

Fortschritte bei KI, Virtualisierung und Plattform-Sicherheit



Andreas Timm

Intel Director Sales
Central Europe, Intel



Hans Roth

Senior Vice President and
General Manager, EMEA, Red Hat

Lead sponsor



Open-source software: Intel is committed

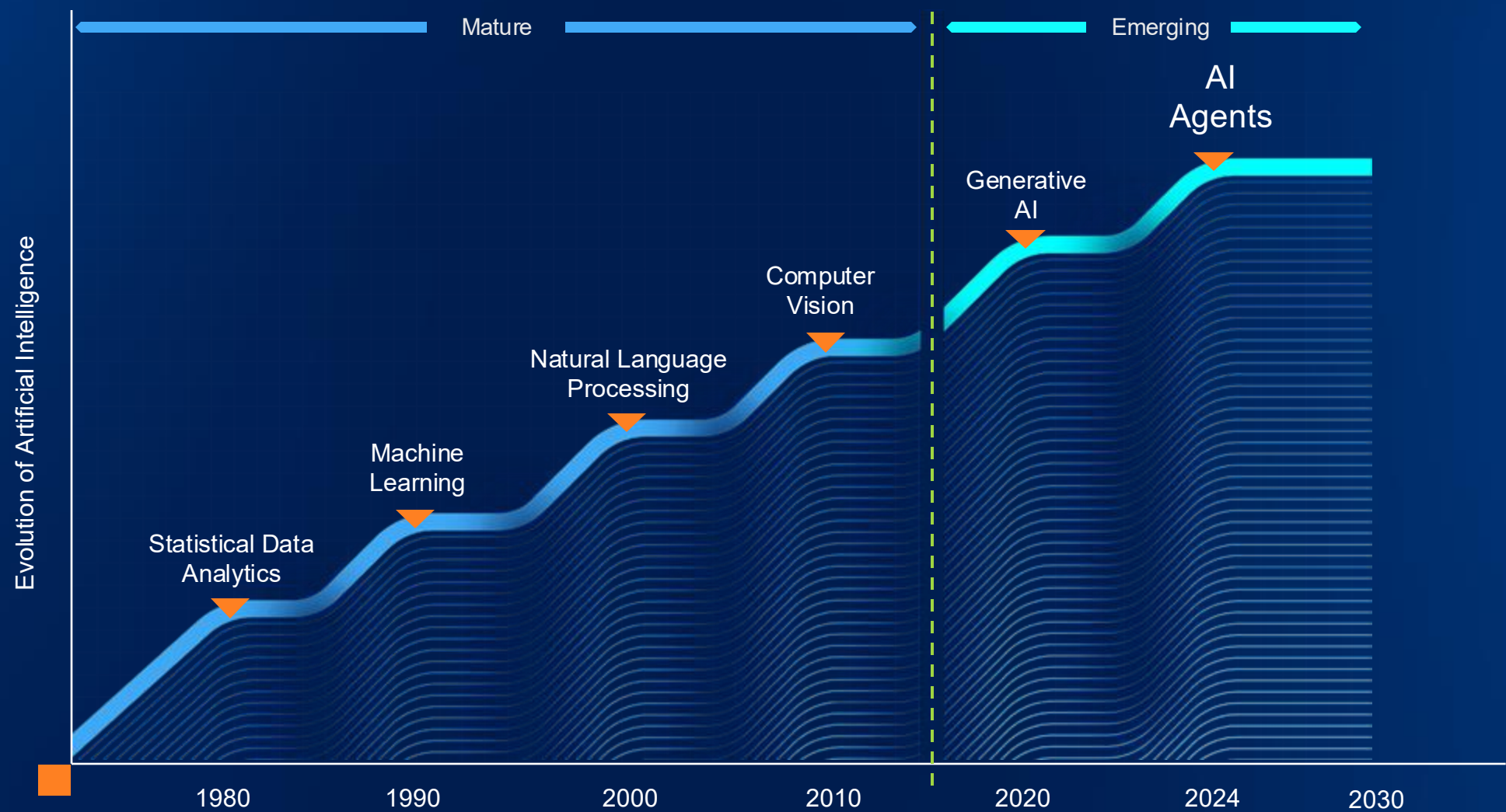
Intel® has a long history with Linux®, actively participating in open-source development and collaboration with the Linux community, to ensure hardware is well-supported and delivers optimal performance on Linux-based systems.

Intel contributes to more than 100 different open-source projects, from the Linux kernel to cloud orchestration and plugins for Kubernetes.



Intel is one of the key contributors to Linux
Linux Kernel Corporate Contributor since 2007¹

Evolution of AI Applications in Enterprise Use Cases



The AI Hierarchy: Mapping ML, DL, and GenAI with Intel

Discover how Intel® processors fuel AI workloads across inference, training, and next-generation GenAI applications.



30-year history of ML on CPUs

- Regression
- Classification
- Clustering
- Decision trees
- Data generation



Meet customer service-level agreements (SLAs) with general-purpose hardware

- Image processing
- Computer vision
- Natural language processing (NLP)
- Recommender systems



Inference <20B parameters



Inference and Fine-Tuning Small to large-sized LLMs

- Subset of AI that focuses on creating new, original content

AI

Machine learning

Deep learning

GenAI



Resolve Customer Queries Faster with More Concurrent Users in Your **LLMs** and **Agents**

■ Get superior performance for batch, real-time inference, and training for small and medium language models with Intel® Xeon® processors.

