



intel.  
XEON®

Wie man das Maximum aus  
skalierbaren Intel® Xeon®  
Prozessoren mit integrierten  
Beschleunigern herausholt

E-Guide

## Inhaltsverzeichnis

Was ist integrierte Beschleunigung und warum sollten Sie sie nutzen?	3
Die praktischen Vorteile der integrierten Beschleuniger von Intel	4
Welche integrierten Beschleuniger von Intel sind für Ihr Unternehmen geeignet?	5
KI – Intel® Deep Learning Boost	6
HPC – Intel® Advanced Vector Extensions 512	8
Sicherheit – Intel® Software Guard Extensions	10
Die nächste Generation integrierter Beschleuniger von Intel	12
Fazit	12



# Was ist integrierte Beschleunigung und warum sollten Sie sie nutzen?

---




Wie wäre es, wenn Sie nicht jedes Mal neue Geräte kaufen müssen, wenn Sie neue Funktionen benötigen, sondern auf bereits in Ihre CPU integrierte Technik zurückgreifen könnten? Mit den skalierbaren Intel® Xeon® Prozessoren ist das möglich. Diese CPUs besitzen als integrierte Beschleuniger bekannte Funktionen, die Ihren Workloads zusätzlichen Nutzen bringen.

Alle skalierbaren Intel® Xeon® Prozessoren verfügen über die größte und beste Auswahl an integrierten Beschleunigern. Diese steigern die Leistung ebenso wie die Effizienz und verringern den Bedarf an zusätzlicher Spezialhardware. Sowohl in Cloud- als auch in lokalen Umgebungen unterstützen diese speziellen Funktionen die heute gängigsten und anspruchsvollsten Workloads wie KI, Sicherheit, HPC, Analytik, Storage und Netzwerk.

**In diesem E-Guide konzentrieren wir uns auf KI, HPC und Sicherheit.**

# Die praktischen Vorteile der integrierten Beschleuniger von Intel

Ganz gleich, ob Sie skalierbare Intel® Xeon® Prozessoren für Ihre Workloads lokal, in der Cloud oder am Edge einsetzen, unsere integrierten Beschleuniger helfen Ihrem Unternehmen, neue Ziele zu erreichen. Sie bieten eine Reihe von Vorteilen wie schnellere Sicherheitsverarbeitung, stärkeren Datenschutz und bessere Auslastung der Infrastruktur. Vor allem sorgen diese integrierten Beschleuniger für **eine höhere Leistung bei Anwendungen, niedrigere Kosten und eine verbesserte Energieeffizienz:**

<b>Leistung</b> 	Als eigens entwickelte Komponente können die integrierten Beschleuniger von Intel oft mehr Leistung für einen bestimmten Workload bereitstellen. <sup>1</sup>
<b>Kosteneinsparungen</b> 	Mit integrierten Beschleunigern von Intel lässt sich die Leistung steigern, ohne zusätzliche Spezialhardware.
<b>Energieeinsparungen</b> 	Dank integrierter Beschleuniger von Intel kann die Energieeffizienz verbessert werden, da das Server-Rack ohne zusätzliche Kerne auskommt.



# Welche integrierten Beschleuniger von Intel sind für Ihr Unternehmen geeignet?

Die skalierbaren Intel® Xeon® Prozessoren verfügen zwar über eine ganze Reihe an integrierten Beschleunigern, aber bestimmte Beschleuniger sind für bestimmte Aufgaben/Workloads besonders geeignet. Um Ihnen die Entscheidung zu erleichtern, welche Intel® Technik Ihrem Unternehmen am nützlichsten ist, werfen wir einen Blick auf drei unserer wichtigsten Angebote in den Bereichen KI, HPC und Sicherheit.

## KI

**Intel® Deep Learning Boost (Intel® DL Boost)** sorgt für erhebliche Leistungssteigerungen bei der Inferenz und dem Trainieren gängiger KI- und HPC-Workloads.<sup>2</sup>

## HPC

**Intel® Advanced Vector Extensions 512 (Intel® AVX-512)** wurde speziell dafür entwickelt, die Leistung für die anspruchsvollsten Rechen-Workloads in Wissenschaft, Wirtschaft und darüber hinaus zu steigern.

## Sicherheit

**Intel® Software Guard Extensions (Intel® SGX)** schützt Daten während ihrer Nutzung durch einzigartige Technologie zur Isolierung von Anwendungen.



