# Developer Tools for Intel® Xeon® 6 Processors

Maximize Performance for AI & Accelerated Compute Workloads on Intel Xeon 6 Processors with P-cores

Accelerate AI Frameworks & Applications

Accelerate Generative AI/LLM, and other deep learning, data science pipelines using the Intel® oneAPI Base Toolkit & Intel AI Tools.

- Intel oneAPI library optimizations are regularly up streamed to the latest versions of PyTorch,* TensorFlow,* and other leading deep learning frameworks, enabling developers to achieve orders of magnitude performance improvements on Intel hardware using their existing AI workflows.

- Intel® oneAPI Deep Neural Network Library (oneDNN) accelerates deep learning and generative AI models on Intel® Xeon 6 Processors with P-Cores, the first Intel CPU platform supporting AI acceleration with Intel® Advanced Matrix Extensions through FP16 and complex FP16 instructions (building on already existing int8 & BF16 support).
    - up to 3x better Llama2 performance vs. prior generation for large-language-model (LLM)[1]
    - up to 1.86x gen-to-gen performance improvement in AI inferencing[2]
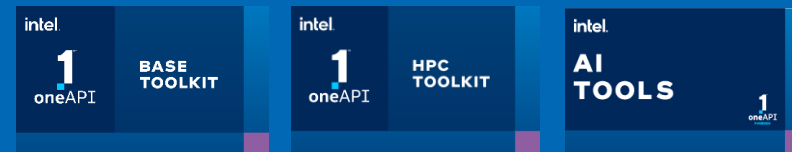
Accelerate AI & General Compute Workloads

Build, analyze, optimize, and scale applications with the latest techniques in vectorization, multithreading, multi-node parallelization, and memory, using the Intel® oneAPI Base Toolkit, Intel® Distribution for Python, Intel® oneAPI HPC Toolkit.

- Accelerate math functions across multiple domains such as BLAS, LAPACK and FFT with Intel® oneAPI Math Kernel Library (oneMKL) performance tuning for up to 2.5x better HPCG performance vs. prior generation with MRDIMM[3]

- Push your application's efficiency further with Intel® oneAPI DPC++/C++ Compiler's improved data access through preloading cache reducing latency & Intel® Advanced Matrix Extensions-FP16 instruction support leveraged by Intel® oneAPI Deep Neural Network Library (oneDNN).

- Intel® Fortran Compiler supports backend code generation and enriched performance tuning for latest Intel Xeon 6 processors.

- Intel® MPI Library now supports 128-core tuning and optimizations for scale out and scale up.

- Intel® VTune™ Profiler's new features such as hotspots, microarchitecture and memory access, I/O and platform diagram makes identifying performance bottlenecks and memory issues easier.

- Intel® Threading Building Blocks (oneTBB) is enhanced to scale parallel execution performance on Intel Xeon 6 processor's higher CPU core count to accelerate multi-threaded applications.

1) See [9A2] at intel.com/processorclaims: Intel® Xeon® 6
2) See [9A3] at intel.com/processorclaims: Intel® Xeon® 6
3) See [9H10] at intel.com/processorclaims: Intel® Xeon® 6