contracting vehicle, and any local invoice requirements.

5. Support, Training, and Installation Services

The following service rate card applies globally. On-site engagements add travel and per diem at actual cost (GSA rates for US federal customers). All service rates are billed in USD and converted to local currency at point of invoice.

Service	Rate	Terms
Remote Support or Training (ad-hoc)	\$350 per hour	4-hour minimum. VPN or cloud-connected delivery.
On-Site Support or Training (ad-hoc)	\$350 per hour + travel	4-hour minimum, plus travel, per diem, and billable travel time.
Remote Deployment Package (env. ready)	\$3,000 fixed	Kubernetes and supporting environment already in place. Scope: 1 day setup + 1 day training.
Remote Deployment Package (K8s assist)	\$5,000 fixed	Includes Kubernetes setup and configuration. Scope: 2 days setup + 1 day training.
On-Site Deployment Engagement	\$7,000 + travel	Scope: 2 days setup + 1 day training on customer site.

Typical enablement day includes approximately 8 hours of hands-on technical training and 4 hours of executive training. Installation is generally quick when infrastructure is already available; most effort is coordination, access, tool integration, connection setup, and knowledge transfer. Delays are typically caused by missing prerequisites or access gaps rather than software installation itself.

6. DoD and Federal Scenario — CLIN-Structured Pricing

The companion calculator includes a dedicated DoD Scenario and CLIN sheet that structures pricing in a format recognizable to federal contracting officers. Six reference scenarios are provided — Small Program Office, Mid-Size Command, and Large Enterprise, each in both Private Cloud (GovCloud) and On-Premise (classified enclave) variants.

6.1 Reference Scenarios

Scenario	Size	Impact Level	Deployment	Year 1 (Rough Order)	5-Year TCV
Small Program Office	250 users / 50M tokens	IL4 (CUI) / GovCloud	Private Cloud	≈ \$186K	≈ \$663K
Small Program Office	250 users / 50M tokens	IL5 (NSS) / On-Prem	On-Premise	≈ \$211K	≈ \$733K
Mid-Size Command	2,500 users / 500M tokens	IL5 (NSS) / GovCloud	Private Cloud	≈ \$687K	≈ \$2.80M
Mid-Size Command	2,500 users / 500M tokens	IL5 (NSS) / On-Prem	On-Premise	≈ \$707K	≈ \$2.82M
Large Enterprise	10,000 users / 2.5B tokens	IL5-IL6 / GovCloud	Private Cloud	≈ \$2.27M	≈ \$9.59M
Large Enterprise	10,000 users / 2.5B tokens	IL6 / SIPR On-Prem	On-Premise	≈ \$2.34M	≈ \$9.66M

All scenarios assume bring-your-own compute (Government-furnished tokens or GPUs). Pricing is rough-order magnitude from the calculator; final pricing requires scope confirmation, architecture review, and integration count.

6.2 CLIN Structure

Each scenario is broken out as a four-CLIN structure for base year plus four option years, consistent with FAR 52.217-9 (Option to Extend the Term of the Contract). This matches the structure most program offices expect to see in a firm-fixed-price proposal.

CLIN 0001 — Software Subscription (Annual)

- 0001AA — Core ZTA Subscription (per-user, per-year)
- 0001AB — AI HealthCheck Add-On (+20% of Core)
- 0001AC — Shadow AI Protection Add-On (+40% of Core)

CLIN 0002 — Implementation and Deployment (One-Time, Base Year)

- 0002AA — Deployment Engagement (15% Private Cloud / 20% On-Prem of Year-1 Subscription)
- 0002AB — On-Site Deployment Package (\$7,000 fixed — 2 days setup + 1 day training)
- 0002AC — Travel and Per Diem (NTE ceiling, GSA rates, actual expense with receipts)

CLIN 0003 — Training Services

- 0003AA — Technical Training (1 day per 100 users, minimum 1 day)
- 0003AB — Executive Training (2 sessions: kickoff and mid-year)
- 0003AC — Train-the-Trainer Program (3 days, internal admin enablement)

