ICS 35.240.01
CCS L78

T/CSAC

团 体 标 准

T/CSAC 024—2025

数据中心人工智能加速芯片安全技术规范

Technical Specification for Security of Artificial Intelligence Acceleration Chips in Data Centers

2025 - 11 - 28 发布 2026- 05- 28 实施

中国网络空间安全协会 发 布

目　　次

前　言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国网络空间安全协会归口。

本文件由中国网络空间安全协会人工智能安全治理专业委员会与中国电子技术标准化研究院联合提出。

本文件起草单位：中国电子技术标准化研究院、华为技术有限公司、上海燧原科技有限公司、中科寒武纪科技股份有限公司、中国科学院信息工程研究所、北京航空航天大学、上海天数智芯半导体股份有限公司、阿里云计算有限公司、平头哥半导体有限公司、河南科技大学、北京三快科技有限公司、联通数字科技有限公司、中国移动通信集团有限公司、中移(杭州)信息技术有限公司、科大讯飞股份有限公司、三六零数字安全科技集团有限公司、北京市商汤科技开发有限公司、北京金山办公软件股份有限公司、广西电网有限责任公司、浙江蚂蚁密算科技有限公司、杭州海康威视数字技术股份有限公司、用友网络科技股份有限公司、远光软件股份有限公司。

本文件主要起草人：王健兵、郝春亮、夏文辉、吕飞霄、张妍婷、许晓耕、严敏瑞、梅敬青、王顺利、王蕊、关振宇、张志勇、张宗洋、边松、余雪松、蔡倩楠、胡铭珊、杨穷千、宋文娣、荆丽桦、张艺伯、权高原、张立尧、费凡芮、王思善、江为强、李德超、胡占锋、胡科开、卢孝新、王麟、岳龙广、梅瑞、白晓媛、谢铭、曾明霏、杜皓华、殷宇辉、张晶晶、李超、闫皓楠、季晟宇、潘登、李思璇、李瑾、杨万禄、肖青、王卫、王少康、陈勇。

数据中心人工智能加速芯片安全技术规范

1 范围

本文件规定了适用于数据中心环境的人工智能加速芯片，在硬件安全、接口安全、固件安全、安全存储单元、密码技术机制、故障检测与诊断和数据保护七个方面的安全功能要求，并给出了相应的测评方法。

本文件为适用于为数据中心环境的人工智能加速芯片的研发与应用提供技术依据，也为开展相应的安全评估和检测认证活动提供参考。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GM/T 0008—2012 安全芯片密码检测准则

3 术语和定义

下列术语和定义适用于本文件。

3.1

人工智能加速芯片 artificial intelligence accelerating chip

人工智能加速处理器 artificial intelligence accelerating chip

具备适配人工智能算法的运算微架构，能够完成人工智能应用运算处理的集成电路元件。

注：典型的人工智能加速处理器有图形处理器（GPU）、神经网络处理器（NPU）和张量处理器（TPU）。

[来源：GB/T 41867—2022，3.1.5，有修改]

3.2

逻辑接口 logic interface

能够实现数据交换功能但在物理上不存在，需要通过配置来建立的接口。

[来源：GM/T 0008—2012，3.1.19，有修改]

3.3

物理接口 physical interface

涉及各种传输介质或传输设备的接口。

[来源：GM/T 0008—2012，3.1.18]

3.4

硬件可信根 hardware root of trust

嵌入在硬件内部、受物理保护的安全模块或功能组件，作为建立计算系统信任链的初始可信基点，能够防篡改，用于安全地生成、存储和处理密钥及敏感数据，并提供可验证的信任度量功能。

3.5

侧信道攻击 side-channel attack

通过观测和分析系统运行过程中泄露的、与内部敏感操作或数据相关的物理信息（如执行时间、功耗、电磁辐射、声音、缓存访问模式等），而非直接攻击其算法或逻辑漏洞，从而推断出系统敏感信息（例如，密钥）的攻击方法。

3.6

可信执行环境 trusted execution environment

基于硬件级隔离及安全启动机制，为确保安全敏感应用相关数据和代码的机密性、完整性、真实性和不可否认性目标构建的一种运算环境。

[来源：GB/T 41388—2022，3.3，有修改]