# KI – Intel® Deep Learning Boost

## Was ist Intel® Deep Learning Boost?

Intel® DL Boost ist ein Beschleuniger, der dazu dient, die Leistung zu steigern und KI- und Deep-Learning-bezogene Aufgaben und Workloads effizienter zu machen.

Bei den skalierbaren Intel® Xeon® Prozessoren wurde spezielle KI-Beschleunigung 2019 mit Intel® Vector Neural Network Instructions (Intel® VNNI) – jetzt Intel® DL Boost – eingeführt.

Die auf dem Beschleuniger Intel® AVX-512 basierende VNNI-Komponente von Intel® DL Boost vereint drei Befehlssätze zu einem. Das verkürzt die Zeit für die Erledigung einer Aufgabe stark.

## Was sind die gängigsten Anwendungsfälle für Intel® DL Boost?

Intel® DL Boost beschleunigt verschiedene KI-Inferenz-Aufgaben wie Bildklassifizierung, Sprachübersetzung und Objekterkennung.

## Wie nutzen Unternehmen Intel® DL Boost in der Praxis?



Der Softwarehersteller [rinf.tech](#) nutzte Intel® DL Boost für eine schnellere und genauere Bildanalyse, die eine bessere Echtzeit-Entscheidungsfindung bei Anwendungsfällen im Einzelhandel, in der Automobilindustrie, bei der Videoüberwachung und bei Business Intelligence unterstützt. Die Inferenzleistung konnte um das bis zu 7,4-fache gesteigert werden.<sup>3</sup>



[Huiyi Huiying Medical Technology \(HYHY\)](#) hat Intel® DL Boost dazu genutzt, die Leistung von KI-Komplettlösungen für die medizinische Bildgebung zu optimieren. Dank der Vorteile der synergetischen Software-Hardware-Beschleunigung konnte das Unternehmen die Inferenzgeschwindigkeit bei Bildanalyseszenarien wie COVID-19-Screening und Brustkrebserkennung wesentlich verbessern.<sup>4</sup>



[Konfoong Biotech International Co., Ltd. \(KFBIO\)](#), das auf die Entwicklung und Herstellung digitaler Pathologiesysteme spezialisiert ist, konnte mittels Intel® DL Boost die Untersuchung und Diagnose von M. tuberculosis-Proben bis zu 11,4 mal schneller erledigen.<sup>5</sup>



## Welche Leistungsvorteile bietet Intel® DL Boost?

---

**Kunden, die die Optimierungen von Intel** für TensorFlow\* und Intel® DL Boost nutzen, **erzielen mehr als**

**11x** mehr KI-Inferenzleistung

bei skalierbaren Intel® Xeon® Prozessoren der 3. Generation im Vergleich zu skalierbaren Intel® Xeon® Prozessoren der 2. Generation.<sup>6</sup>

---

**Intel® DL Boost bietet zudem KI-Workloads mehr Leistung pro Watt, wodurch Unternehmen ihre Kosten und ihren Energieverbrauch senken können.**

# HPC – Intel® Advanced Vector Extensions 512

## Was ist Intel® Advanced Vector Extensions 512?

Intel® AVX-512 ist ein leistungssteigernder Allzweck-Beschleuniger mit einem breiten Anwendungsspektrum. Dank seiner 512 Bit breiten Vektoroperationen eignet sich Intel® AVX-512 besonders für die anspruchsvollsten Rechenaufgaben im HPC-Bereich.

## Was sind die gängigsten Anwendungsfälle für Intel® AVX-512?

Intel® AVX-512 wird im Bildungs-, Kommunal-, Finanz-, Unternehmens-, Technik- und Medizinbereich für eine Vielzahl von komplexen Aufgaben eingesetzt. Dazu gehören KI, wissenschaftliche Simulationen, 3D-Modellierung und -Analyse, Finanzanalytik, Audio- und Videoverarbeitung, Kryptografie und Datenkomprimierung.



Intel® AVX-512 ermöglicht Echtzeit-Analytik für Finanzdienstleistungs-Workloads – für ein besseres Kundenerlebnis, verbesserte Compliance und höheren Datenschutz.



Dank Intel® AVX-512 können komplexe Workloads auf bestehender Hardware ausgeführt werden. Zudem wird die Leistung für Aufgaben wie 3D-Modellierung und -Simulation gesteigert.

