

TC-AICC可信计算

产品简介



腾讯云

【 版权声明 】

©2013–2026 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

文档目录

产品简介

产品概述

产品优势

应用场景

产品简介

产品概述

最近更新时间：2026-10-28 15:46:00

产品定义

TC-AICC可信计算（Artificial Intelligence Confidential Computing, AICC）是一种应用在 AI 交互、训练、推理、计算等场景的云安全技术方案；基于可信执行环境（Trusted Execution Environment, TEE）、远程证明等机密计算技术，在 AI 处理全链路中为数据与模型提供“使用中”的加密保护，实现数据“可用不可见”，帮助解决数据隐私与 AI 效能难以兼顾的难题。

主要功能介绍

端到端加密

混合抗量子加密：客户端加密请求，模型在可信执行环境内解密并完成推理，结果加密返回。密钥交换采用经典椭圆曲线和抗量子密码混合的密钥算法，经典椭圆曲线密钥交换算法（X25519）提供成熟、高效的经典密码学保护，抗量子加密算法（ML-KEM-768）用于抵御量子计算对传统公钥密码体系的威胁，预防先截获、后解密的长周期攻击。

机密计算 TEE 内存加密：由处理器内置的内存加密引擎对 TEE 内的代码与数据进行透明加密，数据在内存中仅以密文形式存在。

密钥与运行环境安全

密钥安全管理：会话密钥统一由密钥管理系统（Key Management System, KMS）进行硬件级管控，覆盖密钥的生成、存储、分发、使用与销毁等全生命周期。

自带密钥（Bring Your Own Key, BYOK）：支持使用您提供的密钥材料，密钥由您自行管理，平台侧不以明文形式持有。

多重隔离：对推理镜像进行安全改造，实现对网络通信、登录访问、磁盘读写等行为的管控约束，使推理节点保持密文进、密文出。

安全启动：基于推理节点的硬件信任根实现可信启动，保证推理环境从启动阶段即具备可验证的身份与状态。

可信证明

远程证明推理：推理节点运行时，由 Agent 采集推理节点的网络状态、登录状态、可信度量等安全状态，用于证明其数据完整性与不可篡改，并在运行期间持续监控。

安全证明报告：证明范围覆盖端侧采集、远程证明、密钥、容器、模型调用至运维的完整链路。远程证明逐项校验各项环境度量指标（当前覆盖 13 项以上），未通过校验时链路不通；提供哈希摘要锚定至芯片级硬件签名证明。环境发生篡改时，可由下游校验发现。可以导出报告，对推理节点做独立安全审计。

接入方式

客户端双通道灵活接入：支持安全代理（AICC-Proxy）接入与软件开发工具包（Software Development Kit, SDK）接入两种方式，两者均兼容 OpenAI 与 Anthropic 标准协议。安全代理适用于不便修改业务代码的存量应用；SDK 适用于需要在客户端集成端到端加密能力的新建应用。

模型支持

支持大模型：当前支持腾讯混元、DeepSeek、Kimi、智谱 GLM、通义千问等多款大模型。由于模型迭代较快，完整的支持清单请以实际开通展示为准。TC-AICC 作为可信推理底座能力，与模型本身解耦，新模型的适配将随平台模型的上线节奏同步推进。

产品优势

最近更新时间：2026-10-28 15:46:00

高效验证，安全可证明

AICC 提供技术验证平台与配套测试工具，支持您对安全推理全链路进行验证并输出审计日志报告，可用于内部合规审查，并为外部机构调查沟通提供技术材料。

- **全链路验证**：验证范围覆盖数据传输、密态推理与结果返回各环节，验证过程与结果以报告形式留存，便于按需调取。
- **远程证明**：每次推理前返回芯片级远程证明报告，用于核验推理环境的硬件、固件与代码完整性；运行期间持续采集环境度量信息并与基线比对。
- **密钥与凭据可核对**：支持自带密钥（Bring Your Own Key, BYOK）与密钥管理系统（Key Management System, KMS）托管，会话密钥在推理任务结束后自动销毁，密钥的生成、使用与销毁时机均可核对。
- **审计日志记录留存**：提供调用日志与操作记录查询能力，为合规审查提供过程性材料。

