HPE Tech Fusion 2025

Making data your differentiator with the power of Intel Inside on your Data Center

Fernando Salado
Industry Technical Sales Specialist

Hewlett Packard Enterprise

intel

Data Center Requirements Are Evolving

Varying uses require unique optimization vectors

Crecimiento de datos y necesidad de trasnformación

Los datos empresariales crecen a una tasa compuesta anual del 30% entre 2023 y 2028.

Los datos son generados y almacenados en **entornos híbridos**.

IA como motor de transformación

IDC posiciona la IA (incluida la generativa) como prioridad estratégica en Argentina y LATAM, con beneficios en productividad, decisiones más rápidas y rentabilidad

Seguridad

IDC resalta la automatización de la seguridad como clave ante la complejidad creciente y los costos de intrusiones

Data Center Requirements Are Evolving

Varying uses require unique optimization vectors

AI Everywhere

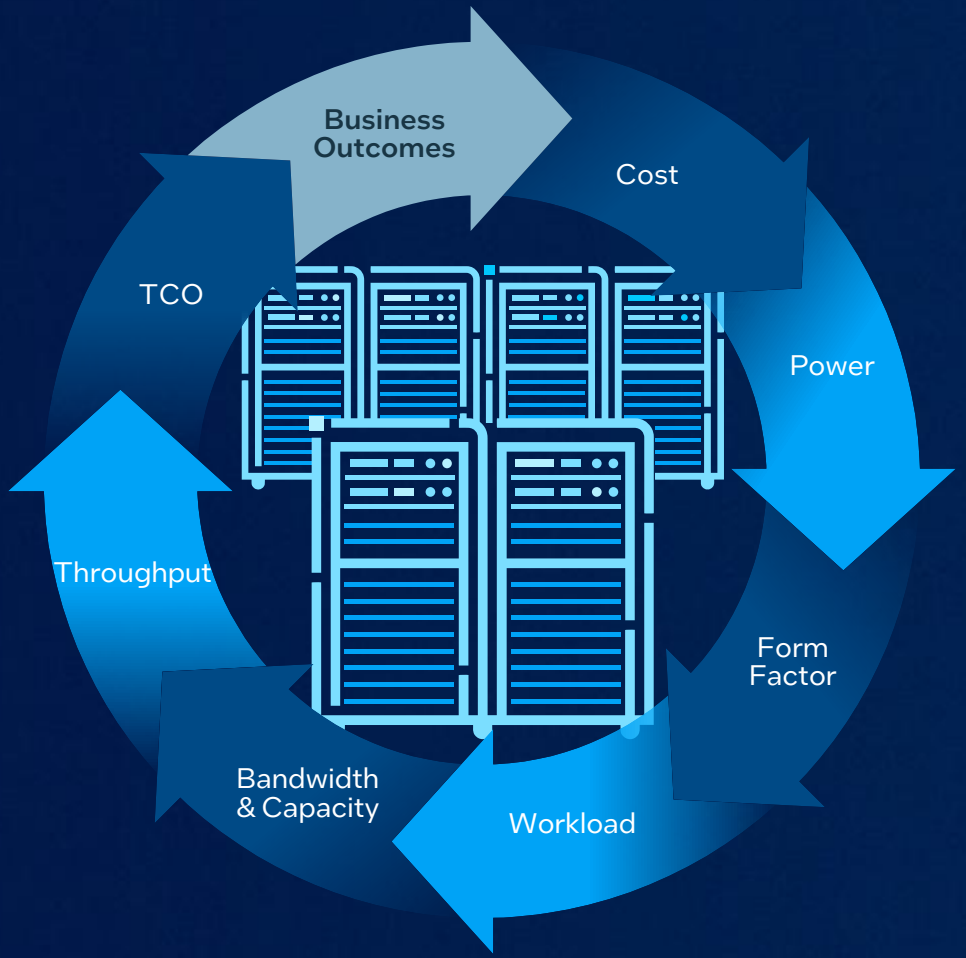
Enable performance at scale with accelerated hardware and open standards-based software

Throughput & Latency

Maximize system-level performance with best-in-class response time

Security, Quality & Reliability

Provide security, quality, and reliability for at-scale deployments



Efficiency & TCO

Increase rack density while meeting power efficiency requirements to improve TCO (total cost of ownership)

Sustainability

Minimizing carbon emissions through improved energy efficiency and circular product design