4 概述

人工智能加速芯片的典型组成，包括加速计算核、控制单元、存储单元和互联接口等,是人工智能应用系统中的算力根基。

本文件从芯片组成及功能作用等方面，将人工智能加速芯片安全分为硬件安全、接口安全、固件安全、安全存储单元、密码技术机制、故障检测与诊断和数据保护七个方面。

具体的安全要求与测评方法映射关系见附录A。本文件中使用“应”修饰的条款是必须满足的，是纳入标准符合性判断的，使用“宜”修饰的条款是推荐满足的，不纳入标准符合性判断，但仍给出相应的测评方法，支撑测评活动，参测方可自愿参与测评。

5 安全要求

5.1 硬件安全

人工智能加速芯片硬件安全，满足如下要求：

a) 应具有唯一标识；
b) 宜基于硬件可信根生成身份证书，支撑实现接入认证等安全保护；
c) 使用受硬件保护的根密钥加解密数据或由根密钥派生密钥过程中，宜能防止通过计时攻击、能量分析攻击或电磁分析攻击等，窃取芯片的关键信息（例如，密钥）；
d) 硬件上电启动校验过程中，宜能防止通过电压、频率或温度等实施故障注入攻击，获得芯片的关键权限（例如，Boot 权限）进而窃取敏感信息。

5.2 接口安全

人工智能加速芯片接口安全，满足以下要求：

a) 应提供逻辑或物理调试接口关闭机制；
b) 不应对外提供绕过安全保护机制直接或间接访问芯片内部存储单元的物理或逻辑接口；
c) 应对调试接口、串口等接口调用提供鉴权机制，并且保护鉴权信息的机密性和完整性。

5.3 固件安全

人工智能加速芯片固件安全，满足以下要求：

a) 固件不应存在权威漏洞库（例如，CNNVD、CVE、CNVD 等）六个月前已公布的高危漏洞,并且应建立漏洞响应和恢复机制；

b）应具备固件备份和恢复机制，防止可能的篡改和损坏影响；
c）应提供固件升级校验机制，检验固件内容完整性、来源真实性等，校验失败则停止升级；
d）应提供固件防回滚机制，防止固件版本回退；
e）上电启动过程中，宜校验固件、系统引导程序等的内容完整性、来源真实性，校验不通过则停止启动或产生告警；
f）芯片运行过程中，宜对启动固件进行访问控制，防止恶意程序篡改启动固件。

5.4 安全存储单元

人工智能加速芯片应对密钥等敏感参数，提供片内安全存储单元（例如，一次性烧写，仅允许授权的进程读取等），保护密钥等敏感参数的机密性、完整性。

5.5 密码技术机制

人工智能加速芯片包含的密码功能应至少符合GM/T 0008—2012标准“安全等级1”所述要求。

5.6 故障检测与诊断

人工智能加速芯片应支持检测芯片硬件掉电、链路中断和固件损坏等故障，并产生告警或日志记录。

5.7 数据保护

人工智能加速芯片在数据保护方面，满足以下要求：
a）宜预置受硬件保护的根密钥；
b）宜支持基于受硬件保护的根密钥派生密钥，加密保护模型、数据集等敏感数据；
c）宜提供硬件可信执行环境，保护模型参数、用户数据或密钥等加载到人工智能加速芯片内存中运算处理过程中的机密性和完整性。

6 测评方法

6.1 硬件安全

6.1.1 唯一标识

人工智能加速芯片的唯一标识要求，测评方法如下：
a）测评指标：人工智能加速芯片应具有唯一标识。
b）测评步骤：检查芯片研制单位提供的芯片唯一标识举证文件，或通过操作命令查询验证芯片唯一标识。
c）预期结果：举证文件或者芯片研制单位通过命令行实操演示表明受测的人工智能加速芯片具有唯一标识。

6.1.2 硬件融合身份

人工智能加速芯片的硬件融合身份要求，测评方法如下：
a）测评指标：人工智能加速芯片宜基于硬件可信根生成身份证书，支撑实现接入认证等安全保护。
b）测评步骤：
1）向设备发起身份验证请求，能够获取设备身份证书链；
2）获取生成的设备身份证书链，校验设备身份证书和证书链。
c）预期结果：