适配信创，支持国产化算力

AICC 的安全能力基于可信执行环境（Trusted Execution Environment, TEE）与可信平台模块（Trusted Platform Module, TPM）构建，与具体加速卡型解耦。国产加速卡与 NVIDIA 生态的加速卡均可纳入统一的可信集群，为有国产化要求的场景提供算力选择空间。

- **卡型解耦**：安全能力不绑定特定卡型，无需关注底层卡型差异；算力资源可在统一算力池中调度。
- **国产化算力**：支持以国产加速卡承载机密推理服务，适用于国产化场景要求。
- **供应链风险控制**：算力来源不依赖单一供应商，可降低因未知供应链风险带来的业务连续性隐患。
- **统一运维口径**：不同卡型的可信集群在创建、变更与删除流程上保持一致，不因卡型差异引入额外的运维流程。

便捷接入，低性能损耗

AICC 提供安全代理接入模式，您可在业务侧部署 AICC Proxy，并将模型调用地址（BaseURL）指向该代理，无需修改业务代码即可接入机密计算环境，远程证明与端到端加密均由代理自动完成。已覆盖多种模型即服务（Model as a Service, MaaS），并原生集成于高性能应用服务 HAI、Tokenhub、TI-ONE、WorkBuddy 等产品。通过硬件加速及异步验证策略，将对业务延时的影响控制在较低水平，兼顾推理吞吐与并发能力。

- **双模接入**：同时提供安全代理与 SDK 两种接入方式。安全代理适用于不便修改业务代码的存量应用；SDK 适用于需要在客户端集成端到端加密能力的新建应用。
- **已集成场景**：覆盖 MaaS 模型调用与智能体应用等场景，相关产品可在既有调用链路中直接启用机密推理能力。
- **性能开销控制**：通过硬件加速与异步验证策略，将安全能力引入的额外开销控制在较低水平，在启用机密计算的同时兼顾推理延时与并发能力。

- **协议兼容：**兼容 OpenAI 与 Anthropic 协议，现有应用可使用标准接口发起调用，降低改造过程中的适配工作量。

成本优化，减少重资产投入

AICC 采用全托管形态，安全更新、密钥轮转与远程证明由云端自动完成，您无需进行高额硬件和运维投入。

- **减少硬件投入：**无需采购专用服务器、建设机房及配套制冷供电，即可使用具备机密计算能力的推理算力。
- **减少运维投入：**安全更新、密钥轮转、远程证明等运维工作由平台承担，减少针对可信计算场景的运维投入。
- **模型更新不受本地环境限制：**使用云端模型能力，减少因本地环境资源受限导致的模型更新滞后。

应用场景

最近更新时间：2026-10-28 15:46:00

TC-AICC 可信计算（Artificial Intelligence Confidential Computing，AICC）适用于在使用大模型处理私有数据与敏感数据的场景。这类场景的共性在于：数据在推理链路中需要保持机密，推理环境需要具备可核验的技术证据，且企业不希望为此自建算力集群。

AICC 在可信执行环境（Trusted Execution Environment，TEE）内完成推理，明文仅在可信执行环境（TEE）内出现。

场景	核心需求	AICC 方案
AI Coding 场景	保护企业核心代码资产	在 TEE 内完成密态推理，代码上下文不离开安全域
RAG 托管场景	文档向量化与检索过程保持密态	Embedding、检索、推理全链路在 TEE 内加密
智能体调用场景	编排、工具调用与上下文全程加密	Agent 编排与工具调用在 TEE 内隔离执行
AI 助手协同调用场景	用户隐私数据在端侧加密后上云推理	端到端加密，远程证明，支持 BYOK（Bring Your Own Key，自带密钥）
企业 AI 办公场景	办公文档与会议内容不离开企业可控环境	透明代理接入，低改造，数据不出域

