

# 面向企业IT的高效 PC 性能测试指南

使用基准来帮助您的组织做好采用 AI 的准备



Safety. Science. Transformation.™

© 2025 UL LLC. 保留所有权利。

# 目录

前言	03	从纷繁复杂.....	10	为什么选择 Procyon?	19
人工智能的应用类型有哪些?	04	...到简单明了	12	Procyon AI 推理基准测试	20
揭秘为何“此举迫在眉睫”	05	基准测试如何降低 IT 采购预算	13	其他 Procyon 基准测试	21
什么是基准测试?	06	供应商中立的基准测试	15	立即开始测试 PC 性能	23
企业 IT 层面的基准测试	08	哪些因素会影响设备的 AI 性能?	16		
Procyon 基准测试套件简介	09	硬件	16		
		AI 推理精度	17		
		AI 推理引擎	18		





# 面向企业的人工智能

人工智能 (AI) 在企业真实场景中的应用正在迅速发展。其中一个快速增长领域是将 AI 工具部署在 PC、笔记本和手机上, 通过本地设备即可完成数据处理。此类AI服务可协助用户处理工作流自动化、概述会议内容、生成图像等任务。

当前, 大多数 AI 处理是由超大规模数据中心 (Hyperscaler) 在云端完成的, 但在云端运行计算密集型 AI 服务会产生高昂的维护成本。因此, 技术供应商正着手将较简单的 AI 任务转移到本地设备, 以降低云计算成本。为了实现这一目标, 预计支持 AI 能力的 PC 份额将从 2024 年的17% 增长到 2025 年的 43% (Statista, 2025 年)。

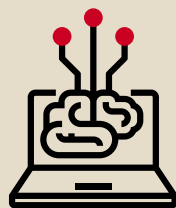
对企业而言, 本地 AI 计算的优势在于, 业务数据可保密存储于企业数据中心或直接留存设备端, 且用户离线时仍能使用 AI 功能。

