

## 借助英特尔® 高级矩阵扩展 (英特尔® AMX) 加速人工智能 (AI) 工作负载

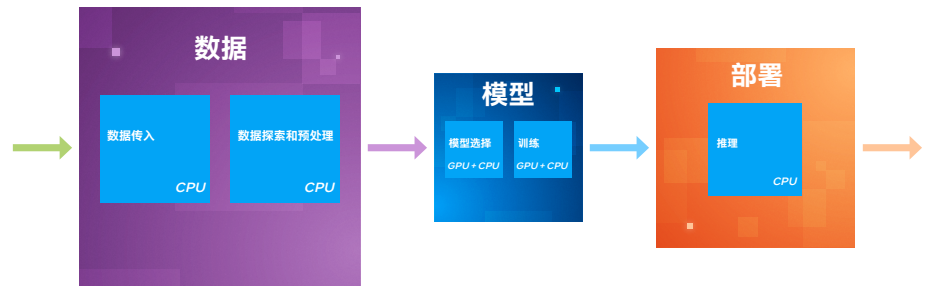
英特尔推出的第四代英特尔® 至强® 可扩展处理器及其内置的英特尔® 高级矩阵扩展 (Intel® Advanced Matrix Extensions, 英特尔® AMX) 可进一步提高 AI 功能, 实现较上一代产品 3 至 10 倍的推理和训练性能提升<sup>1</sup>。



### 优化 AI 流水线

从图书和电影推荐系统到驱动大型电商网站的数字化零售软件, 再到支持聊天机器人和机器翻译的自然语言处理 (NLP) 功能, 企业可通过在不同场景中应用人工智能 (AI) 获得诸多收益。AI 真正的价值在于其解析复杂环境和海量数据的特性, 以及解决以往难解之题的能力, 而这正是不断推进企业革新的关键潜能。据研究显示, 到 2025 年, 90% 新发布的企业应用版本都将包含嵌入式 AI 功能<sup>2</sup>。

### AI 流水线



3 个外方框表示 AI 流水线各阶段。  
5 个内方框表示 AI 工作负载。  
方框大小表明 AI 流水线中处理器活动的相对水平。

图 1. AI 流水线中的 AI 工作负载和处理器活动

第四代英特尔® 至强® 可扩展处理器内置 AI 加速器——英特尔® AMX, 是企业和机构优化 AI 流水线的理想选择。平衡推理是 CPU 在 AI 应用中的主要用例, 英特尔® AMX 专为该用例设计并且具备更多训练能力 (见图 1)<sup>3</sup>。目前, 在所有运行 AI 推理工作负载的已装机数据中心处理单元中, 英特尔® 至强® 可扩展处理器的占比高达 70%; 因此, 为新的 AI 部署选择内置英特尔® AMX 的第四代英特尔® 至强® 可扩展处理器, 是一种既高效又具有成本效益的 AI 工作负载加速方式<sup>4</sup>。

## 选择内置加速器的理由

目前，采用内置英特尔® 深度学习加速技术 (Intel® Deep Learning Boost, 英特尔® DL Boost) 的第三代英特尔® 至强® 可扩展处理器支持 AI 部署，可满足 IT 部门履行客户服务级别协议 (SLA) 的要求，而内置英特尔® AMX 的第四代英特尔® 至强® 可扩展处理器则将带来新的变革。

图 2 所示为英特尔® AMX 在代际间实现高达 5.7 至 10 倍的 PyTorch 实时推理性能提升的情况；图 3 所示为英特尔® AMX 在代际间实现高达 3.5 至 10 倍的 PyTorch 训练性能提升的情况<sup>5</sup>。凭借更强性能，英特尔® AMX 将进一步提升客户满意度。英特尔® AMX 加速器内置于企业和机构业已熟悉的 CPU 解决方案之中，可显著提升多方面性能，让您在选择适合 AI 应用的 CPU 时更轻松。