1）执行上述测评步骤 1），能够获取设备身份证书；
2）执行上述测评步骤 2），能够验证设备身份证书的有效性。

6.1.3 防侧信道攻击

人工智能加速芯片的防侧信道攻击要求，测评方法如下：

a）测评指标：使用受硬件保护的根密钥加解密数据或由根密钥派生密钥过程中，宜能防止通过计时攻击、能量分析攻击或电磁分析攻击等，窃取芯片的关键信息（例如，密钥）。

b）测评步骤：

1）审阅芯片研制单位的举证文件，包括但不限于具备国家主管部门授予检测资质的第三方机构出具的测试报告等；

2）使用侧信道分析仪器，包括功耗（电流）、电磁辐射分析装置，芯片研制单位需提供准确的触发信号和硬件执行位置等信息，支持准确的采样时空定位，采样次数不小于 100000 次，确认是否存在密钥或敏感信息泄露。

c）预期结果：执行上述测评步骤1)-2)，明确芯片研制单位已设计、测试过防侧信道攻击相关能力，未发生密钥或敏感信息泄露。

6.1.4 防故障注入

人工智能加速芯片的防故障注入要求，测评方法如下：

a）测评指标：硬件上电启动校验过程中，宜能防止通过电压、频率或温度等实施故障注入攻击，获得芯片的关键权限（例如，Boot 权限）进而窃取敏感信息。

b）测评步骤：

1）审阅芯片研制单位的举证文件，包括但不限于具备国家主管部门授予检测资质的第三方机构出具的测试报告等；

2）使用故障注入分析仪器，包括电压、频率或温度等实施故障注入，芯片研制单位需提供准确的触发信号，支持准确的注入时间定位、注入次数不小于 10000 次，确认在最优注入条件下，是否存在安全启动绕过。

c）预期结果：执行上述测评步骤 1)-2),明确芯片研制单位已设计、测试过防故障注入相关能力，无法通过故障注入绕过安全启动机制。

6.2 接口安全

6.2.1 接口关闭

人工智能加速芯片的接口关闭要求，测评方法如下：

a）测评指标：应提供逻辑或物理调试接口关闭机制。

b）测评步骤：

1）审查芯片研制单位给出的技术方案，明确调试的串口、JTAG 口等是以何种方式关闭；

2）验证逻辑或物理调试接口关闭机制的有效性。

c）预期结果：人工智能加速芯片的逻辑或物理调试接口均已提供了有效的关闭机制。

6.2.2 无旁路接口

人工智能加速芯片的无旁路接口要求，测评方法如下：

a）测评指标：不应对外提供绕过安全保护机制直接或间接访问芯片内部存储单元的物理或逻辑接口。

b）测评步骤：

1）扫描人工智能加速芯片，观测存在的逻辑或物理接口；
2）尝试通过观测到的逻辑或物理接口，绕过接口鉴权机制，获取系统数据或日志。
c）预期结果：被测芯片中，无法通过其暴露的逻辑或物理接口在旁路鉴权机制的条件下，获取芯片内部数据。

6.2.3 接口鉴权

人工智能加速芯片的接口鉴权要求，测评方法如下：
a）测评指标：应对调试接口、串口等接口调用提供鉴权机制，并且保护鉴权信息的机密性和完整性。
b）测评步骤：
1）审查芯片研制单位给出的技术方案，明确调用芯片调试接口、串口调用是否存在鉴权机制；
2）尝试绕过鉴权机制或暴力猜测接口鉴权信息调用芯片调试接口和串口。
c）预期结果：被测芯片中的调试接口、串口调用均需经过鉴权，并且鉴权机制无法被绕过，鉴权信息有机密性和完整性保护机制。

6.3 固件安全

6.3.1 漏洞扫描

人工智能加速芯片的固件漏洞扫描要求，测评方法如下：
a）测评指标：固件不应存在权威漏洞库（例如，CNNVD、CVE、CNVD 等）六个月前已公布的高危漏洞,并且应建立漏洞响应和恢复机制；
b）测评步骤：
1）使用软件工具扫描固件，查验是否存在 CNNVD、CNVD 或 CVE 等权威漏洞库中六个月前已公布的高危漏洞；
2）审查芯片研制单位提供的漏洞响应与恢复制度文件，确认已建立漏洞响应与恢复机制。
c）预期结果：确认被测芯片固件不存在 CNNVD、CNVD 或 CVE 等权威漏洞库中六个月前已公布的高危漏洞，并且已建立漏洞响应与恢复机制。