■ Use your CPU for cost-effective model updates.

Large language models (LLMs)

Large language models (LLMs)				
Intel Xeon 6 vs. AMD EPYC Turin	Intel Xeon 6 vs. 5th Gen Intel Xeon			5th Gen Intel Xeon vs. 3rd Gen Intel Xeon
Llama2-7B	GPTJ-6B	Llama-13B	Llama 2-7B	Llama2-13B
Up to 1.38x higher throughput	Up to 2x higher performance	Up to 2x higher performance	Up to 2.3x higher training performance	Up to 2.1x real-time inference performance speedup
with Intel Xeon 6980P vs. AMD EPYC 9965 ¹	Intel Xeon 6980P vs. Intel Xeon 8592+ ²	Intel Xeon 6980P vs. Intel Xeon 8592+ ²	Intel Xeon 6980P vs. Intel Xeon 8592+ ³	5th Gen Intel Xeon vs. 3rd Gen Intel Xeon ⁴

^{1,2,3,4} See [9A231], [9A10], [9A6], and [A2] at [intel.com/processorclaims](https://www.intel.com/processorclaims). Intel® Xeon® 6 and 5th Generation Intel® Xeon® Scalable Processors. Results may vary.

Intel® Gaudi® 3 AI Accelerator: AI Inferencing

Price Performance Advantage

Up to
43%

Higher throughput
(tokens per second)

on IBM Granite-3.1-8B-Instruct
vs. leading GPU competitor
with small context sizes

Up to
120%

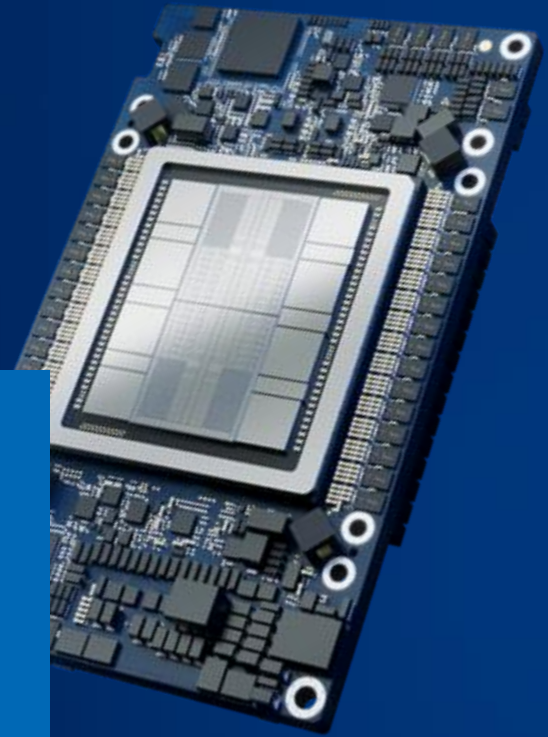
More cost efficient
(tokens per dollar)

on Mixtral-8x7B-Instruct-v0.1
vs. leading GPU competitor
with long input and short output sizes

Up to
92%

More cost efficient
(tokens per dollar)

on Llama-3.1-405B-Instruct-FP8
vs. leading GPU competitor
with large context sizes



Source: [Signal65 Lab Insight Whitepaper - Intel Gaudi 3 AI Accelerator at Scale on IBM Cloud](#), Intel-commissioned study by Signal65, published April 22, 2025. Reported numbers are inferencing results on Intel Gaudi 3 vs. NVIDIA H200. See source for workloads and configurations. Results may vary.



intel
xeon

Intel® Xeon® 6 Processor

1.9x

higher performance per watt at a typical
40% server utilization vs. prior
generation

Designed for
Efficiency

2.5x

higher HPC performance
vs. prior generation

Significant
Performance Leaps

5.5x

higher AI Inferencing performance
vs. AMD EPYC

Unmatched
Performance

Confidential Computing and Post Quantum Crypto for Information & Data Security

Intel® Software Guard Extensions (Intel® SGX)

Smallest Trust Boundary - Confidential data access is restricted to attested application code

Intel® Trust Domain Extensions (Intel® TDX)

Virtual machine isolation from cloud stack, admins, and other tenants

Post Quantum

Intel adds Quantum attack protection while providing 1.89 Tb IPsec throughput.

Performant Post-Quantum Cryptography (PQC) leveraging the Intel NetSec Accelerator and Arqit SKA-Platform™ for PQC.

Apply for a **free** Gaudi 3 Proof of Concept in **30 seconds**

Choose your GenAI or Virtualization PoC:

- ▶ Building Inference, RAG, AgenticAI, Model-as-a-Service, and other AI Use Cases with Intel Gaudi and Xeon
- ▶ Optimize fine tuning with intel Gaudi

Why work with Intel + Red Hat?:

- ▶ Benefit from access to free highly qualified experts from Red Hat and Intel and free access to the latest hardware to build your AI use case / application.

If selected, a Intel / Red Hat representative will contact you via email.



The Intel logo is centered on a solid dark blue background. It features the word "intel" in a white, lowercase, sans-serif font. A small, bright blue square is positioned above the first vertical stroke of the letter "i".

intel