

ICS
CCS

团 体 标 准

T/CIE XXX-2025

软件中试通用要求与方法

General requirements and method for software pilot testing

(征求意见稿)

2025-XX-XX 发布

2025-XX-XX 实施

中国电子学会 发布

目次

前 言	II
软件中试通用要求与方法	1
1 范围	1
2 引用文件	1
3 术语和定义	1
4 一般要求	2
4.1 中试目的	2
4.2 中试对象	2
4.3 中试环境	2
4.4 中试流程	2
4.5 中试结果记录	2
4.6 中试结论判定	3
5 中试内容	3
5.1 软件质量验证	3
5.2 适配验证	3
5.3 应用场景验证	3
5.4 技术来源验证	3
5.5 软件成熟度评估	4
6 中试方法	4
6.1 标准符合性验证	4
6.2 基准案例验证	4
6.3 性能基准测试	5
6.4 物理实测对比验证	5
6.5 竞品对比验证	5
附 录 A（资料性） 需开展软件中试的软件产品	7
A.1 具备的特征	7
A.2 产品分类	7
附 录 B（资料性） 软件中试方法选择矩阵与应用指南	8
B.1 验证方法选择矩阵	8
B.2 方法组合应用原则	8
B.3 典型场景映射表	8
B.4 实施注意事项	8

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。

本文件由中国电子学会提出并归口。

本文件主要起草单位：

本文件主要起草人：

软件中试通用要求与方法

1 范围

本文件规定了软件产品中试的一般要求、中试内容和中试方法。

本文件适用于第三方软件测评机构、开发厂商、用户及有关单位等进行的软件产品中试活动。

2 引用文件

下列文件中的有关条款通过引用而成为本文件的条款。凡注日期或版次的引用文件，其后的任何修改单(不包括勘误的内容)或修订版本都不适用于本文件，但提倡使用本文件的各方探讨使用其最新版本的可能性。凡不注日期或版次的引用文件，其最新版本适用于本文件。

GB/T 11457 信息技术 软件工程术语

GB/T 15532 计算机软件测试规范

GB/T 8567 计算机软件文档编制规范

GB/T 25000.10 系统与软件工程 系统与软件质量要求和评价（SQuaRE）系统与软件质量模型

GB/T 22239 信息安全技术—网络安全等级保护基本要求

3 术语和定义

GB/T 11457界定的以及下列术语和定义适用于本文件。

3.1

软件中试 software pilot testing

特定软件在正式投产使用前系统性验证其生产级部署能力的技术活动。

注 1：特定软件指在国民经济重要领域发挥基础性、关键性、普适性作用的软件。

注 2：生产级部署能力指能够稳定、高效、安全地投入实际生产环境运行并持续提供服务的综合能力，它不仅关注技术实现，更强调工程化、可维护性与业务可持续性。

注 3：通常在实际生产环境或高仿真环境下验证产品的部署能力与实际运行条件的适配性。

3.2

软件成熟度 software maturity level

软件产品在功能完备性、运行稳定性、生产环境适配性、可维护性等维度的综合水平。

注 1：功能完备性，包含需求覆盖度与业务场景支持能力。

注 2：运行稳定性，包含容灾、故障恢复、高并发负载下的可靠性。

注 3：生产环境适配性，包含与目标生产环境的硬件、网络及第三方系统的兼容能力。

注 4：可维护性，包含系统监控、迭代升级、配置管理的可操作性。

3.3

中试场景 pilot scenarios profile

为验证生产级部署能力设计的参数化条件集合，包括环境配置、业务操作、数据规则及异常机制等，用于中试用例设计的输入。

注 1：各要素参数（如并发用户数、故障类型）可按需调整，以适配不同验证目标。

4 一般要求

4.1 中试目的

软件中试的核心目的是通过系统性验证软件产品的生产级部署能力，缩短产品迭代周期，加速产品熟化。

4.2 中试对象

软件中试对象为通过开发阶段测试的软件产品版本。需开展软件中试的产品参见附录 A。

中试对象应具备完整的测试报告、版本基线及部署文档。若开发阶段的测试覆盖不足，应在中试前补充完成相关的测试。

中试对象采用的软件中试方法选择参见附录 B，

4.3 中试环境

软件中试环境应为软件产品的生产运行环境或高仿真环境（硬件配置、网络拓扑、数据规模等与生产环境一致）。若选择仿真环境或模拟环境，应通过技术评审并获授权批准，且须在报告中明确环境差异及风险说明。