### 第四代英特尔® 至强® 可扩展处理器内置英特尔® AMX，实现高达 5.7 至 10 倍的代际实时推理性能提升 (越高越好)

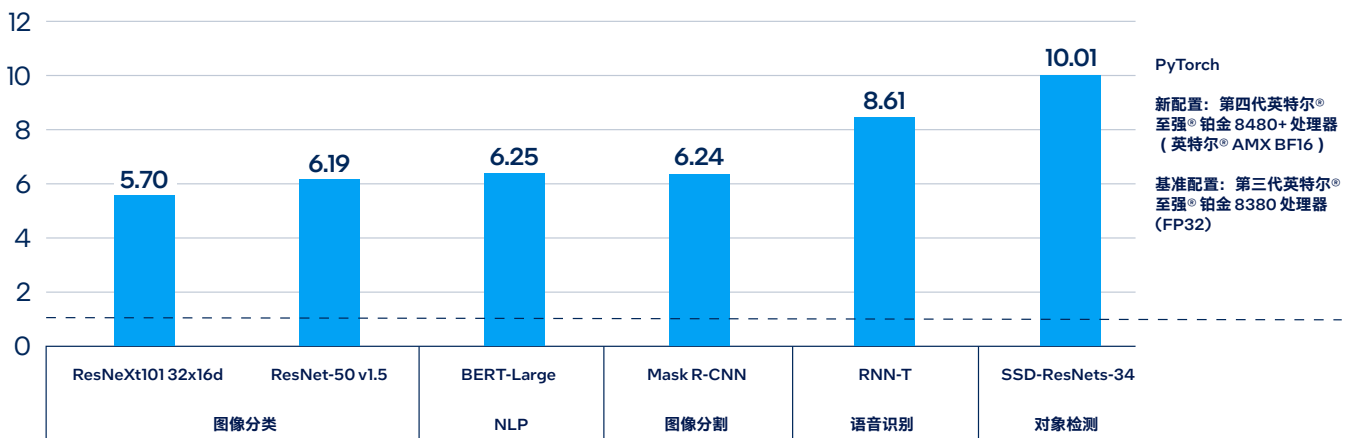


图 2. 第四代英特尔® 至强® 可扩展处理器内置英特尔® AMX，加速 PyTorch 实时推理<sup>5</sup>

### 第四代英特尔® 至强® 可扩展处理器内置英特尔® AMX，实现高达 3.5 至 10 倍的代际训练性能提升 (越高越好)

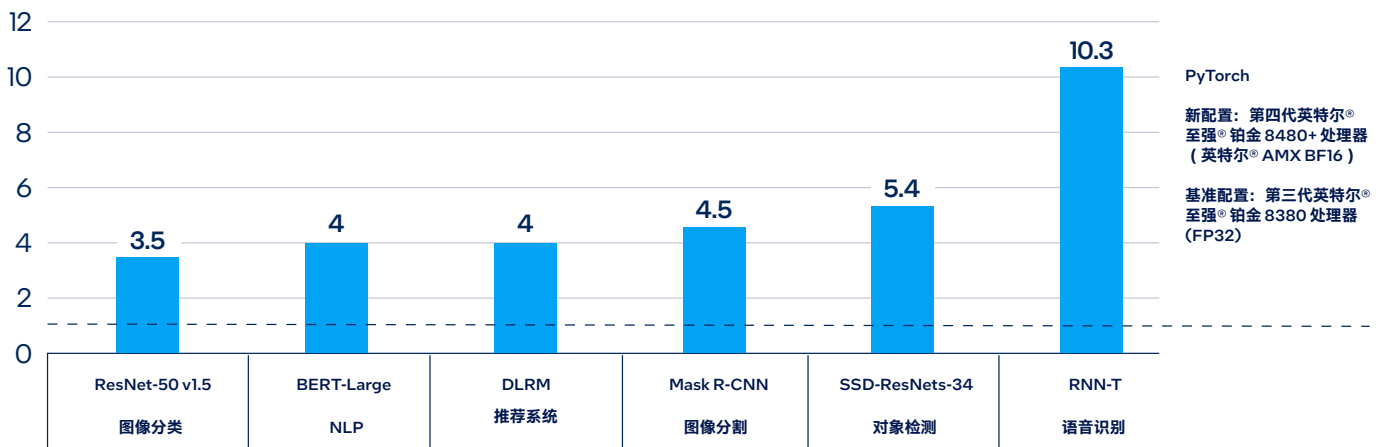


图 3. 第四代英特尔® 至强® 可扩展处理器内置英特尔® AMX，加速 PyTorch 训练<sup>5</sup>

通过图 4 可以看出英特尔® AMX 带来的性能提升远大于每一代产品 (从第一代英特尔® 至强® 可扩展处理器开始) 通过增加内核所实现的性能提升。

## 摩尔定律与加速器

为工作负载匹配合适的计算引擎

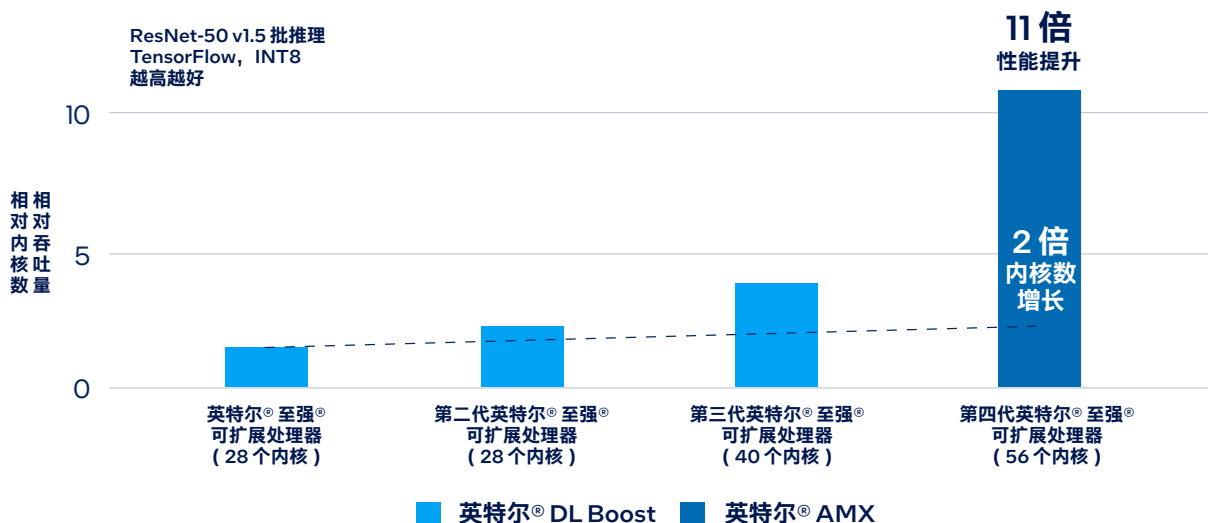


图 4. 以第一代英特尔®至强®可扩展处理器为基准，与前几代产品相比，英特尔®AMX 实现非线性性能提升。

### 英特尔®AMX 是什么？

英特尔®AMX 是内置于第四代英特尔®至强®可扩展处理器中的加速器，可优化深度学习 (DL) 训练和推理工作负载。借助英特尔®AMX，第四代英特尔®至强®可扩展处理器可在优化通用计算和 AI 工作负载间快速转换。