

인텔® HPC Engines
5세대 인텔® 제온® 스케일러블 프로세서

5세대 인텔® 제온® 스케일러블 프로세서와 인텔® HPC Engines를 사용한 HPC 가속

최대

1.31배 더 높은

LAMMPS 성능, 5세대 인텔® 제온® 스케일러블 플랫폼, 이전 세대 대비



고객 성공 사례: 인텔은 Taboola와 협력하여 인텔® 제온® 프로세서 예측 알고리즘을 최적화하고 벤치마크했습니다.

[사례 읽기 >](#)

교토대학교의 컴퓨팅 및 미디어 아카데미 센터에서는 인텔® 제온® 프로세서를 사용해 과학 연구 결과를 더 빨리 얻습니다.

[사례 읽기 >](#)

고성능 컴퓨팅(HPC)은 과학적 발견, 엔지니어링 시뮬레이션 및 복잡한 시스템의 모델링에 필수적입니다. 가속은 CPU 코어 수나 다른 하드웨어 및 소프트웨어 솔루션을 늘리지 않고 고성능을 실현하는 효율적이고 효과적인 대안이 될 수 있습니다. 5세대 인텔® 제온® 스케일러블 프로세서는 HPC 워크로드 성능과 전력 효율을 높일 수 있는 특정 목적을 위해 만든 가속기가 탑재된 상태로 제공됩니다.

인텔® HPC Engines를 통한 고성능 컴퓨팅

지난 몇 년 동안 HPC 소유자들이 컴퓨팅 속도 및 액세스를 늘려야 하는 필요성에 대응하면서 비용을 통제하는 방법에 변화가 있었습니다. 여러 업계에서는 HPC를 비즈니스 인사이트를 더 빨리 얻고 중요한 비즈니스 결정을 내리면서 비용도 절감하기 위한 도구로 점점 많이 사용하고 있습니다.

인텔® HPC Engines가 탑재된 5세대 인텔® 제온® 스케일러블 프로세서의 향상된 신규 기능은 시뮬레이션 및 모델링과 같이 가장 빠르게 증가하는 워크로드 유형 전반에서 성능을 향상시킬 수 있습니다. 5세대 인텔® 제온® 스케일러블 프로세서는 인텔® Advanced Matrix Extensions(인텔® AMX), 인텔® Data Streaming Accelerator(인텔® DSA), 인텔® QuickAssist Technology(인텔® QAT)와 같은 내장 가속기를 통해 대상 워크로드에 향상된 성능, 효율성 및 비용 절감 효과를 제공합니다.

인텔® Advanced Matrix Extensions

머신러닝(ML) 기술은 워크로드를 더 효율적이고 효과적이고 더 많은 인사이트를 얻을 수 있게 바꾸고 있습니다. 업계 동향에 따라, 고객은 비즈니스 결과를 개선하기 위해 HPC 및 AI 기반 솔루션을 사용하고 있습니다. 인텔® AMX는 AI 성능을 높이며, 인텔은 AI에 관한 자사의 전문성을 HPC 및 AI 솔루션을 모두 이용하는 고객들과 공유하고 있습니다.

5세대 인텔® 제온® 스케일러블 프로세서에 통합된 내장형 가속기 엔진 중 하나인 인텔® AMX는 딥러닝 추론 및 학습 성능을 높이는 인텔의 최신 발전 기술입니다. 인텔은 이전 인텔® 제온® 스케일러블 프로세서의 내장형 AI 가속기 기능을 확장한 인텔® AMX를 사용하여 큰 행렬 곱셈 연산을 완전히 바꿀 수 있었습니다. 인텔® AMX는 또한 2차원 레지스터 파일을 사용하여 더 큰 데이터 청크를 저장하기도 합니다. AI 워크로드를 가속하도록 만든 인텔® AMX는 HPC와 AI가 만나는 워크로드의 성능에 중요합니다.

인텔® Storage Engines: 스토리지 관련 워크로드에 사용되는 내장형 가속기

워크로드 가속기 엔진을 CPU에 통합하면 크게 세 가지 이점이 있습니다. 첫째, 내장형 가속기는 드롭인 가속기 카드와 외부 어플라이언스 고유의 I/O 병목 및 지연 시간을 해결합니다. 둘째, 특정 관련 워크로드를 CPU만 사용하는 경우보다 더 빨리 처리합니다. 셋째, 가속기를 사용해 CPU의 작업 부하를 덜고 여유 용량을 확보하여 성능이 더 높은 컴퓨팅 리소스가 필요한 워크로드에 사용할 수 있습니다.