# 人工智能应用类型有哪些？



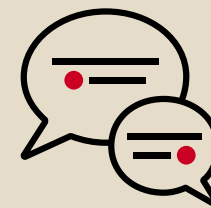
## 语音转文字转录

语音转文字转录/语音识别 (STT) 是一项经典的 PC 任务。使用 AI 推理可显著提升 STT 准确率, 并可将其卸载至专用硬件处理。



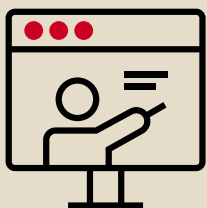
## AI 助手

AI 模型在 STT 的基础上进一步创新, 可以用于解析语音, 以识别和执行命令, 便于用户通过语音控制设备。



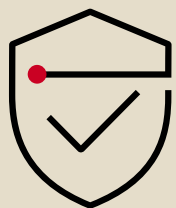
## 实时语言翻译

AI 语言翻译模型还可以处理 STT 生成的文本, 从而实现本地化的实时语言翻译。



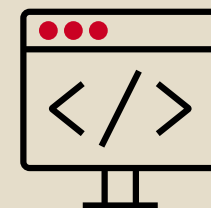
## 改进视频会议/视频处理

专注于视频处理的 AI 模型具备提升画质、降噪和背景移除等功能, 可优化视频会议体验。



## 安全应用

经过训练的 AI 模型搭载本地高效硬件, 可通过扫描识别新型恶意软件等安全威胁。



## 内容创作 (生成式 AI)

训练有素的 AI 模型能根据提示词生成文本及音视频内容。此类工作负载对计算资源需求极高, 几乎必须依赖专用硬件才能达到合格的性能。

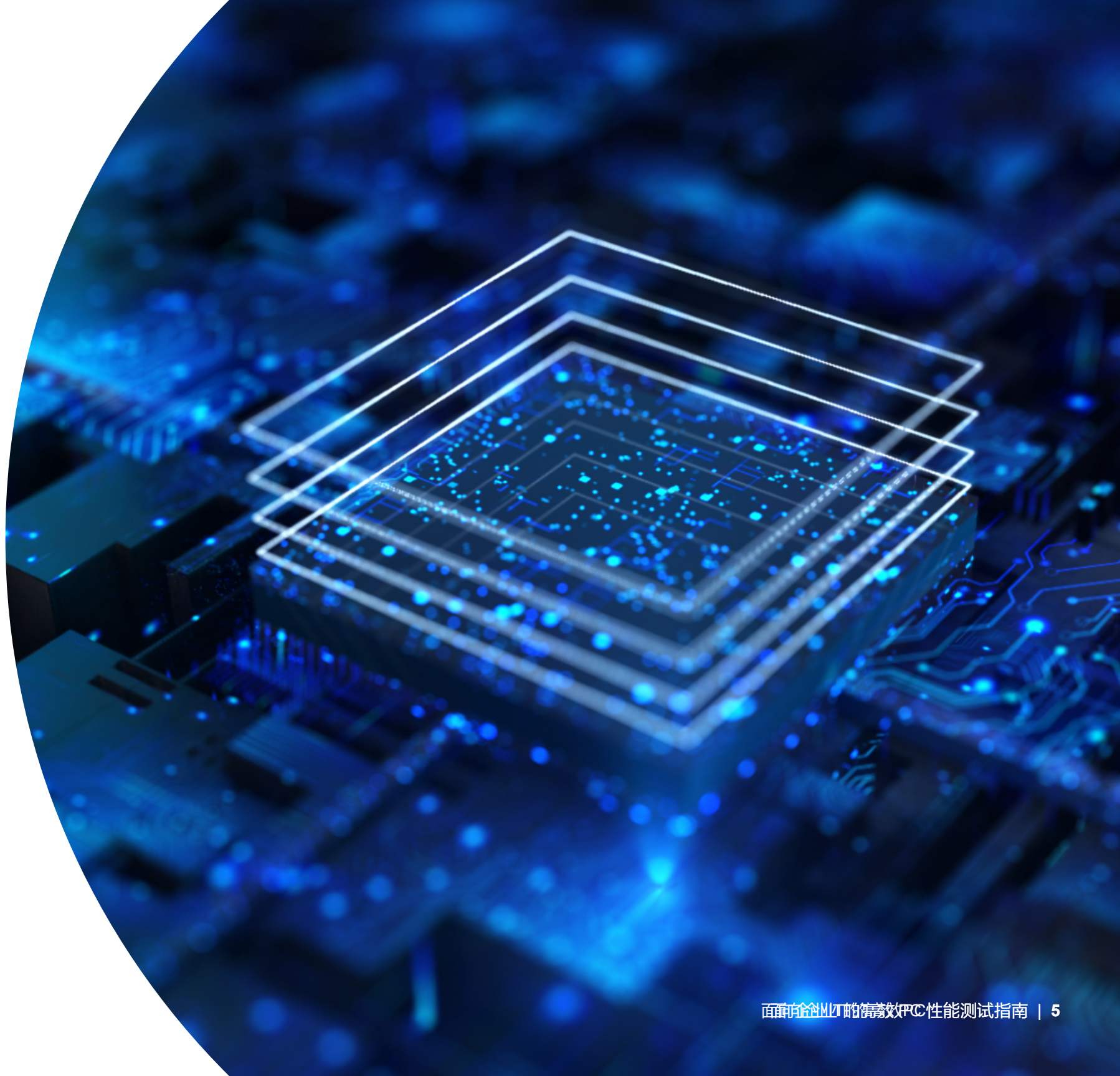


# 揭秘为何“此举迫在眉睫”

正如仅凭 GHz 无法全面衡量办公场所的计算机性能, 单一的 AI 指标 (例如广泛使用的每秒浮点运算次数 (FLOPS) ) 同样不能全面反映 AI 设备的性能。

在实际应用中, AI 处理器的真实推理性能受到多重设计因素的影响。

因此, 如要更好地评估和比较计算机系统性能, 可采用专门针对 AI 性能设计的综合性性能基准测试。





# 什么是基准测试?

基准测试用于评估产品执行特定功能的性能水平, 并与同类产品的性能进行比较, 得出相关对比数据。

该测试为 PC 性能测试提供了量化的差异化标准。计算机基准测试方案通过运行一系列明确定义的测试来衡量 PC 性能, 并对 PC 系统执行常见任务的性能进行评分: 分数越高, 性能越优。

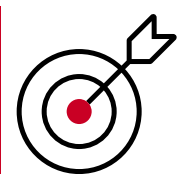
与复杂的技术规格相比, 基准测试分数有利于快速制定明智的决策以交付 PC 性能, 降低硬件成本, 节省测试时间。