## Wie nutzen Unternehmen Intel® AVX-512 in der Praxis?



**Die University at Buffalo** nutzt an ihrem Center for Computational Research Intel® AVX-512 und bietet damit Unternehmen im westlichen Bundesstaat New York umfangreiche Rechenressourcen – wie zum Beispiel Marion Surgical, einem Unternehmen, das virtuelle Realität (VR) und Augmented Reality (AR) nutzt, um Chirurgen in komplexen Abläufen zu schulen.<sup>7</sup>



**Das Broad Institute of MIT and Harvard** nutzt Intel® AVX-512, um bei Genomik-Workloads auf Google Cloud\* N1 und N2 Instanzen die Verarbeitungsgeschwindigkeit zu steigern und die Kosten zu senken.<sup>8</sup>



Wissenschaftler:innen, die bei **CERN**, der Europäischen Organisation für Kernforschung, Elementarteilchen untersuchen, haben Intel® AVX-512 dafür genutzt, Simulations-Workloads mittels Quantifizierung zu beschleunigen. Das führte zu einer **1,8-mal so hohen Leistung** und einer leicht verbesserten Genauigkeit.<sup>9</sup>



## Welche Leistungsvorteile bietet Intel® AVX-512?

Da HPC-Workloads riesige Datenmengen verarbeiten, ist die CPU-Leistung entscheidend dafür, genaue Ergebnisse zeitnah bereitzustellen.

Intel® AVX-512 verbessert die Vektorverarbeitungsleistung von Intel® Xeon® CPUs im Vergleich zu Lösungen der vorherigen Generation und ermöglicht es Unternehmen, intensive Workloads mit höherer Geschwindigkeit zu verarbeiten. Bei den entsprechenden Programmen können pro Sekunde bei jedem Takt 32 Fließkommaoperationen mit doppelter Genauigkeit und 64 mit einfacher Genauigkeit innerhalb der 512-Bit-Vektoren sowie acht 64-Bit- und sechzehn 32-Bit-Integer-Werte mit bis zu zwei 512 Bit breiten FMA-Einheiten (Fused-Multiply Add) verarbeitet werden. Dabei sind im Vergleich zu Intel® Advanced Vector Extensions 2 (Intel® AVX2) die Datenregister doppelt so breit, es sind doppelt so viele Register vorhanden und die FMA-Einheiten sind ebenfalls doppelt so breit.

Im Vergleich zu Lösungen von AMD boten die skalierbaren Intel® Xeon® Prozessoren der 3. Generation mit Intel® AVX-512 eine überragende Leistung bei 12 HPC-bezogenen Benchmarks und praxisorientierten Anwendungen.<sup>10</sup>

Beim biomolekularen Simulationsalgorithmus NAMD boten die skalierbaren Intel® Xeon® Prozessoren der 3. Generation

**1,27x** mehr Leistung

im Vergleich zu.m **AMD Milan\***.<sup>11</sup>

Bei dem für 3D-Bildgebung genutzten Algorithmus RELION boten die skalierbaren Intel® Xeon® Prozessoren der 3. Generation

**1,32x** mehr Leistung

im Vergleich zum **AMD Milan\***.<sup>12</sup>

Beim Simulationsalgorithmus für Finanzinvestitionen Monte Carlo FSI boten die skalierbaren Intel® Xeon® Prozessoren der 3. Generation

**1,50x** mehr Leistung

im Vergleich zum **AMD Milan\***.<sup>13</sup>

Beim Modellierungs- und Simulationsalgorithmus LINPACK boten die skalierbaren Intel® Xeon® Prozessoren der 3. Generation

**1,18x** mehr Leistung

im Vergleich zum **AMD Milan\***.<sup>14</sup>

# Sicherheit – Intel® Software Guard Extensions

## Was ist Intel® Software Guard Extensions?

Intel® SGX bietet eine hardwarebasierte Sicherheitslösung, die Daten während ihrer Nutzung durch Technologie zur Isolierung von Anwendungen schützt. Durch den Schutz ausgewählter Codeteile und Daten vor Einsicht oder Modifizierung können Entwickler sensible Datenoperationen innerhalb von Enklaven ausführen und so die Anwendungssicherheit erhöhen oder die Daten besser schützen. Das bildet eine zusätzliche Abwehrebene, die die Angriffsfläche des Systems verringert.

