

4세대 인텔® 제온® 스케일러블 프로세서 (코드명 사파이어 래피즈)



가장 많이 내장된 가속기로 최고의 성능을 구현합니다.

조직이 규모를 확장하고 비용을 절감하며 새로운 서비스를 제공하려고 할 때 따라 기술이 비즈니스 가치를 제공하는 것이 그 어느 때보다 중요해졌습니다. 더욱 복잡해지는 새로운 애플리케이션에 맞게 시스템을 맞춤화 하는 대신에 기업은 스케일러블 플랫폼을 통해 현재와 미래의 다양한 시스템 구축을 하기 위한 필요한 성능을 달성할 수 있습니다. 4세대 인텔® 제온® 스케일러블 프로세서는 인공지능(AI), 데이터 분석, 네트워킹, 스토리지 및 고성능 컴퓨팅(HPC)에서 가장 빠르게 성장하는 워크로드 전반에서 성능을 가속화하도록 설계되었습니다.

4세대 인텔® 제온® 스케일러블 프로세서에는 시중에 나와 있는 프로세서 중 가장 많은 내장 가속기를 포함하고 있습니다. 고급 보안 기술을 통해 민감하거나 규제된 데이터가 있는 경우에도 비즈니스 협업 및 인사이트를 위한 새로운 기회를 열어주는 동시에 제로 트러스트 보안 전략을 실현하는 데 도움이 됩니다. 여러 클라우드 및 에지에서 확장하여 배포 요구 사항을 충족할 수 있습니다. 또한 4세대 인텔® 제온® 스케일러블 프로세서는 원활한 애플리케이션 이식성과 함께 클라우드 선택에 있어 가장 많은 선택권과 유연성을 제공합니다.

인텔 가속기 엔진은 성능을 재정의합니다.

여러분이 프로세서에 기대하는 것을 재정의하십시오. 내장된 가속 기능은 CPU 코어 수를 늘리는 것보다 더 높은 성능을 달성할 수 있는 보다 효율적인 대체 방법입니다. 내장형 가속기 및 소프트웨어 최적화를 통해 이전 세대 인텔® 제온® 스케일러블 프로세서는 목표로 하는 실제 워크로드에서 와트당 최고의 성능을 제공하는 것으로 나타났습니다. 이를 통해 보다 효율적인 CPU 활용, 낮은 전력 소비 및 높은 수익을 얻을 수 있게 함으로써 기업이 지속 가능성 목표를 달성할 수 있도록 지원합니다.

■ 인텔® 고급 매트릭스 익스텐션(인텔® AMX)은 자연어 처리(NLP), 추천 시스템 및 이미지 인식과 같은 딥 러닝(DL) 추론 및 교육 워크로드를 가속화합니다.

■ 인텔® 데이터 스트리밍 가속기(인텔® DSA)는 스트리밍 데이터 이동 및 변환 작업을 개선하여 스토리지, 네트워킹 및 데이터 집약적인 워크로드를 위한 고성능을 제공합니다.

■ 인텔® 인메모리 분석 가속기(인텔® IAA)는 분석 성능을 향상시키는 동시에 CPU 코어에서 작업을 오프로드하여 데이터베이스 쿼리 처리량 및 기타 워크로드를 가속화합니다.

■ 인텔® 다이내믹 로드 밸런서(Intel® DLB)는 시스템 로드가 변화함에 따라 여러 CPU 코어에 네트워크 데이터를 동적으로 분산하여 효율적인 하드웨어 기반 로드 밸런싱을 제공합니다.

■ vRAN용 인텔® 고급 벡터 익스텐션(인텔® AVX)은 동일한 전력 엔벨로프로 이전 세대에 비해 vRAN(가상 무선 액세스 네트워크) 밀도를 최대 2배까지 증가시킵니다.

■ Intel® Advanced Vector Extensions 512 (Intel® AVX-512)는 최대 2개의 FMA(Fused-Multiply Add) 장치를 지원하고 까다로운 컴퓨팅 작업을 위한 성능을 가속화하는 최적화를 포함합니다.

■ 인텔® 퀴어시스트 기술(인텔® QAT)은 암호화, 암호 해독 및 데이터 압축을 가속화하여 프로세서 코어에서 이러한 작업을 오프로드하여 시스템 리소스 소비를 줄이는 데 도움을 줍니다.

■ 인텔® 크립토 가속은 SSL(Secure Sockets Layer) 웹 서버, 5G 인프라 및 VPN/방화벽과 같은 암호화에 민감한 워크로드의 성능을 향상시키고 퍼베이시브 데이터 암호화를 구현하는 데 따른 불이익을 줄여줍니다.