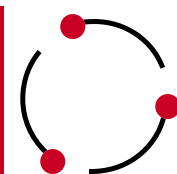


良好的基准测试有三个重要特质：



**准确性**

一致地产生真实而精确的结果。



**相关性**

测量最重要的性能要素。



**中立性**

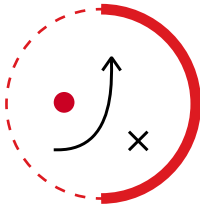
对产品或供应商不带任何偏见。





# 企业 IT 层面的基准测试

基准测试贯穿 PC 资产生命周期的每个阶段, 助力 IT 团队简化 PC 生命周期管理。



## 规划和采购

简化 PC 性能对比和成本效益分析。



## 验证和标准化

对照用户定义的基线, 测试并比较新 PC 性能。



## 运营和管理

通过高效的远程自动化测试, 提供可靠的性能洞察与报告。



## 优化或更换

基于中央数据库的基准测试结果, 制定明智的 PC 生命周期决策。



# Procyon 基准测试套件简介

UL Solutions 提供简易的 AI 设备性能测试和测量方案, 助力您洞察多元化真实场景中的 AI 性能表现:

## Procyon AI 推理基准测试

Procyon 基准测试套件在主流科技公司的参与下开发而成, 每个测试共享通用的设计与功能集, 以便提供熟悉的一致体验。每个基准测试均针对相关的实际 AI 用例而设计, 并尽可能采用广泛应用的 AI 技术。

Procyon AI 基准测试涵盖广泛的相关用例, 同时简化深度测试。测试结果提供直观易懂的分数, 以及详细的硬件性能与推理指标。



# ● 从纷繁复杂.....

● AI性能内部测试简介



## 1. 识别所有可能的用例。

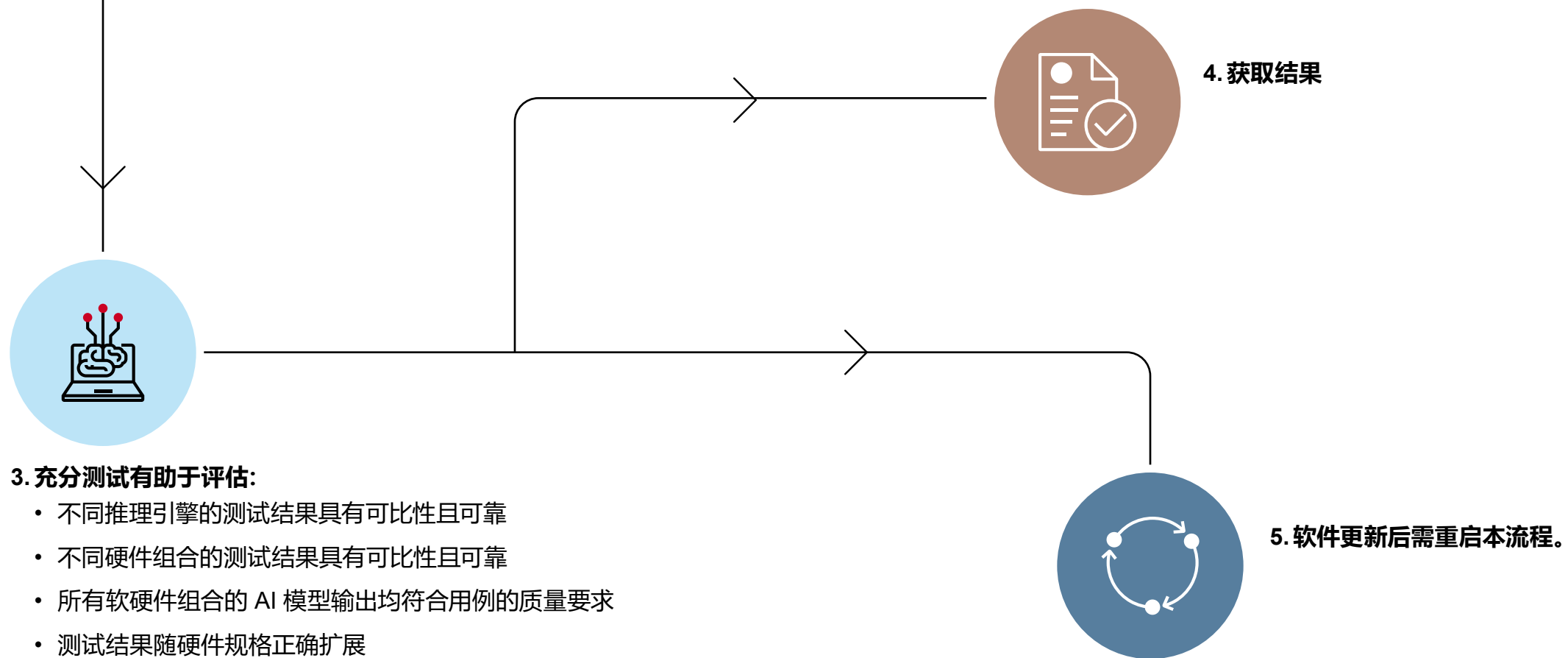
- 识别硬件
- 识别推理引擎
- 识别所需的 AI 模型
- 指定提示词和文档提示词
- 确定需要衡量哪些性能指标
- 根据输出质量要求规定量化设置