软件中试工具一般要求是经过认可的工具，工具版本及配置参数需全程记录，确保测试结果可复现。

4.4 中试流程

软件中试流程一般包括以下活动：

- a) 中试策划：明确中试目标，制定中试计划，列出风险控制清单，并评估是否具备中试启动条件，包括软件版本、测试报告、部署文档、运行环境等；
- b) 中试设计：分析中试需求，明确验证维度和内容；制定中试技术方案，包含场景设计、工具选型、资源计划等；基于场景生成可执行的测试用例集；
- c) 中试实施：全量执行测试用例，记录原始数据；对发现的问题，按影响程度标记为阻断项、严重项、一般项；问题修复后需回归验证；
- d) 中试总结：整理和分析相关数据和结果，形成中试报告。报告中需基于验证结果量化评分给出软件成熟度评估结果，并列岀风险清单以及投产建议。

4.5 中试结果记录

软件中试结果应记录以下信息：

- a) 原始数据（含时间戳和环境参数）；
- b) 数据处理算法及工具版本；
- c) 不确定度评估报告；

- d) 差异分析结论;
- e) 验证局限说明。

4.6 中试结论判定

中试结论判定应遵循风险可控原则，按优先级排序：

- a) 一票否决项：存在阻断性问题直接判定为“不通过”；
- b) 分级通过准则：
 - 通过：严重问题 100%修复，一般问题遗留率 $\leq 5\%$ ；
 - 有条件通过：严重问题修复率 $\geq 90\%$ ，需附限期整改承诺；
 - 不通过：严重问题修复率 $< 90\%$ 或一般问题遗留率 $> 10\%$ 。
- c) 特殊判定：若需调整准则，需经相关方联合评审并签署书面协议。

5 中试内容

5.1 软件质量验证

软件质量验证应覆盖文档、程序和数据三个维度：

- a) 文档质量：
 - 验证软件产品说明、用户手册、API 文档等内容的完整性、一致性与可读性，确保其符合用户要求和 GB/T 8567 等相关标准；
 - 重点检查文档版本与软件版本的一致性，可通过工具辅助审查关键术语准确性。
- b) 程序质量：验证软件程序的功能性、性能效率、可靠性、安全性、交互能力、兼容性、灵活性、可维护性、功能安全性等是否满足设计要求和 GB/T 25000.10 等相关标准要求。
- c) 数据质量：验证基础数据的完整性、一致性、时效性，并符合数据治理规范。

5.2 适配验证

适配验证需验证软件在异构环境中的集成与运行能力，分为两类：

- a) 横向适配：验证与上下游系统的接口兼容性，包括文件格式、API 协议、数据格式等；测试高并发场景下的数据流稳定性。
- b) 垂直适配：验证与操作系统、数据库等基础软件的兼容性；验证与 CPU、GPU 等硬件的适配性。

5.3 应用场景验证

应用场景验证需基于用户业务模型构建多维度验证场景，应包括：

- a) 核心业务场景：模拟高频、高价值业务流程，验证功能完整性与性能稳定性，数据规模需与生产环境一致。
- b) 边界场景：针对极端输入、临界条件进行健壮性测试，可采用故障注入工具模拟异常。
- c) 长周期场景：如持续运行 24 小时以上，检测内存泄漏、资源碎片化等累积性缺陷，配合监控工具记录资源消耗趋势。

5.4 技术来源验证

技术来源验证需识别软件供应链潜在风险，应包括：

- a) 开源组件风险：
 - 检测开源许可证冲突、开源代码率；

- 扫描已知漏洞及过期组件版本，使用 SCA 工具生成成分分析报告；
- b) 二进制风险：验证闭源 SDK、动态库的合法性及有无恶意代码植入；
- c) 供应链攻击防护：检查构建管道完整性，防范恶意包篡改，并进行构建溯源。