AI(인공지능)

가속화된 벡터 명령과 매트릭스 다중 연산을 통해 4세대 인텔® 제온® 스케일러블 프로세서는 탁월한 AI 추론 및 훈련 성능을 제공합니다. 인텔 AMX는 권장 시스템, NLP, 이미지 인식, 미디어 처리 및 전달, 미디어 분석과 같은 딥러닝 워크로드의 성능을 크게 향상시킬 수 있습니다.

HPC(고성능컴퓨팅)

4세대 인텔® 제온® 스케일러블 프로세서는 제조 시뮬레이션, 분자 역학, 지구 시스템 모델링, AI 추론 및 교육에서 발견되는 HPC 워크로드에서 흔히 발생하는 스레드가 많은 코드의 성능을 개선할 준비가 되어 있습니다. 내장된 가속기는 AI 데이터 유형의 처리 속도를 높이는 동시에 높은 수준의 정밀도를 제공합니다. 그리고 DDR5 메모리, PCIe Gen5, Intel® UPI(Ultra Path Interconnect) 2.0 및 Compute Express Link(CXL)에 대한 지원도 전체 데이터 처리량을 향상시킵니다.

데이터 분석

내장된 가속기는 메모리 내 데이터베이스, 빅 데이터, 데이터 웨어하우징, 비즈니스 인텔리전스(BI), 전사적 자원 관리(ERP) 및 운영 분석의 성능을 향상시킵니다. 인텔 DSA는 데이터 처리 집약적인 애플리케이션에서 일반적인 스트리밍 데이터 이동 및 변환 작업을 개선하는 반면, 인텔 IAA는 CPU 코어에서 작업을 오프로드하여 데이터베이스 작업 처리량을 가속화합니다.

네트워크와 스토리지

인텔 DLB는 코어 간의 작업 균형을 유지하고 네트워크 패킷 우선 순위를 제공합니다. 인텔 DSA는 데이터 복사 및 일반 데이터 변환 작업을 오프로드하여 코어 주기를 확보합니다. 이러한 내장형 가속기는 효율적인 네트워크 데이터 배치, 엔터프라이즈 스토리지 데이터 이동을 가능하게 하고 클라우드 컴퓨팅의 향상된 메모리 관리 작업을 통해 클라우드 컴퓨팅을 향상시킵니다.

암호화

이제 4세대 인텔® 제온® 스케일러블 프로세서에 내장된 인텔 QAT는 암호화 및 압축을 가속화합니다. 인텔 QAT는 CPU 효율성과 응용 프로그램 처리량을 크게 높이는 동시에 데이터 공간과 전력 사용을 줄여 조직이 성능 저하 없이 암호화를 강화할 수 있도록 합니다.

보안

인텔® 소프트웨어 가드 익스텐션(인텔® SGX)는 최근에 데이터 센터에서 가장 많이 연구, 업데이트 및 배포된 기밀 컴퓨팅 기술로, 데이터 센터의 기밀 컴퓨팅 기술 중 가장 적은 신뢰 경계를 갖고 있습니다.

최대 **1.53배**

이전 세대에 비해 평균 성능 향상

최대 **10배**

이전 세대(FP32)에 비해 내장 인텔 AMX(BF16)를 통해 실시간 추론 및 교육을 위한 더 높은 PyTorch 성능 구현

최대 **3배**

이전 세대(FP32)에 비해 내장 인텔 AMX(BF16)를 통해 실시간 추론 및 교육을 위한 더 높은 PyTorch 성능 구현

최대 **1.6배**

이전 세대에 비해 통합 인텔 DSA를 사용하여 더 높은 IOPS(초당 입/출력 작업) 및 대용량 패킷 순차 읽기 대기 시간 최대 37% 감소

최대 **2배**

이전 세대 프로세서와 비교하여 vRAN 워크로드에 대해 동일한 전력 범위에서의 용량

최대 **95%**

이전 세대에 비해 통합 Intel QAT를 사용하여 더 적은 코어 및 2배 더 높은 레벨-1 압축 처리량 구현

기술 소개

4세대 인텔® 제온® 스케일러블 프로세서는 이전 세대보다 코어당 성능이 더 높은 새로운 아키텍처를 특징으로 합니다. 또한 소켓당 최대 60개의 코어와 시스템당 1개, 2개, 4개 또는 8개의 소켓을 제공합니다. 이러한 코어 수 증가의 균형을 맞추기 위해 플랫폼은 메모리 및 입/출력(I/O) 하위 시스템의 동반 발전을 제공합니다. DDR5 메모리는 DDR4보다 최대 1.5배의 대역폭과 속도(4,800MT/s)를 제공합니다. 또한 이 플랫폼은 소켓당 80개의 PCIe Gen5 레인을 갖추고 있어 이전 플랫폼에 비해 I/O가 크게 향상되었습니다.

