



Workstations powered by Intel for AI developer tasks: Q&A

This document is a companion to the Principled Technologies report “[Workstations powered by Intel can play a vital role in CPU-intensive AI developer tasks.](#)” Please read that report to learn more and explore the [science behind the report](#) for a deeper dive into the hands-on testing we carried out.

Questions

Why should I care about AI?.....	2
Do we have to use the cloud or outfit our data center with servers that have GPUs?.....	2
Which parts of the AI development process might be well suited to workstations featuring Intel processors?.....	2
Why would we want to consider refreshing to new Intel processor-powered workstations for AI use cases?.....	3
How do we decide between tower workstations with processors from the Intel Xeon W-3400 processor collection and mobile workstations with 13th Gen Intel Core processors?.....	4
How can we maximize performance and minimize memory utilization when using tower workstations for AI tasks?	4
What other tools does Intel offer to help enhance my organization’s use of AI?.....	5

Why should I care about AI?

More and more organizations are adopting artificial intelligence (AI), and while chatbots and image generators might be the AI topics you most see in the news, the possible applications of AI are numerous. AI has the potential to speed medical diagnoses, improve the accuracy of safety and security solutions, and offer new capabilities in city planning, among other benefits. According to one survey, 95 percent of businesses are either already using AI or plan to start doing so in the next year.¹

Do we have to use the cloud or outfit our data center with servers that have GPUs?

It's true that applications and services that run in the cloud or on servers with purpose-built GPUs are useful for AI workflows, particularly inference and training workloads. However, these platforms are costly and in some cases may have components that are difficult to obtain. By employing on-site workstations for common development and prototyping AI tasks—which typically require less computing power than inference and training—organizations can avoid paying for more power than they need for less demanding tasks. Principled Technologies designed and executed three AI development workflows on tower and mobile workstations from three vendors, with each workflow utilizing only the computation capabilities of the Intel CPU cores, and found that these platforms were suitable for carrying out these AI tasks. We did not use GPUs for these tasks.

Which parts of the AI development process might be well suited to workstations featuring Intel processors?

The AI development workflows we used in our testing, which comprised several CPU-intensive activities, can run well on workstations with Intel processors. The first workflow characterized documents, added them to a vector database running on the system, and indexed this content. The second workflow used the disparate data sources in a multi-modal large language model (LLM) and combined them to determine the characteristics of a painting by Leonardo da Vinci. The third workflow involved standardizing images to a common scale and precision, and using a machine learning (ML) approach to find features in the images.

Why would we want to consider refreshing to new Intel processor-powered workstations for AI use cases?

Each generation of Intel processors introduces new efficiencies and features. The three tower workstations we tested feature processors from the new Intel Xeon W-3400 processor collection. According to Intel, platforms featuring these processors are “the ultimate workstation solution for professional creators, delivering outstanding performance, security, and reliability along with expanded platform capabilities for VFX, 3D rendering, complex 3D CAD, and AI development & edge deployments.”²

In several studies over the past four years, we at Principled Technologies have proved that users can see strong performance improvements by choosing new workstations powered by Intel, for both AI and other use cases. You can see several examples of this work here:

Optimize creative and design workflows and enjoy a better user experience with the Dell Precision 5680

In a series of tests, a Dell Precision 5680 Intel-based workstation delivered better results to rendering, faster than a 15-inch Apple MacBook Pro.

- Up to 24K and 60FPS performance
- Up to 3.2x the 3D rendering

Optimize for creative and design workflows and enjoy a better user experience with the Dell Precision 5680

HP ZBook G10 mobile workstation: Get the performance you need to take projects and action plans to the next level

HP ZBook G10 mobile workstation delivers the performance you need to take projects and action plans to the next level.

HP ZBook G10 mobile workstations: Get the performance you need to take projects and action plans to the next level

Lenovo ThinkBook 16p Gen 4: Spend wisely and be a pacesetter

The new ThinkBook 16p Gen 4 is a powerful workstation, perfect for professionals and students alike. It's the ultimate workstation for those who want to be a pacesetter.

Lenovo ThinkBook 16p Gen 4: Spend wisely and be a pacesetter

Upgrade to a mobile workstation with Windows 10 Pro for Workstations and an Intel Xeon E processor and save time on compute-intensive activities

- 68% less time to launch Adobe Photoshop
- 35% less time to launch Adobe Photoshop
- 30% more throughput
- Up to 60% higher CPU performance

Upgrade to a mobile workstation with Windows 10 Pro for Workstations and an Intel Xeon E processor and save time on compute-intensive activities

Champion big decisions and gutsy moves with the new HP Z8 Fury G5 Workstation Desktop PC

HP Z8 Fury G5 Workstation Desktop PC

The new HP Z8 Fury G5 Workstation Desktop PC: Crunch through demanding workloads in considerably less time