## 2. 创建环境

- 下载依赖项
- 下载 AI 模型
- 下载适用于所支持硬件的推理引擎与执行组件 (假设所有硬件都是相同的)。
- 验证硬件和软件报告的测量结果是否可靠且具有可比性。
- 验证是否所有经过测试的推理引擎均支持所选量化设置。



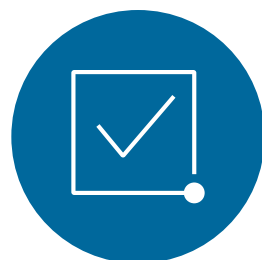
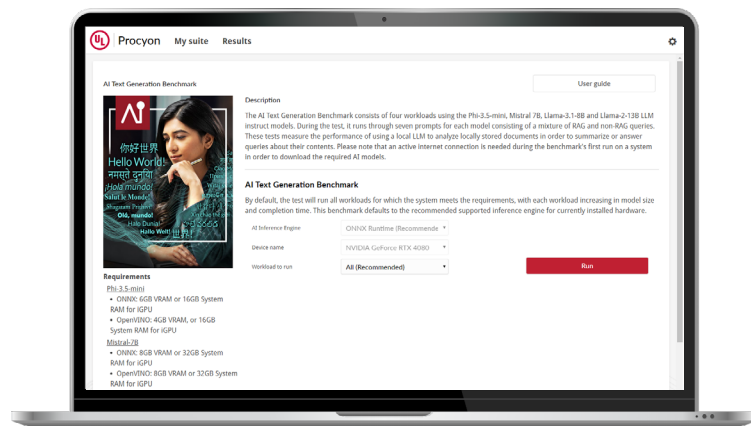


...到简单明了

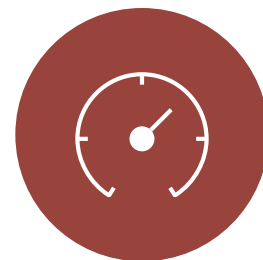
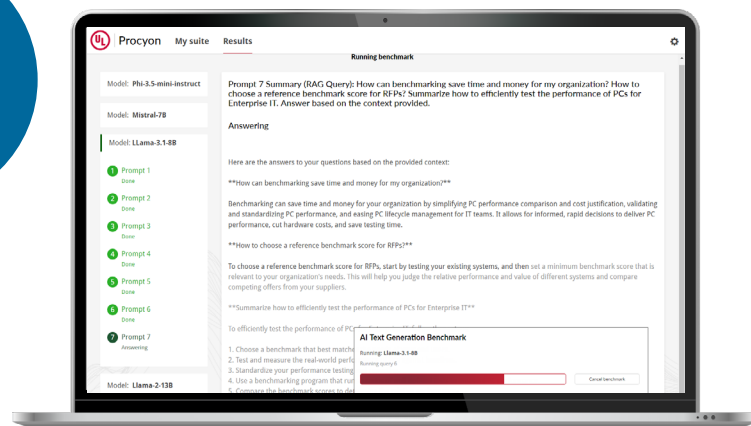
使用 Procyon 测试 AI 性能