인텔® Advanced Vector Extensions 512 (인텔® AVX-512) - 더 빠른 HPC를 위한 기초

모든 x86 CPU는 공통 명령어 집합 아키텍처(ISA)를 공유합니다. 인텔은 기본 x86 명령어를 새로운 워크로드에 확장하고 이런 명령어의 기능을 2011년의 인텔® Advanced Vector Extensions(인텔® AVX)부터 시작해 세대마다 발전시켰습니다. 이제 그런 기존의 인텔® AVX 명령어들과 그 후신인 인텔® AVX-512 및 인텔® AVX2는 일반 컴퓨팅, AI 프로세싱 및 수학 집약적인 HPC 워크로드를 가속합니다.

더 적은 단계는 더 빠른 처리를 의미

인텔® AVX-512의 “확장 기능”은 일반적인 컴퓨팅 작업을 더 적은 단계로 압축하고 결합하고 융합합니다. 간단한 예로 CPU에 $3 \times 3 \times 3 \times 3 \times 3$ 을 계산하라고 명령할 수 있는데, 이 계산에는 다섯 클럭 사이클이 소요됩니다. 아니면 CPU가 한 사이클에 수행할 수 있는 3^5 을 계산하라는 명령어를 만들 수도 있습니다. 인텔® AVX-512는 이 논리를 수백 가지의 작업별 연산에 적용합니다. 인텔® 제온® 스케일러블 프로세서에는 FMA(단일 곱셈-누산기) 유닛이 코어당 최대 2개까지 있어 곱셈과 덧셈을 연산 하나로 합쳐서 연산 속도를 높입니다.

인텔® QuickAssist Technology

5세대 인텔® 제온® 스케일러블 프로세서의 내장형 가속기인 인텔® QAT로 컴퓨팅 집약적인 워크로드의 부하를 덜어 사용 가능한 공간을 확보하고 비용을 절감하십시오. 인텔® QAT는 가속된 암호화, 키 보호 및 데이터 압축을 제공하여 시스템 리소스 사용량을 줄입니다. 그럼으로써 빅데이터 및 데이터베이스 애플리케이션에서 더 많은 Gbps 및 Ops/Sec 성능을 제공하여 고객에게 베니핏을 제공합니다.

인텔® QAT는 암호화 및 압축과 종종 연관되는 오버헤드를 줄여 궁극적으로 클러스터 성능을 개선하는 데 중요한 역할을 합니다. 따라서 암호화 및 데이터 압축 성능이 개선됨과 동시에 데이터 발생량이 줄어 각 코어가 더 많은 클라이언트를 서빙할 수도 있습니다.

인텔® Data-Streaming Accelerator

메모리, 스토리지 및 네트워킹 하위 시스템 안팎으로 이동하는 데이터의 여정은 CPU에 큰 부담이 될 수 있습니다.

인텔® 제온® 프로세서에 내장된 가속기인 인텔® DSA는 스트리밍 데이터 이동 및 변환 작업을 개선하여 스토리지, 네트워킹 및 데이터 집약적인 워크로드를 위한 고성능을 제공합니다. 인텔® DSA는 CPU, 메모리 및 캐시 사이와 모든 연결된 메모리, 스토리지 및 네트워크 장치 사이의 데이터 이동 속도를 높이는 데 도움이 됩니다.

고객은 OVS 부하를 인텔® Infrastructure Processing Unit (인텔® IPU)으로 보내 성능을 개선하고 CPU 효율을 더욱 더 최적화할 수 있습니다.

5세대 인텔® 제온® 스케일러블 프로세서에서는 HPC 가속이 기본으로 탑재되어 있습니다

HPC 가속을 위한 핵심적인 기초가 모든 인텔® 제온® 스케일러블 프로세서에 내장되어 있고, 대부분의 소프트웨어 프로그램과 함께 사용할 수 있습니다. HPC 고객은 큰 노력을 기울이지 않고도 이 기술의 이점을 누릴 수 있습니다.

인텔® oneAPI 기본 툴킷의 추가 기능인 인텔® HPC 툴킷은 최신 벡터화, 멀티스레딩, 멀티노드 병렬화 및 메모리 최적화 기법을 사용해 HPC 애플리케이션을 빌드하는 데 사용됩니다. 이 툴킷은 개방형 메시지 전달 인터페이스(Open MPI) 라이브러리 기반 클러스터 분석 및 튜닝 도구를 포함합니다. 한편, 인텔® oneAPI Math Kernel Library는 인텔® CPU와 인텔® GPU에 고도로 최적화되고 빠르고 완전한 수학 기능 라이브러리를 제공합니다.