높은 페브리크 대역폭과 연결된 가속기 효율성을 지원하는 CXL 1.1을 제공합니다. 4세대 인텔® 제온® 스케일러블 프로세서는 워크로드 요구 사항이 변경됨에 따라 확장 및 적용할 수 있는 기술을 지원합니다. 또한 다음을 수행할 수 있습니다.

- 인텔® 인프라스트럭처 프로세스 유닛(인텔® IPU)에 무거운 작업을 오프로드하여 CPU 활용도를 개선하는 동시에 네트워크, 스토리지 및 컴퓨팅 성능을 추가로 향상
- 인텔 UPI 2.0으로 멀티 소켓 대역폭 증가(초당 최대 16GT/s)
- 인텔® 스피드 셀렉트 기술(인텔® SST)로 특정 워크로드 요구 사항을 충족하도록 CPU 구성
- 공유 마지막 레벨 캐시(LLC) 증가
(모든 코어에서 공유되는 최대 100MB LLC)
- 하드웨어 강화 보안으로 보안 태세 강화
- 인텔® CPU 가상 RAID(인텔® VROC)로 별도의 RAID 카드가 필요하지 않음

인텔 제온 스케일러블 프로세서를 통한 선택의 폭과 유연성 확장

하드웨어에서 시스템, 소프트웨어에 이르기까지 인텔은 조직이 계속 확장되는 비즈니스 목표를 달성하는 동시에 데이터를 더 안전하게 유지하도록 지원하도록 설계된 신뢰할 수 있는 기술 기반을 제공합니다. 더 큰 컴퓨팅 밀도를 제공하여 전력 사용량을 줄이고, AI 워크플로우를 가속화하거나 클라우드 네이티브 아키텍처로의 전환을 지원하는 등 인텔 제온 스케일러블 프로세서는 가장 중요한 비즈니스 과제를 해결하는 동시에 최고의 클라우드 선택과 애플리케이션 이식성을 제공합니다.

4세대 인텔® 제온® 스케일러블 프로세서에 적용된 새로운 기능

PCI Express Gen5 (PCIe 5.0)

CPU와 연결된 장치 사이에서 가능한 최고의 처리량을 가능하게 하는 기회를 통해 새로운 I/O 속도를 경험하십시오. 4세대 Intel Xeon Scalable 프로세서는 빠른 네트워킹, 고대역폭 가속기 및 고성능 저장 장치에 이상적인 최대 80개의 PCIe 5.0 레인을 갖추고 있습니다. PCIe 5.0은 PCIe 4.0,9의 I/O 대역폭을 두 배로 늘리고 이전 버전과의 호환성을 유지하며 CXL을 위한 기본 슬롯을 제공합니다.

DDR5

더 높은 메모리 대역폭으로 데이터 병목 현상을 극복하여 컴퓨팅 성능을 개선합니다. DDR5는 DDR4보다 최대 1.5배 향상된 대역폭을 제공하여 성능, 용량, 전력 효율성 및 비용을 개선할 수 있는 기회를 제공합니다. 4세대 인텔® 제온® 스케일러블 프로세서는 DDR5로 최대 4,800MT/s(1DPC) 또는 4,400MT/s(2DPC)를 제공합니다.

CXL

차세대 워크로드를 위한 CXL 1.1로 데이터 센터의 컴퓨팅 대기 시간을 줄이고 총소유비용(TCO)을 낮추십시오. CXL은 표준 PCIe 물리적 계층에서 실행되는 대체 프로토콜이며 동일한 링크에서 표준 PCIe 장치와 CXL 장치를 모두 지원할 수 있습니다. CXL은 CPU와 가속기 사이에 통합되고 일관된 메모리 공간을 생성하는 중요한 기능을 제공하며 데이터 센터 서버 아키텍처가 구축되는 방식을 혁신적으로 변화시킬 것입니다.