Software Compatibility

ISA consistency for software ecosystem compatibility

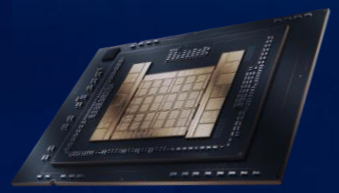
Addressing Tomorrow's Computing Needs

Across data center, network & edge

Intel® Xeon® 6 Processors

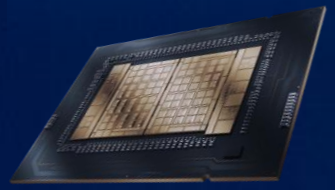
Maximum Efficiency with E-cores

Performance-per-watt for high-density compute and scale-out workloads



Peak Performance with P-cores

Per-core-performance for compute-intensive workloads



Built for Networking and Edge

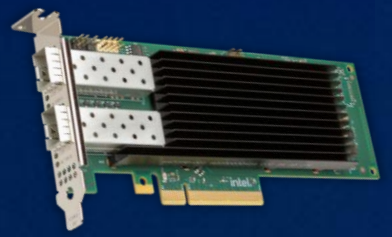
More power-efficient, servers with Intel® vRAN Boost and media acceleration, with networking built in

Performance Boost for Small Business

Affordable performance for business-critical services

Intel Ethernet Controllers & Adapters

Standards-based, high-performance networking



HPE ProLiant Compute Gen12

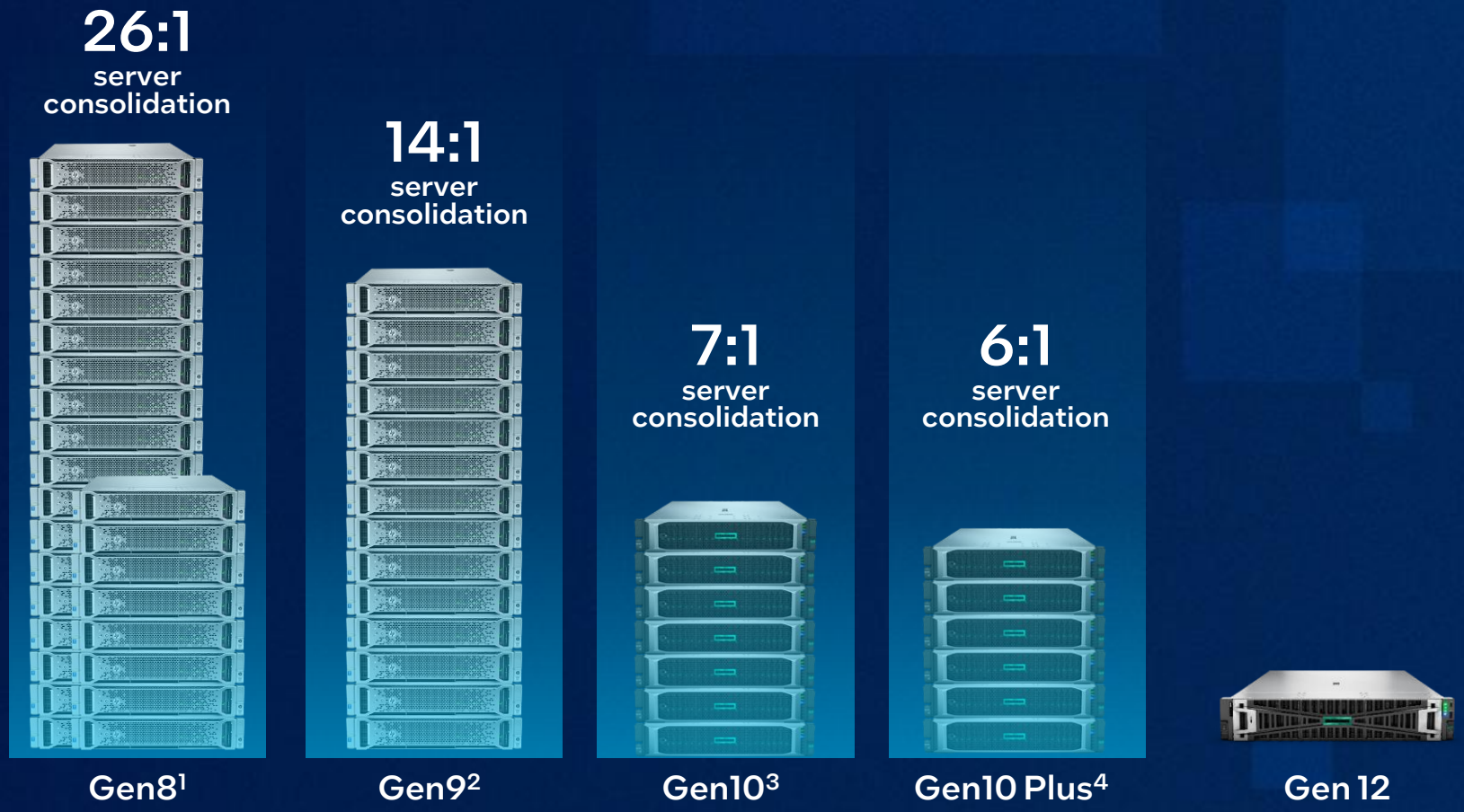
The most compelling case for server refresh yet

41% higher performance per watt versus Gen10[6]