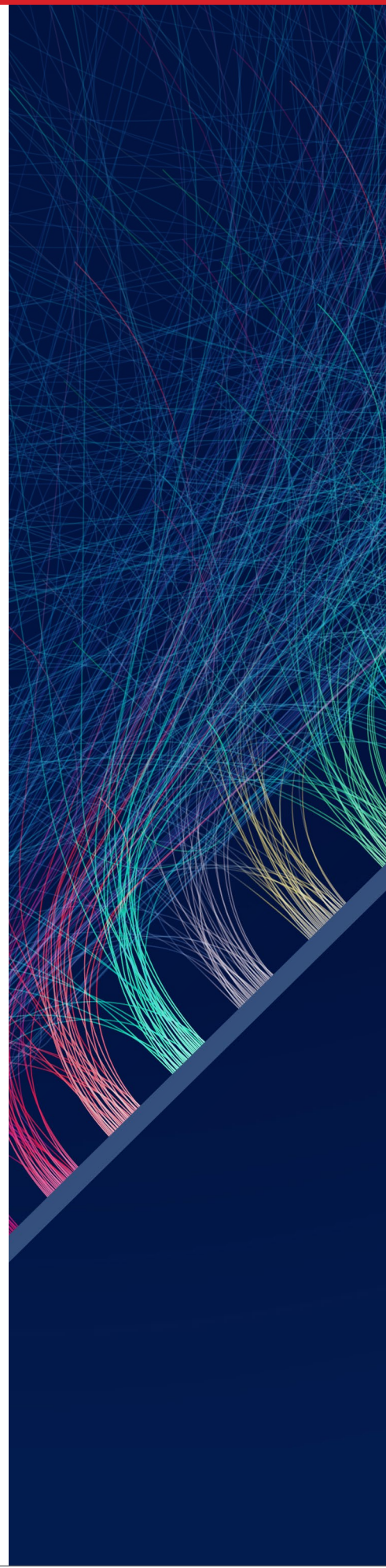
How do we decide between tower workstations with processors from the Intel Xeon W-3400 processor collection and mobile workstations with 13th Gen Intel Core processors?

In our testing of AI development workflows on tower and mobile workstations from three vendors, running two of the workflows on the tower workstations took roughly half as much time as on the mobile workstations. For the third workflow, the tower workstations could take advantage of processor features not available in the mobile workstation processors, and the difference was even more pronounced. (We explore this more in the following question.) The tower workstations also had significantly more memory than the mobile systems. Across the board, the mobile workstations performed well, but the tower workstations brought extra firepower to the tasks.

Based on these findings, we conclude that the tower workstations would be appropriate for a development environment for more complex models with a greater volume of data, while the mobile workstations would be well-suited for data scientists fine-tuning simpler models.

How can we maximize performance and minimize memory utilization when using tower workstations for AI tasks?

The processors in the tower workstations we tested support 16-bit floating point (FP) precision natively using AVX-512 extensions. When we ran one of our workflows with this precision as well as the standard 32-bit FP precision, we learned that using 16-bit precision reduced the time for the workflow by 26 percent. Memory usage improved even more dramatically with 16-bit precision, with an average reduction of 44 percent. These findings indicate that some of your AI workflows are also likely to benefit from using 16-bit FP precision, allowing you to attack larger problems or see faster results. (Note that for any given workload, you'll need to check that using reduced precision—16-bit rather than 32-bit—provides results at an acceptable accuracy.)



What other tools does Intel offer to help enhance my organization's use of AI?

AI Tools from Intel, formerly known as the Intel AI Analytics Toolkit, aim to maximize performance at all stages of the AI pipeline, from preprocessing through machine learning, and support efficient model development through interoperability.³ According to Intel, AI Tools "give data scientists, AI developers, and researchers familiar Python tools and frameworks to accelerate end-to-end data science and analytics pipelines on Intel architecture."⁴

AI Tools offer optimized frameworks for TensorFlow and PyTorch, model optimization tools, pre-trained models, analytics and AI optimization from Intel, and much more.⁵ In our testing, we used the Intel Python distribution and some of the optimized versions of Python packages that Intel AI Tools provide for ML and AI. Learn more about Intel AI Tools at <https://www.intel.com/content/www/us/en/developer/tools/>.

1. Kelly Lindenau, "The Big 'Book' Of Artificial Intelligence Statistics For B2B Practitioners," accessed June 26, 2024, <https://www.demandgenreport.com/uncategorized/the-big-book-of-artificial-intelligence-statistics-for-b2b-practitioners/47715/>.
2. Intel, "Intel® Xeon® W Processors," accessed April 10, 2024, <https://www.intel.com/content/www/us/en/products/details/processors/xeon/w.html>.
3. Intel, "AI Tools," accessed June 26, 2024, <https://www.intel.com/content/www/us/en/developer/tools/oneapi/ai-analytics-toolkit.html>.
4. Intel, "AI Tools."
5. Intel, "AI Tools."

Read the report at <https://facts.pt/8xoaOpQ>



Facts matter.®

Principled Technologies is a registered trademark of Principled Technologies, Inc. All other product names are the trademarks of their respective owners. For additional information, review the science behind this report.

This project was commissioned by Intel.