5.5 软件成熟度评估

软件产品成熟度评估需基于软件质量验证、适配验证、应用场景验证和技术来源验证的结果量化评估成熟度等级。

- a) 评估模型：可采用加权评分法，按优先级分配权重，如：
 - 质量特性达标率：功能/性能/安全等测试通过率；
 - 适配性缺陷密度：每千行代码的环境适配问题数；
 - 场景覆盖率：核心、边界、长周期场景覆盖比例；
 - 供应链风险指数：高危漏洞数量+许可证违规严重性。
- b) 成熟度等级：
 - 可投产：如总分 ≥ 90 分，无高危缺陷；
 - 有条件投产：如 $80 \leq \text{总分} < 90$ ，中危缺陷 ≤ 3 项；
 - 需优化：如总分 < 80 ，或存在未修复高危缺陷。
- c) 输出物：形成《软件成熟度评估报告》，包含缺陷清单、等级结论及投产建议。

6 中试方法

6.1 标准符合性验证

6.1.1 适用对象

适用于具备明确需求规格说明且测试用例可完整设计的通用型软件产品。

6.1.2 验证要求

- a) 可依据 GB/T 15532 附录 A 规定的测试方法实施，包括但不限于：功能测试、性能测试、接口测试、容错性测试等。
- b) 若采信已有测试结果时，需满足以下条件：
 - 测试报告出具机构具有 CNAS 认可资质；
 - 被测系统名称和版本等信息一致；
 - 测试环境与被测系统运行环境具有等效性；
 - 测试结果有效期一般不超过 12 个月。

6.2 基准案例验证

6.2.1 适用对象

适用于工程分析类软件（如 CAD/CAE/CAM 系统）、科学计算软件等具有明确输入输出关系的产品。

6.2.2 验证要求

- a) 标准案例库应满足：
 - 案例覆盖核心功能的典型应用场景；
 - 输入参数应标注量纲及允许误差范围；
 - 基准结果应通过第三方专业机构评估认证。

- b) 案例执行应满足：
 - 输入文件格式应符合相关标准；
 - 输出数据应包含完整的元数据描述；
 - 结果比对应采用相对误差（如 $\leq 1\%$ ）的收敛准则。

6.3 性能基准测试

6.3.1 适用对象

适用于数据库管理系统、操作系统等基础软件的性能验证。

6.3.2 验证要求

- a) 基准数据集应满足：
 - 数据规模应符合 TPC 基准测试标准；
 - 数据分布应包含典型业务场景特征；
 - 应包含异常数据样本。
- b) 测试执行应遵循：
 - 按 SPEC 标准实施预热期；
 - 连续测试时间不少于规定时间；
 - 性能指标采集间隔应符合要求。

6.4 物理实测对比验证

6.4.1 适用对象

适用于 CAE 仿真软件、工业控制软件等与物理系统强耦合的产品。

6.4.2 验证要求

- a) 物理试验要求：
 - 试验设备应通过校准；
 - 试验环境参数应记录温度、湿度、振动等要素。
- b) 数据对比要求：
 - 建立关键指标映射关系表。

6.5 竞品对比验证

6.5.1 适用对象

适用于存在成熟商业竞品的应用软件产品。

6.5.2 验证要求

- a) 对标对象选择：
 - 市场占有率前 3 位的同类型产品；
 - 版本号差异不超过两个迭代周期；
 - 需取得合法授权使用许可。
- b) 验证维度应包括：
 - 功能覆盖度；
 - 性能吞吐量；
 - 安全漏洞等级；