设想一辆汽车，既有非常出色的城市道路行驶表现，也可以迅速转换模式，展现 F1 赛事级别的竞速性能。第四代英特尔®至强®可扩展处理器就具备这种灵活性。开发人员可以编写非 AI 功能代码来利用处理器的指令集架构 (ISA)，也可编写 AI 功能代码，以充分发挥英特尔®AMX 指令集的优势。英特尔已将其 oneAPI DL 引擎——英特尔®oneAPI 深度神经网络库 (Intel®oneAPI Deep Neural Network Library, 英特尔®oneDNN) 集成至包括 TensorFlow、PyTorch、PaddlePaddle 和 ONNX 在内的多个主流 AI 应用开源工具当中。

### 英特尔®AMX 架构

英特尔®AMX 架构由两部分组件构成 (见图 5)：

- 第一部分为 TILE，由 8 个 1KB 大小的 2D 寄存器组成，可存储大数据块。
- 第二部分为平铺矩阵乘法 (TMUL)，它是与 TILE 连接的加速引擎，可执行用于 AI 的矩阵乘法计算。



图 5. 英特尔® AMX 架构由 2D 寄存器文件 (TILE) 和 TMUL 组成

英特尔® AMX 支持两种数据类型：INT8 和 BF16，两者均可用于 AI 工作负载所需的矩阵乘法运算。

- 当推理无需 FP32 (AI 经常使用的单精度浮点格式) 的精度时可使用 INT8 这种数据类型。由于该数据类型的精度较低，因此单位计算周期内运算次数就更多。
- BF16 这种数据类型实现的准确度足以达到大多数训练的要求，必要时它也能让 AI 推理实现更高的准确度。