4세대 인텔 제온 스케일러블 프로세서 소개

인텔 제온 플래티넘 8400 프로세서 시리즈는 보안이 지원되고 민첩한 하이브리드 클라우드 데이터 센터를 위한 프로세서입니다. 이 프로세서는 고급 데이터 분석, AI, 고밀도 인프라 및 멀티클라우드 워크로드용으로 설계되었습니다. 이 프로세서는 높은 수준의 성능, 플랫폼 기능 및 업계 최고의 워크로드 가속화를 제공합니다. 일부 인텔 제온 플래티넘 8400 프로세서 시리즈에서 최대 8소켓 프로세서까지 향상된 하드웨어 기반 보안과 뛰어난 멀티 소켓 처리 성능을 제공합니다. 신뢰할 수 있고 하드웨어로 강화된 데이터 서비스 제공과 새로운 I/O 및 연결 기술을 통해 이 프로세서는 I/O, 메모리, 스토리지 및 네트워크 기술의 개선을 제공하여 점점 더 데이터에 의존하는 세계에서 실행 가능한 통찰력을 활용합니다. 아래의 기술들이 포함되어 있습니다.

- 인텔 제온 스케일러블 프로세서당 최대 60개 코어
- 최대 4,800 MT/s(1 DPC)의 프로세서당 8개의 메모리 채널
- 인텔 AMX를 통한 AI 가속화로 딥러닝 추론 및 교육 성능의 큰 도약

최대 4개의 소켓 확장성을 갖춘 인텔 제온 골드 6400 프로세서 시리즈 및 인텔 골드 5400 프로세서 시리즈는 요구사항이 많은 메인스트림 데이터 센터, 멀티 클라우드 컴퓨팅, 네트워크 및 스토리지 워크로드에 최적화되어 있습니다. 더 빠른 메모리 속도와 향상된 메모리 용량을 지원하는 이 프로세서는 향상된 성능, 향상된 메모리 기능, 하드웨어 강화된 보안 및 워크로드 가속을 제공합니다.

인텔 제온 실버 4400 프로세서 시리즈는 필수적인 성능, 향상된 메모리 속도 및 전력 효율성을 제공합니다. 엔트리 레벨 데이터 센터 컴퓨팅, 네트워크 및 스토리지에 필요한 하드웨어 향상된 성능을 제공합니다.

인텔의 새로운 프로세서가 귀하의 비즈니스를 발전시키는 방법에 대한 자세한 내용은 intel.com/xeonscalable 및 intel.com/4thGenXeon를 참조하십시오.



최대 8개 프로세서 소켓 지원

16GT/s Intel UPI 포트 4개

CXL 포함한 PCIe 5.0 레인 80개

최대 4,800MT/s(채널당 DIMM 1개) 또는 4,400MT/s(채널당 DIMM 2개)의 DDR5 지원

인텔 AVX-512(512비트 FMA 2개)

인텔 HT기술 및 인텔 터보 부스트 기술

인텔 딥러닝 부스트 및 인텔 AMX

인텔 SST

고급 신뢰성, 가용성 및 서비스(RAS) 기능

Intel SGX 최대 128GB 엔클레이브 크기 (일부 SKU에서 최대 512GB 엔클레이브 크기)

인텔 QAT, 인텔 DLB, 인텔 DSA 및 인텔 IAA로 워크로드 가속화

최대 4개 프로세서 소켓 지원

16GT/s Intel UPI 포트 3개

CXL 포함한 PCIe 5.0 레인 80개

최대 4,800MT/s(채널당 DIMM 1개) 또는 4,400MT/s(채널당 DIMM 2개)의 DDR5 지원

인텔 AVX-512(512비트 FMA 2개)

인텔 HT기술 및 인텔 터보 부스트 기술

인텔 딥러닝 부스트 및 인텔 AMX

인텔 SST

고급 신뢰성, 가용성 및 서비스(RAS) 기능

Intel SGX 최대 128GB 엔클레이브 크기

인텔 QAT, 인텔 DLB, 인텔 DSA 및 인텔 IAA로 워크로드 가속화

최대 2개 프로세서 소켓 지원

16GT/s Intel UPI 포트 2개

CXL 포함한 PCIe 5.0 레인 80개

최대 4,800MT/s(채널당 DIMM 1개) 또는 4,400MT/s(채널당 DIMM 2개)의 DDR5 지원

인텔 AVX-512(512비트 FMA 2개)

인텔 HT기술 및 인텔 터보 부스트 기술

인텔 딥러닝 부스트 및 인텔 AMX

Intel SGX 최대 64GB 엔클레이브 크기

인텔 QAT, 인텔 DLB, 인텔 DSA 및 인텔 IAA로 워크로드 가속화