- 可靠性 MTBF 指标;
- 兼容性适配清单。

附 录 A
(资料性)
需开展软件中试的软件产品

A.1 具备的特征

特征	特征说明	典型产品示例
高复杂度系统集成	软件需整合多领域技术栈（如 AI 算法+硬件控制+实时通信），存在跨模块交互的“长尾效应”问题	自动驾驶系统、SCADA 系统等
零容忍的容错要求	软件失效将导致人身安全风险或重大经济损失，需将缺陷率压至极低水平	医疗设备软件、电网调度系统等
硬件-软件深度绑定	软件性能高度依赖特定硬件环境（如算力芯片/传感器），需在真实硬件平台上验证兼容性与资源瓶颈	嵌入式系统、异构计算平台等
强实时反馈的业务闭环	软件系统需在毫秒级响应周期内完成“感知-决策-执行”闭环，且决策结果直接影响物理世界运行	工业控制软件、增强现实（AR）系统等
合规驱动的版本迭代	软件更新需通过法定认证程序，且生产部署必须与合规验证强绑定，形成“部署即审计”机制	航空电子系统、医疗设备软件等

A.2 产品分类

类别	典型产品示例
基础通用	基础软件（如操作系统、数据库）、开发工具链、虚拟化与云基础软件、网络安全软件等
行业专用	工业软件（如 CAD、CAE、DCS、SCADA）、金融软件、医疗软件、能源软件等
生态共用	云原生中间件、开发与运维协同平台、开源生态平台等

附 录 B
(资料性)
软件中试方法选择矩阵与应用指南

B.1 验证方法选择矩阵

软件分类	典型产品示例	主要验证目标	推荐验证方法	实施要求
工程分析软件	CAD/CAE/CAM 系统	算法准确性、模型收敛性	基准案例验证 (6.2) 物理实测对比验证 (6.4)	强烈推荐
基础平台软件	操作系统、数据库	性能稳定性、扩展能力	性能基准测试 (6.3) 标准符合性验证 (6.1)	强烈推荐
工业控制软件	PLC 编程系统、SCADA 系统	实时性、安全性	物理实测对比验证 (6.4) 竞品对比验证 (6.5)	推荐
科学计算软件	数值模拟工具、算法库	计算精度、并行效率	基准案例验证 (6.2) 性能基准测试 (6.3)	推荐
商业应用软件	ERP、CRM 系统	功能完整性、兼容性	竞品对比验证 (6.5) 标准符合性验证 (6.1)	推荐

B.2 方法组合应用原则

1.基础组合

- a) 所有软件产品均应至少包含标准符合性验证 (6.1) 作为基础验证;
- b) 涉及物理系统交互的软件需叠加物理实测对比验证 (6.4)。

2.增强组合

- a) 对安全性要求 \geq Level 3 (按 GB/T 22239) 的软件, 需同时采用:
标准符合性验证 (6.1) + 竞品对比验证 (6.5) + 性能基准测试 (6.3)。
- b) 工程类软件的认证级验证应采用:
基准案例验证 (6.2) \times 物理实测对比验证 (6.4) \rightarrow 交叉验证。

3.替代规则

- a) 当基准案例验证不可行时, 允许采用 3 组不同来源的基准数据集替代;
- b) 竞品对比验证中若无市场前 3 位产品, 可扩展至行业头部 5 家用户定制化案例。

B.3 典型场景映射表

验证场景	优先方法	补充方法
航空 CAE 软件适航认证	基准案例验证 (全量案例)	物理实测对比验证 (关键工况)
金融数据库性能评测	TPC-C 基准测试	72 小时压力测试
智能驾驶控制软件验证	SIL 2 级物理台架试验	对标 AUTOSAR 参考实现
工业互联网平台兼容性	竞品对比验证 (主流协议栈)	混沌工程测试

B.4 实施注意事项

1.数据一致性要求

- a) 跨方法验证时，测试数据应遵循相关数据质量标准；
- b) 物理试验与仿真数据的时间同步精度需满足相关要求；

2.工具链认证

- a) 案例库执行工具应通过第三方评估认证；
- b) 基准测试工具需提供随机性测试报告。

3.结果有效期

- a) 标准符合性报告：12 个月；
- b) 物理试验报告：24 个月（需附加环境参数稳定性证明）。

4.人员资质

- a) 验证主管需具备 ISTQB 高级认证或相当资格；
 - b) 物理试验操作员需持有相关认证资格。
-