## Was sind die gängigsten Anwendungsfälle für Intel® SGX?

Intel® SGX ermöglicht es Confidential-Computing-Lösungen, Daten lokal, am Edge oder in der Cloud besser zu schützen. Dieser Beschleuniger unterstützt Unternehmen beim Schutz von sensiblen Daten und sensiblem Code, damit diese die Erfüllung der Vorschriften in Bezug auf Schutz, Souveränität und Vertraulichkeit von Daten sicherstellen können.

## Wie nutzen Unternehmen Intel® SGX in der Praxis?



Die britische Bausparkasse **Nationwide Building Society** nutzte Intel® SGX für die Entwicklung von „Know Your Customer“, einem System, das **eine sicherere Verarbeitung von mehreren vertraulichen Datenbeständen innerhalb einer Enklave ermöglicht** – unter Einhaltung der datenschutzrechtlichen Bestimmungen.<sup>15</sup>



**Die Swiss Re Gruppe**, einer der weltgrößten Rückversicherer, nutzt Intel® SGX erfolgreich, um den **Schutz vertraulicher Daten** mehrerer Parteien mit einem Proof of Concept für Confidential Computing auszubauen.<sup>16</sup>



Die **University of California San Francisco** nutzte Intel® SGX zur Entwicklung einer datenschutzwahrenden Analytik, die die Entwicklung und Validierung von klinischen Algorithmen beschleunigt. Die Plattform wird eine „Zero Trust“-Umgebung bereitstellen, **um das geistige Eigentum an Algorithmen und die Vertraulichkeit von Gesundheitsdaten zu schützen**.<sup>17</sup>

## Welche Sicherheitsvorteile bietet Intel® SGX?

Intel® SGX bietet einen wichtigen Baustein für Confidential Computing, indem es den Zugang zu sensiblen Daten und sensiblem Code während deren Nutzung einschränkt. Dadurch werden diese vor Einsicht oder Manipulation durch andere Software geschützt. Nur Intel® SGX verfügt über die Flexibilität, virtualisierte Bare-Metal- und cloudnative Container-Bereitstellung zu unterstützen.



# Die nächste Generation integrierter Beschleuniger von Intel

Die integrierten Beschleuniger von Intel finden sich aktuell in den skalierbaren Intel® Xeon® Prozessoren – und die nächste Generation integrierter Beschleuniger ist schon in Sicht.

Die skalierbaren Intel® Xeon® Prozessoren der 4. Generation werden über Intel® Advanced Matrix Extensions (Intel® AMX), Intel® QuickAssist-Technik (Intel® QAT), Intel® Data Streaming Accelerator (Intel® DSA) sowie andere integrierte Technologien verfügen. Hier ist ein kurzer Einblick in das, was Sie erwarten können.

## Intel® AMX

Intel® AMX, unser DL Boost der nächsten Generation zur Steigerung der Deep-Learning-Leistung, verfügt über eine Reihe von Befehlen zur Matrixmultiplikation. Diese verbessern die Inferenz und das Trainieren von KI erheblich, mit einer bis zu 4,5-fachen INT8-Bildinferenz pro Sekunde im Vergleich zur vorherigen Generation.<sup>18</sup>

## Intel® QAT

Sie ist bereits jetzt Teil der skalierbaren Intel® Xeon® Prozessoren und bietet den Nutzer:innen eine noch schnellere Datenverschlüsselung und effizientere Datenkomprimierung für verschiedene Anwendungsbereiche – vom Netzwerk bis zu Unternehmensanwendungen, von der Cloud bis zu Storage und von der Bereitstellung von Inhalten bis zu Datenbanken.

## Intel® DSA

Intel® DSA ist ein leistungsstarker Beschleuniger, der entwickelt wurde, um Streaming- und Transformationsvorgänge wie sie bei Netzwerken, datenverarbeitungsintensiven Anwendungen und Hochleistungs-Storage gängig sind zu optimieren.

## Fazit

Intel kann auf eine lange Geschichte der Innovation und Integration zurückblicken, die uns in die einzigartige Position versetzt hat, diese neue Kategorie von integrierten Beschleunigern mit den skalierbaren Intel® Xeon® Prozessoren zu entwickeln. Diese speziellen integrierten Schaltkreise bringen den Kund:innen größeren geschäftlichen Nutzen bei verschiedenen Workloads.