6.3.2 备份和恢复

人工智能加速芯片的固件备份和恢复要求，测评方法如下：
a）测评指标：应具备固件备份和恢复机制，防止可能的篡改和损坏影响。
b）测评步骤：模拟固件损坏或篡改，例如，在固件升级过程中，突然断电，模拟固件损坏，观测设备能否自行恢复固件，正常启动运行。
c）预期结果：被测芯片能够自行恢复固件并正常启动运行。

6.3.3 升级校验

人工智能加速芯片的固件升级校验要求，测评方法如下：
a）测评指标：应提供固件升级校验机制，检验固件内容完整性、来源真实性等，校验失败则停止升级。
b）测评步骤：
1）对固件进行篡改后，再进行升级，验证能否成功升级；
2）查验固件发布时，是否带有数字签名与验签或其他有效机制，支撑验证固件发布来源的真实性。
c）预期结果：

1）执行测评步骤 1），结果为篡改后的固件，无法用于设备升级；
2）执行测评步骤 2），结果为固件发布时，附带有数字签名与验签或其他有效机制，能够验证固件发布来源的真实性。

6.3.4 防回滚

人工智能加速芯片的固件防回滚要求，测评方法如下：
a）测评指标：应提供固件防回滚机制，防止固件版本回退。
b）测评步骤：构建一个版本号低于当前的固件，尝试用该低版本的固件升级。
c）预期结果：低版本的固件无法在芯片中完成升级。

6.3.5 安全启动

人工智能加速芯片的安全启动要求，测评方法如下：
a）测评指标：上电启动过程中，宜校验固件、系统引导程序等的内容完整性、来源真实性，校验不通过则停止启动或产生告警。
b）测评步骤：
1）审查芯片研制单位给出的技术方案，了解芯片上电启动过程中启动校验链上涉及的组件，查看芯片启动过程中是否有对相应组件进行校验的日志记录；
2）尝试篡改人工智能加速芯片相应的固件和系统引导程序等，观测人工智能加速芯片能否正常启动；
3）查验人工智能加速芯片相应的固件和系统引导程序等，是否附带有数字签名或其他有效机制，能够用于验证其来源真实性。
c）预期结果：
1）执行测评步骤 1），结果为能够看到芯片启动过程中，记录的有对固件、系统引导程序等的完整性校验相关日志记录；
2）执行测评步骤 2），结果为篡改人工智能加速芯片相应的固件或系统引导程序后，人工智能加速芯片无法正常启动；
3）执行测评步骤 3），结果为人工智能加速芯片相应的固件和系统引导程序等，均附带有数字签名或其他有效机制，能够验证其来源真实性。

6.3.6 访问控制

人工智能加速芯片的固件访问控制要求，测评方法如下：
a）测评指标：芯片启动运行过程中，宜对启动固件进行访问控制，防范恶意程序篡改启动固件。
b）测评步骤：在芯片启动运行过程中，例如，固件缓存、解压和内存加载过程中，尝试访问、篡改和替换启动固件。
c）预期结果：无法访问、篡改或替换启动固件。

6.4 安全存储单元

人工智能加速芯片的安全存储单元要求，测评方法如下：
a）测评指标：应对密钥等敏感参数，提供片内安全存储单元（例如，一次性烧写，仅允许授权的进程读取等），保护密钥等敏感参数的机密性、完整性。
b）测评步骤：
1）审查芯片研制单位给出的技术方案，明确是否对密钥等敏感参数提供安全存储单元；
2）尝试篡改或非法访问密钥信息。

c）预期结果：被测芯片为密钥等敏感参数提供安全存储单元，且无法非授权篡改、访问其中的密钥信息。

6.5 密码技术机制

人工智能加速芯片的密码技术机制要求，测评方法如下：

a）测评指标：人工智能加速芯片包含的密码功能应至少符合 GM/T 0008—2012 标准“安全等级 1”所述要求。

b）测评步骤：检查芯片研制单位能否出示由国家密码管理局商用密码检测认证中心颁发的，符合 GM/T 0008—2012 标准“安全等级 1”或更高等级要求的认证证书。