凭借这种新的平铺架构，英特尔® AMX 实现了大幅代际性能提升。与运行英特尔® 高级矢量扩展 512 神经网络指令 (Intel® Advanced Vector Extensions 512 Neural Network Instructions, 英特尔® AVX-512 VNNI) 的第三代英特尔® 至强® 可扩展处理器相比，运行英特尔® AMX 的第四代英特尔® 至强® 可扩展处理器将单位计算周期内执行 INT8 运算的次数从 256 次提高至 2048 次。此外，如图 6 所示，第四代英特尔® 至强® 可扩展处理器可在单位计算周期内执行 1024 次 BF16 运算，而第三代英特尔® 至强® 可扩展处理器执行 FP32 运算的次数仅为 64 次<sup>7</sup>。

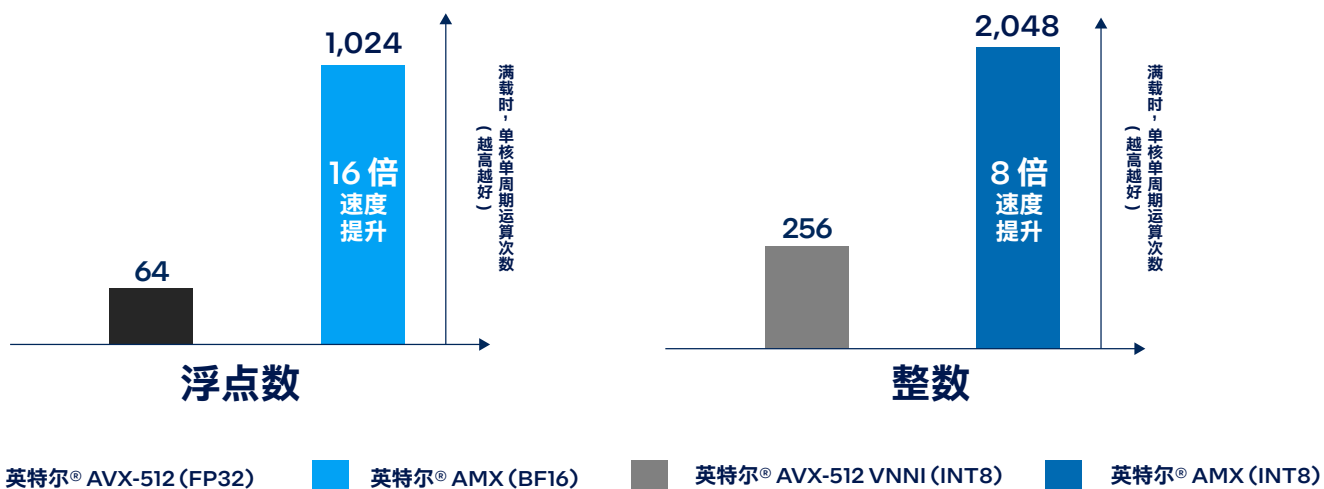


图 6. 与英特尔® AVX-512 VNNI 相比，英特尔® AMX 在处理 INT8 和 BF16 两种数据类型时表现更出色<sup>7</sup>。

## AI 用例

内置英特尔® AMX 的第四代英特尔® 至强® 可扩展处理器可用于广泛的深度学习 (DL) 用例。



### 推荐系统

无论是推荐电影和图书，还是进行广告精准投放，均可带来个性化的终端用户体验。基于深度学习 (DL) 创建可将用户实时行为，以及近乎实时的时间和地点等相关场景特征考虑在内的推荐系统。



### 自然语言处理 (NLP)

语言推理、机器学习 (ML) 等自然语言处理 (NLP) 应用是企业支持和扩展情绪分析、聊天机器人和机器翻译等多种功能的关键所在，预计到 2029 年，自然语言处理应用的全球市场规模将达到 1,618.1 亿美元<sup>8</sup>。



### 零售电商软件解决方案

在 PyTorch、TensorFlow 等已面向 AI 优化的框架基础上，通过使用深度学习 (DL) 推理和训练，大幅缩短交易时间，满足峰值需求，助力零售企业增加收入并提供出色的客户体验。