Egal, ob die Leistung gesteigert, Nachhaltigkeitsinitiativen unterstützt oder die sensibelsten Daten geschützt werden sollen – Intel bietet eine Familie von integrierten Beschleunigern, die eine große Bandbreite an Lösungen ohne zusätzliche Hardware ermöglichen. Wie unsere eigenen internen Tests sowie reale Anwendungsfälle gezeigt haben, bieten diese Beschleuniger ein unerreichtes Preis-Leistungs-Verhältnis im Vergleich zu anderen CPUs auf dem Markt.

**Erfahren Sie mehr über die skalierbaren Intel® Xeon® Prozessoren und besuchen Sie <https://www.intel.de/xeonscalable>.**

<sup>1</sup>Siehe [123] unter <https://www.intel.com/3gen-xeon-config>

<sup>2</sup>Siehe [121] unter <https://www.intel.com/3gen-xeon-config>

<sup>3</sup>Siehe Seite 12 unter <https://www.intel.com/content/dam/www/public/us/en/documents/product-overviews/dl-boost-product-overview.pdf>

<sup>4</sup>Siehe <https://www.intel.de/content/www/de/de/customer-spotlight/stories/hyhy-customer-story.html>

<sup>5</sup>Siehe <https://www.intel.de/content/www/de/de/customer-spotlight/stories/kfbio-ai-customer-story.html>

<sup>6</sup>Siehe [118] unter <https://www.intel.com/3gen-xeon-config>

<sup>7</sup>Siehe <https://www.intel.com/content/www/us/en/customer-spotlight/stories/university-at-buffalo-customer-story.html>

<sup>8</sup>Siehe <https://www.intel.com/content/www/us/en/newsroom/news/broad-institute-intel-google-advance-biomedical-research.html>

<sup>9</sup>Siehe <https://www.intel.de/content/www/de/de/customer-spotlight/stories/cern-inference-customer-story.html>

<sup>10</sup>Siehe [104] unter <https://www.intel.com/3gen-xeon-config>

<sup>11</sup>Siehe [36] unter <https://www.intel.com/3gen-xeon-config>

<sup>12</sup>Siehe [38] unter <https://www.intel.com/3gen-xeon-config>

<sup>13</sup>Siehe [37] unter <https://www.intel.com/3gen-xeon-config>

<sup>14</sup>Siehe [39] unter <https://www.intel.com/3gen-xeon-config>

<sup>15</sup>Siehe <https://www.intel.de/content/www/de/de/customer-spotlight/stories/nationwide-building-society-customer-story.html>

<sup>16</sup>Siehe <https://www.intel.com/content/www/us/en/customer-spotlight/stories/swiss-re-customer-story.html>

<sup>17</sup>Siehe <https://www.intel.com/content/www/us/en/newsroom/news/ucsf-propel-medical-device-innovations.html>

<sup>18</sup>Siehe Session-Benchmark #41 und #42 unter <https://edc.intel.com/content/www/us/en/products/performance/benchmarks/vison-2022/>. Die Ergebnisse können von Fall zu Fall abweichen.

### Hinweise und Disclaimer:

Die Leistung variiert je nach Verwendung, Konfiguration und anderen Faktoren. Weitere Informationen finden Sie auf der [Performance-Index-Website](#).

Die Leistungsergebnisse basieren auf Tests, die an den in den Konfigurationen angegebenen Daten durchgeführt wurden, und berücksichtigen möglicherweise nicht alle öffentlich verfügbaren Sicherheitsupdates. Konfigurationsdetails finden Sie im Backup. Kein Produkt und keine Komponente bietet absolute Sicherheit.

Kosten und Ergebnisse können variieren.

Intel® Technik kann entsprechend geeignete Hardware, Software oder die Aktivierung von Diensten erfordern.

© Intel Corporation. Intel, das Intel Logo und andere Intel Markenbezeichnungen sind Marken der Intel Corporation oder ihrer Tochtergesellschaften. \*Andere Marken oder Produktnamen sind Eigentum der jeweiligen Inhaber.

Intel hat keinen Einfluss auf und keine Aufsicht über die Daten Dritter. Sie sollten andere Quellen heranziehen, um die Richtigkeit zu überprüfen.