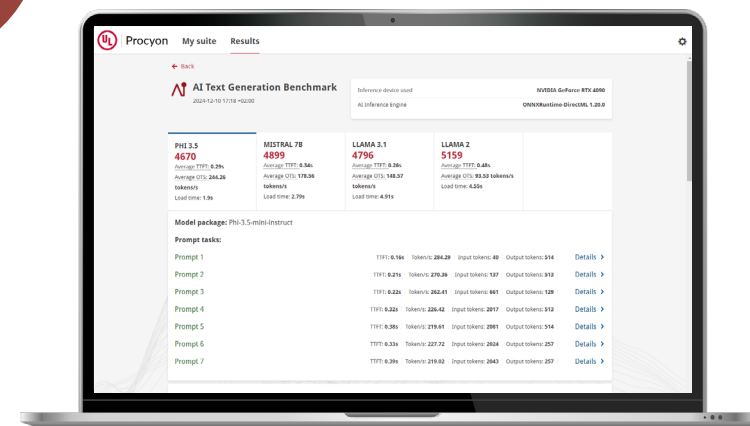
1. 点击“运行基准测试”




2. 实时查看工作完成情况



3. 获取每个用例的明确评分指标, 可选择提供额外的详细信息







# 基准测试如何降低 IT 采购预算

当前, 随着 AI 技术快速发展和不断普及, 在采购流程中借助基准测试来验证本地计算性能并进行产品对比, 已成为评估成本与性能的关键步骤。

在进行成本效益分析时, 使用基准测试可以轻松地将目标性能水平与一系列因素进行权衡, 这些因素包括维护、采用成本、扩展性, 以及一次性采购的资本支出与云端订阅的运营支出预算对比等。

当企业批量采购 PC 时, 节省的成本非常可观。通过增加 PC 招标的竞争性, 可进一步提升 IT 预算的价值。





## 公开竞争

集中向主要供应商采购虽能带来规模效益,但也可能推高成本。向更多供应商开放 RFP (招标书),就更可能获得更有竞争力的报价,进而获得协商增值服务的机会。

## 避免使用限制性语言

避免使用表达品牌偏好的语言或将供应商限制在特定解决方案的语言,以此提升招标的竞争性。采用供应商中立的 RFP,有助于供应商灵活提供更具成本效益且满足自身需求的方案。

## 压缩 IT 预算

为了获得理想的 PC 招标结果,应同时采用公开竞争的方式,使用供应商中立的语言,并规定所需的最低 PC 性能及基准参考分数。这种组合策略有助于供应商提出您可能未曾考虑过的创新解决方案。

## 规定最低性能要求

行业标准基准测试方案可助力招标方明确最低 PC 性能要求。在 RFP 中设定最低基准分数要求,能有效规避采购“低价低配”的 PC 所带来的虚假经济效益。为了更好地比较竞价方案,可要求供应商在投标文件中提供基准测试分数。这种做法以可比较的基准分数来表示 PC 性能,可有效进行成本效益分析,减少因采购过度配置的系统导致的超支风险。