## 使用英特尔® AMX 立启新加速

借助英特尔® AMX，几乎无需费力，即可提升性能。这得益于多个默认使用的框架都经过英特尔® oneDNN 的优化。Windows 和 Linux 操作系统、基于内核的虚拟机 (KVM) 和多个主流虚拟机管理程序都支持英特尔® AMX 指令集。INT8 和 BF16 运算在 TensorFlow 和 PyTorch 等开源框架内可自动优化。开发人员可借助英特尔® 分发版 OpenVINO™ 工具包 (Intel® Distribution of OpenVINO™ toolkit) 实现 AI 推理的自动化、优化、微调和运行，且几乎或者完全不需要具备编码知识。而且，开发人员只需使用英特尔® Neural Compressor 便可将训练模型量化为 INT8 数据类型。

## 使用第四代英特尔® 至强® 可扩展处理器为 AI 加速

使用内置英特尔® AMX 的第四代英特尔® 至强® 可扩展处理器，充分挖掘 AI 无限潜能，更好助力千行百业。在数据中心已广泛部署英特尔® 至强® 可扩展处理器的基础之上更进一步，借助全新矩阵乘法运算加速方案，体验更加出色的 AI 训练和推理性能。

有关英特尔® 人工智能和英特尔® AMX 的更多信息，请访问 [intel.cn/content/www/cn/zh/artificial-intelligence/overview.html](https://intel.cn/content/www/cn/zh/artificial-intelligence/overview.html)。