HPC의 다음 시대를 위한 성능 가속

HPC의 접근성이 높아지고 비용이 낮아짐에 따라 슈퍼컴퓨팅 리소스의 상대적인 가치가 기하급수적으로 상승할 것입니다. 한때 국립 연구소나 글로벌 제조업체에서만 사용할 수 있었던 컴퓨팅 성능은 이제 클라우드 인스턴스와 하이브리드 HPC 클러스터를 통해 사용이 가능해지고 있습니다. 인텔® HPC Engines는 더 많은 조직이 새로운 발견을 하고 혁신하고 더 빨리 시장에 진출하기 위해 필요한 컴퓨팅 리소스에 접근할 수 있도록 HPC 성능을 전반적으로 개선할 수 있습니다.

인텔® 제온® 프로세서에 내장된 인텔® HPC Engines로 가장 까다로운 컴퓨팅 작업도 쉽게 처리하십시오.

인텔® AMX가 성능에 미치는 “깊은” 영향

인텔® AMX가 탑재된 5세대 인텔® 제온® 스케일러블 프로세서와 3세대 인텔® 제온® 스케일러블 프로세서 비교

최대

9.9배 더 높은 실시간 자연어 처리 추론 성능(BERT-large)과 7.7배 더 높은 성능/와트, AMX BF16이 탑재된 5세대 인텔® 제온®, 3세대 인텔® 제온® 프로세서 대비²

최대

2.3배 성능 가속, 5세대 인텔® 제온® 스케일러블 프로세서 사용 시, 3세대 인텔® 제온® 대비, GPT-J 퍼스트 토큰 지연 시간(int8)³

자세한 내용

[인텔® AVX-512 >](#)

[AI와 HPC 통합 >](#)

[인텔® 제온® 스케일러블 프로세서 기반 AI 및 딥러닝 솔루션 >](#)

5세대 인텔® 제온® 스케일러블 프로세서로 지금 클라우드 또는 자체 인프라에서 HPC 워크로드를 가속하기 시작하십시오.

intel.com/hpc를 방문하십시오



1. intel.com/processorclaims: 5세대 인텔® 제온® 스케일러블 프로세서에서 [H14]를 참조하십시오. 결과는 다를 수 있습니다.
2. intel.com/processorclaims: 5세대 인텔® 제온® 스케일러블 프로세서에서 [A19]를 참조하십시오. 결과는 다를 수 있습니다.
3. intel.com/processorclaims: 5세대 인텔® 제온® 스케일러블 프로세서에서 [A1]를 참조하십시오. 결과는 다를 수 있습니다.

*채널, 콘텐츠 제작 마케팅, 기타, 검색, 텍스트 분류, 텍스트 생성의 기하평균

고지 및 면책 조항

성능은 사용, 구성 및 기타 요인에 따라 다릅니다. 자세한 내용은 intel.com/PerformanceIndex에서 확인하십시오.

성능 결과는 구성에 표시된 날짜의 테스트를 기반으로 하며 공개된 모든 업데이트가 반영되어 있지 않을 수도 있습니다. 구성 백업 상세 정보를 확인하십시오. 어떤 제품 또는 구성 요소도 절대적으로 안전할 수는 없습니다.

워크로드 및 구성에 대해서는 www.intel.com/processorclaims에서 5세대 인텔® 제온® 스케일러블 프로세서를 참조하십시오. 결과는 다를 수 있습니다.

인텔® Advanced Vector Extensions(인텔® AVX)는 특정 프로세서 작업의 처리 속도를 높입니다. 여러 프로세서 성능 특성으로 인해, AVX 명령어를 이용하면 a) 일부분이 정격 주파수 미만으로 작동하고, b) 인텔® Turbo Boost Technology 2.0이 있는 일부분이 임의의 또는 최대 터보 주파수에 도달하지 못하는 원인이 될 수 있습니다. 성능은 하드웨어, 소프트웨어 및 시스템 구성에 따라 다르며, 자세한 내용은 다음 웹 페이지에서 확인할 수 있습니다. <https://www.intel.co.kr/content/www/kr/ko/products/details/processors/core.html>

인텔® 기술은 지원되는 하드웨어, 소프트웨어 또는 서비스 활성화가 필요할 수 있습니다.

비용과 결과는 다를 수 있습니다.

인텔은 인권을 존중하고 인권 침해에 연루되지 않도록 하기 위해 노력합니다. 인텔의 [글로벌 인권 원칙](#)을 참조하십시오. 인텔® 제품과 소프트웨어는 국제적으로 인정되는 인권의 침해를 초래하거나 악화시키지 않는 애플리케이션에만 사용해야 합니다.

© 인텔사. 인텔, 인텔 로고 및 기타 인텔 마크는 인텔사 또는 그 자회사의 상표입니다. 기타 명칭 및 브랜드는 해당 소유주의 자산일 수 있습니다. 0922/MP/CMD/PDF

가속기의 가용성은 SKU에 따라 다릅니다. 제품 세부 정보를 더 보려면 [인텔® 제품 사양](#) 페이지를 방문하십시오.