AI Coding 场景

目标用户

使用人工智能（Artificial Intelligence，AI）编程工具进行代码补全、审查与重构的研发团队，尤其是对核心代码资产有管控要求的科技、制造、医药、金融等企业。

痛点与需求

- **代码是企业核心知识产权资产：**企业内部代码库包含知识产权与商业秘密。使用云端 AI 编程工具时，代码上下文需上传云端参与推理，存在被第三方读取的风险。
- **代码审计与安全分析数据敏感：**静态应用安全测试（Static Application Security Testing，SAST）与动态应用安全测试（Dynamic Application Security Testing，DAST）的扫描结果、漏洞修复建议涉及安全缺陷信息，一旦泄露可能被用于定向攻击。
- **依赖外部模型能力但不愿数据外泄：**企业需要持续使用较新的模型能力提升代码质量，而私有化部署在成本与模型迭代速度上存在限制。
- **合规要求代码数据不离开企业可控环境：**涉密与行业合规可能要求代码资产不得离开企业可控环境。

解决方案

- AICC-SDK 集成到 IDE（集成开发环境）或编程工具客户端，代码上下文在发送前自动加密。
- 云端 GPU 在可信执行环境（TEE）内执行 LLM 推理，代码明文仅在可信执行环境内可见。
- 远程证明帮助企业验证推理环境可信，BYOK 助力企业自主掌控密钥。
- 代码补全、审查与重构结果加密返回，全链路代码不出安全域。
- 支持 CodeBuddy Security 与 AICC 联合方案，安全扫描结果也在 TEE 内处理。

RAG 托管场景

目标用户

需要将研报、合同、制度、技术文档等内部知识资产接入大模型问答的企业，以及为多个机构提供知识检索服务的平台方。

痛点与需求

- **企业知识库包含高敏感文档**：在检索增强生成（Retrieval-Augmented Generation，RAG）链路中，研报、合同、制度、技术文档等上传云端构建向量索引后，文档内容以嵌入（Embedding）形式存储在向量库中，存在被逆向还原或越权检索的风险。
- **检索权限越权是重要的安全盲区**：知识库权限与源系统权限不一致时，用户可能检索到无权访问的文档片段；该风险被开放式 Web 应用程序安全项目（Open Worldwide Application Security Project，OWASP）列入大语言模型（Large Language Model，LLM）应用十大风险；长上下文窗口可一次性带出较多敏感信息。
- **Embedding 过程存在明文暴露风险**：文档向量化需上传至 Embedding 模型，若在云端以明文处理，文档内容在过程中可能被读取。
- **检索与推理链路缺少端到端保护**：检索结果与用户提问拼接后送入大模型，链路中可能存在明文流转环节，易发生敏感数据泄露。

解决方案

- 文档 Embedding 在可信执行环境（TEE）内执行，明文文档不离开安全域。
- 向量库加密存储，检索过程在 TEE 内解密并进行相似度计算，结果不出可信执行环境。
- LLM 推理在 TEE 内完成，检索片段与用户 Query 在可信执行环境内拼接推理。
- 细粒度权限管控：在 TEE 内校验用户权限，仅返回授权范围内的检索结果。
- 通过 AICC-Proxy 透明代理接入，企业现有 RAG 架构无需修改业务代码即可升级为机密 RAG。