- 1 详情请见以下网址的 [A16, A17]: <https://edc.intel.com/content/www/cn/zh/products/performance/benchmarks/processors/> (第四代英特尔®至强®可扩展处理器)。结果可能不同。
- 2 Forbes, "Top Artificial Intelligence (AI) Predictions For 2020 From IDC and Forrester"[IDC 和 Forrester 有关 2020 年人工智能 (AI) 领域的重大预测], 2019 年 11 月, [forbes.com/sites/gilpress/2019/11/22/top-artificial-intelligence-ai-predictions-for-2020-from-idc-and-forrester/#4fef9821315a](https://forbes.com/sites/gilpress/2019/11/22/top-artificial-intelligence-ai-predictions-for-2020-from-idc-and-forrester/#4fef9821315a).
- 3 The Next Platform, "With AMX, Intel Adds AI/ML Sparkle to Sapphire Rapids" (英特尔凭借 AMX 使第四代英特尔®至强®可扩展处理器具备出色的 AI/ML 性能), 2021 年 8 月, [nextplatform.com/2021/08/19/with-amx-intel-adds-ai-ml-sparkle-to-sapphire-rapids/](https://nextplatform.com/2021/08/19/with-amx-intel-adds-ai-ml-sparkle-to-sapphire-rapids/).
- 4 基于英特尔对截至 2021 年 12 月运行 AI 推理工作负载的全球数据中心服务器装机容量的市场建模。
- 5 PyTorch 模型性能配置。PT-NLPBERT-Large: 8480: 单节点, 配备 2 个英特尔®至强®铂金 8480 处理器的非量产平台, 总内存 1,024 GB (16 插槽/64 GB/DDR5-4800), ucode 0x2b0000a1, 启用英特尔®超线程技术 (Intel® Hyper-Threading Technology, 英特尔®HT 技术), 启用英特尔®睿频加速技术, CentOS Stream 8, 5.15.0, 1 个 1 TB 英特尔®固态硬盘 SC2KW256G8 (PT)/Samsung 固态硬盘 860 EVO (TF), BERT-Large, 推理: SQuAD1.1 (序列长度=384), bs=1 (4 核/实例), bs=n (单路/实例), bs: FP32=1,56, 英特尔®AMX BF16=1,16, 英特尔®AMX INT8=1,56, Trg: 维基百科 2020 年 1 月 1 日 (序列长度=512), bs: FP32=28, 英特尔®AMX BF16=56 (单实例, 单路), 框架: <https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66>; ModelZoo: <https://github.com/IntelAI/models/tree/spr-launch-public>, PT: 1.13, IPEX: 1.13, oneDNN: v2.7, 基于英特尔 2022 年 10 月 24 日进行的测试。8380: 单节点, 2 个英特尔®至强®铂金 8380 处理器, 总内存 1,024 GB (16 插槽/64 GB/DDR4-3200), ucode 0xd000375, 启用英特尔®HT 技术, 启用英特尔®睿频加速技术, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1 个英特尔®固态硬盘 SC2KG960G8, BERT-Large, 推理: SQuAD1.1 (序列长度=384), bs=1 (4 核/实例), bs=n (单路/实例), bs: FP32=128, 英特尔®AMX BF16=128, 英特尔®AMX INT8=128, 训练: bs: fp32/英特尔®AMX BF16=32K (单实例, 单路), Criteo TB 数据集, 框架: <https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66>; ModelZoo: <https://github.com/IntelAI/models/tree/spr-launch-public>, PT: 1.13, IPEX: 1.13, oneDNN: v2.7, 基于英特尔 2022 年 10 月 24 日进行的测试。PT-DRLM: 8480: 单节点, 配备 2 个英特尔®至强®铂金 8480 处理器的非量产平台, 总内存 1,024 GB (16 插槽/64 GB/DDR5-4800), ucode 0x2b0000a1, 启用英特尔®HT 技术, 启用英特尔®睿频加速技术, CentOS Stream 8, 5.15.0, 1 个 1 TB 英特尔®固态硬盘 SC2KW256G8 (PT)/Samsung 固态硬盘 860 EVO (TF), DLRM, 推理: bs=n (单路/实例), bs: FP32=128, 英特尔®AMX BF16=128, 英特尔®AMX INT8=128, 训练: bs: fp32/英特尔®AMX BF16=32K (单实例, 单路), Criteo TB 数据集, 框架: <https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66>; ModelZoo: <https://github.com/IntelAI/models/tree/spr-launch-public>, PT: 1.13, IPEX: 1.13, oneDNN: v2.7, 基于英特尔 2022 年 10 月 24 日进行的测试。