# 供应商中立的基准测试

公平开放的竞争是节约采购成本的有效途径。但许多IT团队和组织难以使用实用、易懂且供应商中立的方式来描述 PC 性能。

在 RFP 中以公正的基准测试分数明确 PC 性能要求, 是实现供应商中立的简便方法。您将发现, 在采购之前, 对比性能的基准测试分数远对比组件规格更为高效。

公正的基准测试应由技术专家与专业的科技公司合作制定, 采用开放透明的流程, 确保标准公平中立。应寻求获得媒体和业界普遍信任、尊重和采用的基准测试。



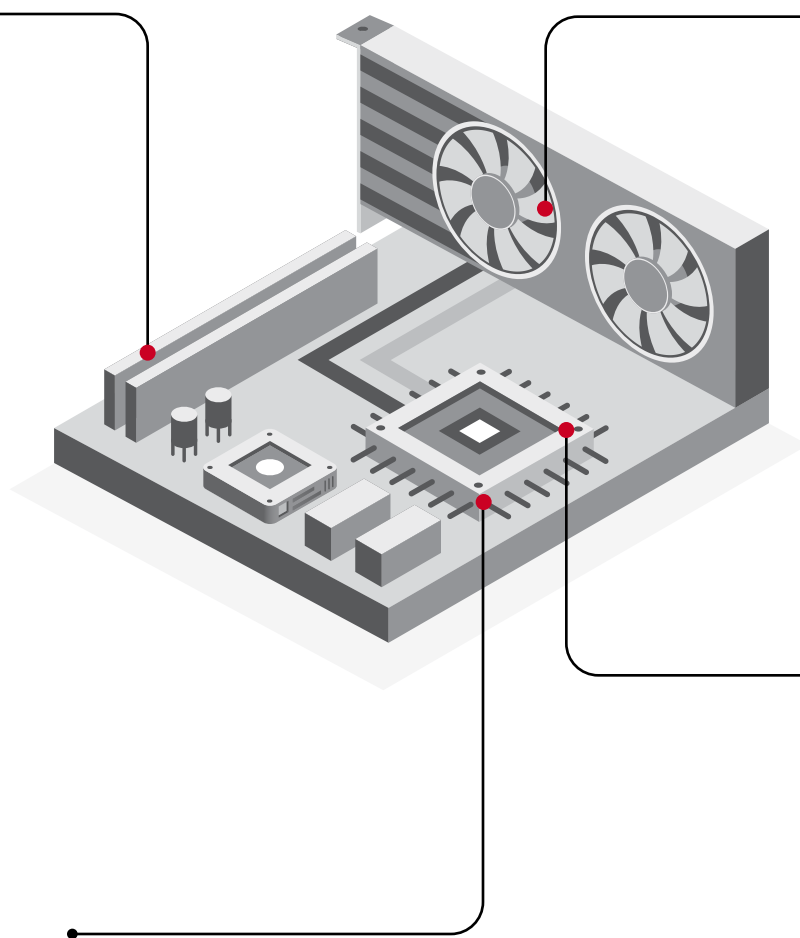
# 哪些因素会影响设备的 AI 性能?

## 内存:

部分 AI 模型规模庞大, 需依赖大量高速低延迟内存来运行。对于拥有 NPU 和集成 GPU 的系统来说, 这种内存是指系统随机存取存储器 (RAM)。对于独立 GPU 来说, 这种内存是指显存 (VRAM)。系统内存不足可能会显著影响执行速度。

## 神经处理单元 (NPU) :

越来越多的设备配备了高度专业化的 NPU, 以高效方式管理 AI 工作负载和加速 AI 任务。这种针对任务定制的芯片在执行大多数 AI 任务时, 效率远超 GPU 和 CPU 等其他硬件, 配合适当的处理器使用时, 可极大提高应用性能、散热效率与电池续航时间。



## 显卡:

尽管许多数据中心利用 GPU 的高性能和多功能性来处理 AI 任务, 但它们的能效远低于神经处理单元 (NPU)。

尽管能效较低, 但 GPU 通常比 NPU 提供更高的计算灵活性。其硬件和支持软件堆栈的专业性较低, 能够更灵活地适应更广泛的算法。

这使得 GPU 在尚未成熟的软件市场极具价值, 但由于能效较低, 开发者需谨慎使用, 尤其在便携设备上。

## 中央处理器 (CPU) :

CPU 通常与专用 NPU 和集成 GPU 共享晶圆空间, 主要用于辅助其他 AI 加速器。对于搭载 NPU 的系统, 含 CPU 的混合执行路径仍是当前部分工作负载的理想选择。



# 哪些因素会影响设备的 AI 性能?

评估AI推理性能的另一关键因素是模型所用数值的精度。现代AI推理硬件支持多种精度, 软件开发者可根据程序需求在精度与性能间权衡:

高精度格式的模型输出更准确, 但需更多内存、带宽及算力来处理。

低精度格式日益广泛 (但并非唯一) 地应用于客户端工作负载, 不仅可以提高能效, 而且能更好地适配客户端系统中的有限资源。



Stable Diffusion 1.5  
INT8 Quantization  
Intel OpenVINO



Stable Diffusion 1.5  
FP16 Quantization  
Intel OpenVINO



AI 推理引擎

# 哪些因素会影响设备的 AI 性能?

软件对 AI 推理性能的影响同等重要。人类在优化大语言模型 (LLM) 推理所使用的算法与库方面已经取得了重大进展, 其中许多进展均来自于硬件制造商。

但由于各企业需求不尽相同, 加之 AI 技术发展迅猛, IT 团队需要耗费大量的时间来定期测试最新的性能更新。

鉴于 AI 技术发展如此迅速, 延迟实施最新更新可能意味着错失关键的 AI 性能提升。





# 为什么选择 Procyon?

## 简单省钱

Procyon 可以消除妨碍您定期测量与比较设备性能的复杂痛点, 助力为您节省时间和成本。

## 深度洞察

查看子项测试分数和硬件监控数据, 进一步了解设备在各环节的性能表现优劣。

## 保存和对比分数

在 Procyon 应用中对比分数并保存, 用于追踪性能随时间推移或系统/软件更新而产生的变化。

## 结合行业专业知识开发

每个 Procyon 基准测试均与行业内专业的硬件公司合作开发, 力保基准测试对所支持的硬件保持中立性。

## 快速轻松测试多个真实场景

Procyon 基准测试以真实用例为设计基础。每个 Procyon AI 推理基准测试都会在真实场景中进行用例测试, 覆盖多种常见的实现方案。

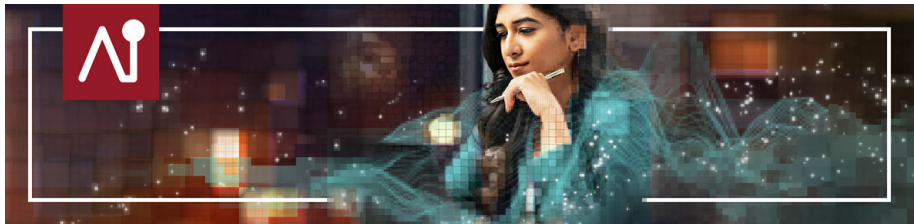
## 持续更新的基准测试

Procyon 基准测试会定期更新, 以支持最新硬件与软件功能, 确保基准测试与您的系统硬件相关。

## 全球专家广泛应用

全球数千家先进的行业、媒体和企业都在采用我们的基准测试, 这些独立标准化工具可助力他们理解与对比设备性能。

# Procyon AI 推理基准测试



## AI 图像生成基准测试

该测试基于当前十分流行的AI图像生成模型 Stable Diffusion, 用于评估本地 AI 加速器的推理性能。多种量化设置支持测试来自 AI 能力强大的计算机硬件行业专家的各种计算机硬件。



## AI 计算机视觉基准测试

该基准测试结合多种流行的计算机视觉模型, 可对比不同厂商的 AI 推理引擎在 Windows PC 和 macOS 硬件上的表现, 对设备端推理操作的性能进行评分。



## AI 文本生成基准测试

该基准测试旨在尽可能减少常见 AI 大语言模型的工作负载变量, 可以更紧凑、更简便地测试多种大语言 AI 模型的文本生成性能, 保证测试过程的可重复性和一致性。



# 其他 Procyon 基准测试

Procyon 提供灵活的许可方式, 支持按需选择单个基准测试模块, 无论是用于专项 AI 性能测量, 还是办公生产率、创造力或电池续航时间等广泛的计算机性能洞察。



## 视频编辑基准测试

该基准测试使用 Adobe Premiere Pro 执行典型的视频编辑工作流程, 助力创作者简便地比较不同系统和规格的性能。



## 电池续航基准测试

比较 Windows 膝上型电脑、笔记本电脑和平板电脑的实际电池续航时间。该基准测试生成电池续航时间概况, 充分展示了不同场景 (包括播放视频、待机以及用于办公) 下的电池续航时间。



## 照片编辑基准测试

该基准测试使用 Adobe Lightroom Classic 和 Adobe Photoshop 来执行典型的照片编辑工作流程 (包括批量处理和图像修饰), 进而衡量 Windows PC 和 Apple Mac 的计算机性能。

# 其他 Procyon 基准测试



## 办公生产率基准测试

使用 Microsoft office 应用程序来测量 Windows PC 和 Apple Mac 在办公生产率工作方面的表现。该基准测试的工作负载基于 Microsoft Word、Excel、PowerPoint 和 Outlook 的相关、真实任务。



## 每小时电池耗电量基准测试

现代设备有许多影响功耗的设置，使用完整的电池续航时间基准测试来逐一测试这些设置会消耗大量时间。该基准测试有助于快速简便地了解驱动程序和操作系统更新、省电设置以及企业设备管理软件对设备电池耗电量的影响。





# 立即开始测试 PC 性能

全球各地的IT专业人士均可以选择 UL Solutions 基准测试, 以精准测量、理解并管理计算机硬件性能。我们提供值得信赖且广泛采用的性能测试方案, 助力 IT 专业人士进行更明智的决策。联系我们, 了解关于自动化性能测试的更多信息: UL Solutions 基准测试服务助力节省时间与成本。

请通过 [Benchmarks.UL.com/Contact](https://Benchmarks.UL.com/Contact) **与我们联系**



**[Benchmarks.UL.com/Procyon](https://Benchmarks.UL.com/Procyon)**

© 2025 UL LLC. 保留所有权利。

RLC25CS2045218\_zhCN