CLIN 0004 — Technical Support Services

- 0004AA — Remote Technical Support (hours pool, approximately 1 hour per 20 users per year, minimum 40 hours; 4-hour minimum per incident)
- 0004AB — On-Site Support (NTE days, 4-hour minimum per visit plus travel and per diem)
- 0004AC — Compliance and ATO Support (scaled by impact level: IL4 = 16 days, IL5 = 24 days, IL6 = 40 days; covers FedRAMP, DoD SRG, STIG, SSP and POA&M artifacts, and ATO package support)
- 0004AD — Integration and Professional Services (80 hours base for custom connectors and API integrations; scales with scope)

6.3 Terms and Assumptions

- Pricing Type — Firm-Fixed-Price (FFP) for subscription, implementation, and training CLINs. Support services priced Time-and-Materials (T&M) against NTE ceilings.
- Period of Performance — 12-month base period plus four 12-month option periods, exercised at Government discretion under FAR 52.217-9.
- Compute — All scenarios require Government-furnished tokens or GPUs. No LLM inference compute is included in the ZeroTrusted.ai subscription.
- Classification — IL4 covers Controlled Unclassified Information (CUI); IL5 covers National Security Systems (non-classified); IL6 covers classified information up to SECRET and SIPR. FedRAMP High is the minimum posture for IL4 and IL5 cloud deployments.
- Travel — GSA rates, invoiced at actual cost with receipts against a per-CLIN NTE ceiling.
- Distributor / Reseller — US federal pricing shown at direct list; applicable GSA schedule, SEWP, or prime-contractor markups are applied separately per the contracting vehicle.
- Tax — US federal pricing is pre-tax. State sales tax applies per jurisdiction where applicable.
- Option-Year Escalation — Option years shown flat to Year 1 as a conservative planning baseline. Customary 3% annual escalation is negotiable.
- Cyber Compliance — Contractor is SOC 2 Type II. FedRAMP Moderate status to be confirmed per award. CMMC Level 2 attestation under DFARS 252.204-7021 applies where CUI is processed.

7. Pricing Decision Guidance

The following framework helps partners and customer-facing teams position the right meter and deployment model for each opportunity.

7.1 Choosing a Meter

- If the customer has a known, stable user population and wants predictable annual licensing, use per-user pricing. This is the default for federal, DoD, and large enterprise.
- If the customer is protecting internal or private model workloads and token volume is the clearest meter of value, use token pricing. This is the default for data-science organizations with heavy internal LLM use.
- If ZeroTrusted.ai is being embedded into customer portals, security tools, or agent platforms via system-to-system API calls, use API pricing. This is the default for ISVs and platform-embedded deployments.
- Mixed meters are supported within a single contract when different parts of the organization consume the platform differently.

7.2 Choosing a Deployment Model

- For Government, DoD, intelligence community, or classified workloads, default to on-premise in a customer-controlled enclave. Quote at 20% implementation.
- For regulated commercial (financial services, healthcare, critical infrastructure) or for enterprise customers with existing cloud strategy, default to private cloud in the customer's own AWS, Azure, GCP, Salesforce, or Red Hat tenancy. Quote at 15% implementation.
- For small organizations, pilots, proofs of concept, or customers who explicitly prefer managed hosting, SaaS is available at 10% implementation. Confirm data-residency and compliance requirements before offering SaaS to any regulated customer.

8. Partner Support and Deal Flexibility

Our goal is to make pricing simple and repeatable for partner teams while still leaving room to align architecture, commercial structure, and delivery scope to each opportunity. ZeroTrusted.ai can support demos, customer briefings, technical materials, tailored messaging, and proof-of-concept engagements.

For larger opportunities — particularly on-premise and private cloud deals where the cost structure is more software-driven and less dependent on recurring hosting overhead — pricing can often be improved further through negotiation. Government and federal engagements with long-term option-year commitments also typically unlock sharper terms.

Please cc contact@zerotrusted.ai on pricing or technical coordination so the full team can respond quickly.

Commercial guidance summary only — final pricing remains subject to scope, architecture, integration count, support requirements, tax treatment, and executed commercial terms. All figures derived from the companion Master Pricing Calculator 2026 v2.