PT-ResNets-34: 8480: 单节点, 配备 2 个英特尔®至强®铂金 8480 处理器的非量产平台, 总内存 1,024 GB (16 插槽/64 GB/DDR5-4800), ucode 0x2b0000a1, 启用英特尔®HT 技术, 启用英特尔®睿频加速技术, CentOS Stream 8, 5.15.0, 1 个 1 TB 英特尔®固态硬盘 SC2KW256G8 (PT)/Samsung 固态硬盘 860 EVO (TF), SSD-ResNet-34, 推理: bs=1 (4 核/实例), bs=n (单路/实例), bs: FP32=1,112, 英特尔®AMX BF16=1,112, 英特尔®AMX INT8=1,112, 训练: bs: FP32/英特尔®AMX BF16=224 (单实例, 单路), COCO 2017, 框架: <https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66>; ModelZoo: <https://github.com/IntelAI/models/tree/spr-launch-public>, PT: 1.13, IPEX: 1.13, oneDNN: v2.7, 基于英特尔 2022 年 10 月 24 日进行的测试。8380: 单节点, 2 个英特尔®至强®铂金 8380 处理器, 总内存 1,024 GB (16 插槽/64 GB/DDR4-3200), ucode 0xd000375, 启用英特尔®HT 技术, 启用英特尔®睿频加速技术, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1 个英特尔®固态硬盘 SC2KG960G8, ResNet-50 v1.5, 推理: bs=1 (4 核/实例), bs=n (单路/实例), bs: FP32=1,64, 英特尔®AMX BF16=1,64, 英特尔®AMX INT8=1,116, 训练: bs: FP32=128, 英特尔®AMX BF16=128 (单实例, 单路), ImageNet (224 x 224), 框架: <https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66>; ModelZoo: <https://github.com/IntelAI/models/tree/spr-launch-public>, PT: 1.13, IPEX: 1.13, oneDNN: v2.7, 基于英特尔 2022 年 10 月 24 日进行的测试。PT-RNN-T: 8480: 单节点, 配备 2 个英特尔®至强®铂金 8480 处理器的非量产平台, 总内存 1,024 GB (16 插槽/64 GB/DDR5-4800), ucode 0x2b0000a1, 启用英特尔®HT 技术, 启用英特尔®睿频加速技术, CentOS Stream 8, 5.15.0, 1 个 1 TB 英特尔®固态硬盘 SC2KW256G8 (PT)/Samsung 固态硬盘 860 EVO (TF), Resnext101 32x16d, 推理: bs=1 (4 核/实例), bs=n (单路/实例), bs: FP32=1,64, 英特尔®AMX BF16=1,64, 英特尔®AMX INT8=1,116, ImageNet, 框架: <https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66>; ModelZoo: <https://github.com/IntelAI/models/tree/spr-launch-public>, PT: 1.13, IPEX: 1.13, oneDNN: v2.7, 基于英特尔 2022 年 10 月 24 日进行的测试。PT-ResNext101: 8480: 单节点, 配备 2 个英特尔®至强®铂金 8480 处理器的非量产平台, 总内存 1,024 GB (16 插槽/64 GB/DDR5-4800), ucode 0xd000375, 启用英特尔®HT 技术, 启用英特尔®睿频加速技术, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1 个英特尔®固态硬盘 SC2KG960G8, Resnext101 32x16d, 推理: bs=1 (4 核/实例), bs=n (单路/实例), bs: FP32=1,64, 英特尔®AMX BF16=1,64, 英特尔®AMX INT8=1,116, ImageNet, 框架: <https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66>; ModelZoo: <https://github.com/IntelAI/models/tree/spr-launch-public>, PT: 1.13, IPEX: 1.13, oneDNN: v2.7, 基于英特尔 2022 年 10 月 24 日进行的测试。PT-MaskRCNN: 8480: 单节点, 配备 2 个英特尔®至强®铂金 8480 处理器的非量产平台, 总内存 1,024 GB (16 插槽/64 GB/DDR5-4800), ucode 0x2b0000a1, 启用英特尔®HT 技术, 启用英特尔®睿频加速技术, CentOS Stream 8, 5.15.0, 1 个 1 TB 英特尔®固态硬盘 SC2KW256G8 (PT)/Samsung 固态硬盘 860 EVO (TF), MaskRCNN, 推理: bs=1 (4 核/实例), bs=n (单路/实例), bs: FP32=1,112, 英特尔®AMX BF16=1,112, 训练: bs: FP32/英特尔®AMX BF16=112 (单实例, 单路), COCO 2017, 框架: <https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66>; ModelZoo: <https://github.com/IntelAI/models/tree/spr-launch-public>, PT: 1.13, IPEX: 1.13, oneDNN: v2.7, 基于英特尔 2022 年 10 月 24 日进行的测试。8380: 单节点, 2 个英特尔®至强®铂金 8380 处理器, 总内存 1,024 GB (16 插槽/64 GB/DDR4-3200), ucode 0xd000375, 启用英特尔®HT 技术, 启用英特尔®睿频加速技术, Ubuntu 22.04 LTS, 5.15.0-27-generic, 1 个英特尔®固态硬盘 SC2KG960G8, Resnet101 32x16d, bs=n (单路/实例), 推理: bs: FP32=1,64, INT8=1,116, 训练: bs: FP32=128 (单实例, 单路), COCO 2017, 框架: <https://github.com/intel-innersource/frameworks.ai.pytorch.private-cpu/tree/d7607bdd983093396a70713344828a989b766a66>; ModelZoo: <https://github.com/IntelAI/models/tree/spr-launch-public>, PT: 1.13, IPEX: 1.13, oneDNN: v2.7, 基于英特尔 2022 年 10 月 24 日进行的测试。