智能体调用场景

目标用户

将智能体（Agent）接入企业内部系统的中大型企业，涉及 CRM、ERP、数据库等多数据源的自动化调用场景。

痛点与需求

- **Agent 访问多数据源，敏感数据全程流转：**智能体编排调用 CRM、ERP、数据库等多个工具，敏感业务数据在 Agent 链路中明文流转，任一环节泄露都存在严重后果。
- **Prompt 注入攻击可诱导数据外泄：**攻击者通过外部内容注入隐藏指令，诱导 Agent 执行未授权操作或外泄敏感数据，传统防护难以检测（OWASP LLM Top1 风险）。
- **Agent 上下文包含大量敏感信息：**Agent 的记忆、规划、工具调用结果均含敏感数据，如明文存储在云端，则可能存在风险。
- **多步推理链路缺乏可信执行保障：**Agent 的多轮推理、工具选择、结果聚合缺乏硬件级隔离，存在篡改风险。

解决方案

- Agent 编排器在可信执行环境（TEE）内运行，规划、决策、记忆均在密态下处理。
- 工具调用经 AICC-Proxy 加密通道执行，敏感参数与返回结果全程加密。
- LLM 推理在 TEE 内完成，Agent 上下文（含敏感数据）的明文仅可信执行环境可见。
- 远程证明保障 Agent 执行环境不被篡改，防范恶意代码注入。
- 审计日志在 TEE 内生成并加密存储，全链路可追溯、可验证。

AI 助手协同调用场景

目标用户

在手机、PC 等终端设备内置 AI 助手，且需要将用户隐私数据上传云端推理的终端厂商。

痛点与需求

- **个人隐私数据上云推理的风险：**手机、PC 端 AI 助手需处理照片、健康数据、通讯录、聊天记录、位置信息等敏感隐私数据，上传云端推理存在被第三方读取的风险。
- **端侧算力不足，需要借助云端 GPU：**大模型推理需要高性能 GPU 算力，端侧设备算力不足以在本地运行，推理需上传云端完成。
- **品牌声誉影响较大：**一旦发生隐私数据泄露事件，会对品牌声誉造成影响，终端厂商需要向用户与监管方提供相应的技术证明。
- **需要向用户提供可核验的隐私保护证明：**仅说明数据处理方式可能难以建立用户信任，厂商希望以可验证的技术手段证明用户数据在云端不被读取。

解决方案

- AICC-SDK 集成到端侧 AI 助手，用户数据在端侧加密后上传，云端无法解密。
- 云端 GPU 在可信执行环境（TEE）内执行推理，数据明文仅在可信执行环境内存在。
- 远程证明让端侧设备验证云端推理环境可信且未被篡改。
- HSM BYOK 让设备厂商/用户掌控密钥，云平台无法访问数据。

- 支持与腾讯云大模型能力组合，用于端侧 AI 助手的机密推理场景。

企业 AI 办公场景

目标用户

在办公场景中使用云端 AI 处理合同、财报、会议纪要等文档的企业。

痛点与需求

- **办公文档包含机密商业信息**：合同、财报、战略规划、人事档案等文档使用云端 AI 处理时，内容以明文上传，存在被读取、留存与不当使用的风险。
- **会议录音与纪要涉及敏感讨论**：会议转录与纪要生成需上传语音或视频内容，涉及商业机密与参会人等信息。
- **员工使用公有云 AI 工具不易统一管控**：员工直接使用 AI 工具处理办公文档时，企业难以管控数据流向；数据一旦进入 Prompt、上下文或日志链路，使用范围不易追溯。
- **合规要求办公数据不离开企业可控环境**：涉密与行业合规可能要求办公文档与内部流程数据不得上传至外部服务器。

解决方案

- AICC-Proxy 透明代理部署在企业网络出口，办公 AI 请求自动加密转发。
- 文档处理、会议转录、邮件分析在可信执行环境（TEE）内执行，明文不离开安全域。
- LLM 推理在 TEE 内完成，办公文档内容的明文仅在可信执行环境内存在。
- 低改造接入：企业现有办公应用（OA、邮件、会议系统）无需修改代码。
- 全链路审计：操作日志在 TEE 内生成并加密存储，助力用户落实合规要求。