c）预期结果：芯片研制单位能够出示由国家密码管理局商用密码检测认证中心颁发的，符合 GM/T 0008—2012 标准“安全等级 1”或更高等级要求的认证证书。

6.6 故障检测与诊断

人工智能加速芯片的故障检测与诊断要求，测评方法如下：

a）测评指标：应支持检测芯片硬件掉电、链路中断和固件损坏等故障，并产生告警或日志记录。

b）测评步骤：

1）审查芯片研制单位给出的技术方案，明确被测芯片支持检测的故障类型；

2）分别模拟芯片硬件掉电、链路中断和固件损坏故障，观测是否产生告警信息或日志记录。

c）预期结果：被测芯片支持对硬件掉电、链路中断和固件损坏等故障的检测和告警或生成日志记录。

6.7 数据保护

6.7.1 根密钥

人工智能加速芯片的根密钥要求，测评方法如下：

a）测评指标：宜预置受硬件保护的根密钥。

b）测评步骤：

1）审查芯片研制单位给出的技术方案，查看是否在芯片中预置受硬件保护的根密钥；

2）验证使用该根密钥加密数据后，将数据密文复制到其他设备上也无法解密；

3）尝试修改、导出该根密钥明文。

c）预期结果：被测芯片中预置受硬件保护的根密钥，经根密钥加密的数据无法在其他设备上解密，且无法修改、导出该根密钥。

6.7.2 模型与数据加密

人工智能加速芯片的模型与数据加密要求，测评方法如下：

a）测评指标：宜基于受硬件保护的根密钥加密保护模型、数据集等敏感数据。

b）测评步骤：验证能够调用该受硬件保护的根密钥直接或间接对模型参数或用户数据等进行加解密保护。

c）预期结果：能够调用被测芯片预置的受硬件保护的根密钥直接或间接加解密模型参数或用户数据。

6.7.3 可信执行环境

人工智能加速芯片的可信执行环境要求，测评方法如下：

a）测评指标：宜提供硬件可信执行环境，保护模型参数、用户数据或密钥等加载到人工智能加速芯片内存中运算处理过程中的机密性和完整性。

b）测评步骤：

1）审查芯片研制单位给出的技术方案，了解该芯片的可信执行环境技术原理、操作步骤等；

2）验证支持用户将模型参数、用户数据或密钥等调入可信执行环境中进行加解密保护，可信执行环境之外无法获取模型参数、用户数据或密钥明文；

3）尝试通过操作系统层、虚拟机监控器层或物理探针等攻击面，读取或篡改可信执行环境中的内存数据。

c）预期结果：人工智能加速芯片提供硬件可信执行环境，且支持用户将模型参数、用户数据或密钥调入可信执行环境保护，可信执行环境之外的进程无法读取或篡改可信执行环境中的内存数据。

附 录 A
（资料性）
安全要求与测评方法映射

本文件对人工智能加速芯片提出的安全要求和相应的测评方法总结见表 A.1 。

表 A.1 安全要求与测评方法映射表

安全要求	条款	测评方法
硬件安全	5.1 a)	6.1.1
	5.1 b)	6.1.2
	5.1 c)	6.1.3
	5.1 d)	6.1.4
接口安全	5.2 a)	6.2.1
	5.2 b)	6.2.2
	5.2 c)	6.2.3
固件安全	5.3 a)	6.3.1
	5.3 b)	6.3.2
	5.3 c)	6.3.3
	5.3 d)	6.3.4
	5.3 e)	6.3.5
	5.3 f)	6.3.6
安全存储单元	5.4	6.4
密码技术机制	5.5	6.5
故障检测与诊断	5.6	6.6
数据保护	5.7 a)	6.7.1
	5.7 b)	6.7.2
	5.7 c)	6.7.3

参 考 文 献

[1] GB/T 45958—2025 网络安全技术 人工智能计算平台安全框架

[2] GB/T 41388—2022 可信执行环境 基本安全规范

[3] GB/T 41867—2022 信息技术 人工智能 术语