推理: ResNet-50 v1.5: ImageNet (224 x 224), SSD ResNet-34: COCO 2017 (1200 x 1200), BERT-Large: SQuAD1.1 (序列长度=384), Resnext101: ImageNet, MaskRCNN: COCO 2017, DLRM: Criteo TB 数据集, RNN: LibriSpeech, 训练: ResNet-50 v1.5: ImageNet (224 x 224), SSD ResNet-34: COCO 2017, BERT-Large: 维基百科 2020 年 1 月 1 日 (序列长度=512), DLRM: Criteo TB 数据集, RNN: LibriSpeech, MaskRCNN: COCO 2017.
- 6 用于 INT8 测试的软件配置: TensorFlow ResNet-50 v1.5, 推理: BS=116 (INT8), 单实例/路。oneDNN v2.7, 面向英特尔®架构优化的 TensorFlow 2.10。基于英特尔分别于 2022 年 10 月 24 日 (第三代和第四代英特尔®至强®可扩展处理器) 和 2022 年 7 月 19 日 (第二代和第一代英特尔®至强®可扩展处理器) 进行的测试。硬件配置: 第四代英特尔®至强®可扩展处理器配置 (已测): 配备 2 个英特尔®至强®铂金 8480 处理器的非量产平台 [56 核, 350 W 热设计功耗 (TDP)], 1 TB (8 通道/64 GB/4800 MHz) 的 DDR5 总内存, 使用 BKC 01, 使用英特尔®AMX/INT8 和 BF16, CentOS Stream 8, 英特尔®AMX 内核 (5.15), 测试结果可能不同。第三代英特尔®至强®可扩展处理器配置 (已测): 单节点, 2 个英特尔®至强®铂金 8380 处理器 (40 核/2.3 GHz, 270 W TDP), 1 TB (8 插槽/64 GB/3,200 MHz) DDR4 总内存, ucode 0xd0002f2, 启用英特尔®HT 技术, 启用英特尔®睿频加速技术, Ubuntu 20.04.2 LTS (Focal Fossa), 5.4.0-73-generic, 1 个英特尔®固态硬盘 SC2CV480A3 操作系统。第二代英特尔®至强®可扩展处理器配置 (已测): 单节点, 2 个英特尔®至强®铂金 8280 处理器, 28 核, 英特尔®HT 技术, 启用英特尔®睿频加速技术, 384 GB 总内存 (12 插槽/32 GB/2,933 MHz), BIOS: SE5C620.86B.02.0113.12152020065 (ucode: 0x500320a), CentOS Stream 8, 4.18.0-383.el8.x86\_64, 英特尔®至强®可扩展处理器配置 (已测): 单节点, 2 个英特尔®至强®铂金 8180 处理器, 28 核, 英特尔®HT 技术, 启用英特尔®睿频加速技术, 384 GB 总内存 (12 插槽/32 GB/2,666 MHz), BIOS: SE5C620.86B.0X.01.0117.021220182317 (ucode: 0x2006b06), Ubuntu 20.04.2 LTS, 5.4.0-73-generic.
- 7 基于假设 CPU 满载情况下单个内核在单个周期内进行矩阵乘法 + 累加运算的架构峰值处理能力。截至 2021 年 8 月。配合工作负载/配置信息请见 [www.intel.cn/PerformanceIndex](http://www.intel.cn/PerformanceIndex) (前往 Events 页下的 Architecture Day 2021)。结果可能不同。
- 8 有关全球 NLP 市场规模信息的来源: Fortune Business Insights, "Natural Language Processing (NLP) Market Size, Share & COVID-19 Impact Analysis, By Deployment (On-Premises, Cloud, Hybrid), By Enterprise Size (SMEs, and Large Enterprises), By Technology (Interactive Voice Response (IVR), Optical Character Recognition (OCR), Text Analytics, Speech Analytics, Classification and Categorization), By Industry Vertical (Healthcare, Retail, High Tech, and Telecom, BFSI) and Regional Forecast, 2022-2029" 按照部署 (本地、云端、混合)、企业规模 (中小企业和大型企业)、技术/交互语音应答 (IVR)、光学字符识别 (OCR)、文本分析、语音分析、分类)、行业垂直领域 (医疗、零售、高科技、电信及银行、金融服务和保险 (BFSI)) 等类别对 2022 年至 2029 年自然语言处理 (NLP) 市场规模、占有率及新进入者的分析及区域预测。2021 年 6 月, [fortunebusinessinsights.com/industry-reports/natural-language-processing-nlp-market-101933#](https://fortunebusinessinsights.com/industry-reports/natural-language-processing-nlp-market-101933#).

实际性能受使用情况、配置和其他因素的差异影响。更多信息请见 [www.intel.cn/PerformanceIndex](http://www.intel.cn/PerformanceIndex)。

性能测试结果基于配置信息中显示的日期进行的测试，且可能并未反映所有公开可用的安全更新。详情请参阅配置信息披露。没有任何产品或组件是绝对安全的。

具体成本和结果可能不同。

英特尔技术可能需要启用硬件、软件或激活服务。

英特尔并不控制或审计第三方数据。请您审查该内容，咨询其他来源，并确认提及数据是否准确。

© 英特尔公司版权所有。英特尔、英特尔标识以及其他英特尔商标是英特尔公司或其子公司的商标。其他的名称和品牌可能是其他所有者的资产。