Free up data center capacity with up to 26:1 consolidation[2]

Reduce energy consumption and achieve up to 84% power savings[1]

Mitigating risk and technical debt from dated operating system and software

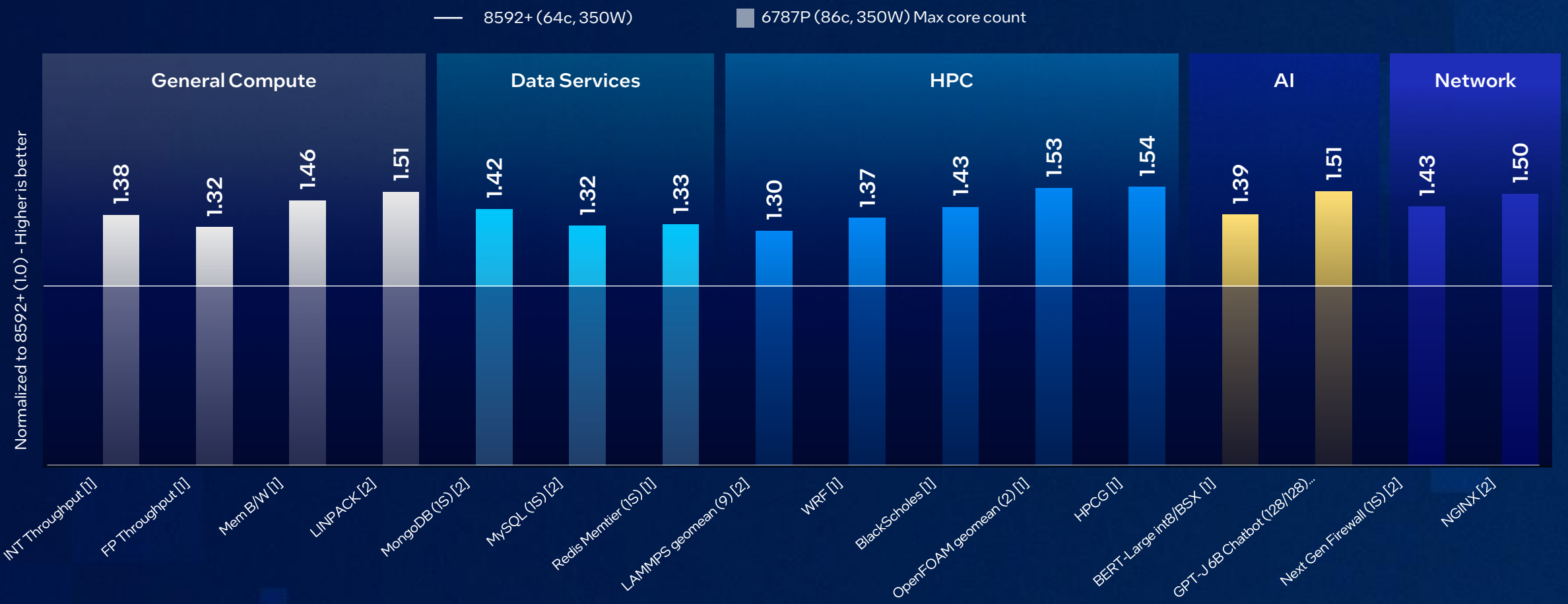


	Gen8[1]	Gen9[2]	Gen10[3]	Gen10 Plus[4]
Power savings per year (up to)	84%	61%	65%	50%

SPEC and the names SPECrate are registered trademarks of the Standard Performance Evaluation Corporation (SPEC). The stated results [SPECrate2017_int_base: #36693 (1), #36691 (2), #20893 (3), #37007 (4)] are published as of 01-01-2025, see spec.org, and compared against a 48-core estimated Gen12 system. All rights reserved. Power savings based on the Thermal Design Power of the systems. (6)The performance per watt advantages are based on internal power and performance measurements on similar configured high energy efficient servers and compared against an estimated 86-core Gen12 system.
Source: HPE. Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Delivering Significant Gains Across Broad Workloads

All within similar or lower power envelope



[1] MRDIMM [2] DDR. See [7G20, 7D21-7D23, 7H21-7H25, 7A21, 7A28, 7N20, 7N25] intel.com/processorclaims: Intel Xeon 6. Results may vary.
This offering is not approved or endorsed by OpenCFD Limited, producer and distributor of the OpenFOAM software via www.openfoam.com, and owner of the OPENFOAM® and OpenCFD® trademark.

Save Power and Money on New Server Purchases

Hewlett Packard Enterprise

intel

Performance advantage and TCO savings vs AMD EPYC 9005 servers

Intel Xeon 6900P Processor-based Servers

Recommendation System
DLRM

1.87x Perf / Server

Computational Fluid Dynamics
OpenFOAM

1.43x Perf / Server

Intel Xeon 6700P Processor-based Servers

Web Services
NGINX TLS (1S) on 6760P

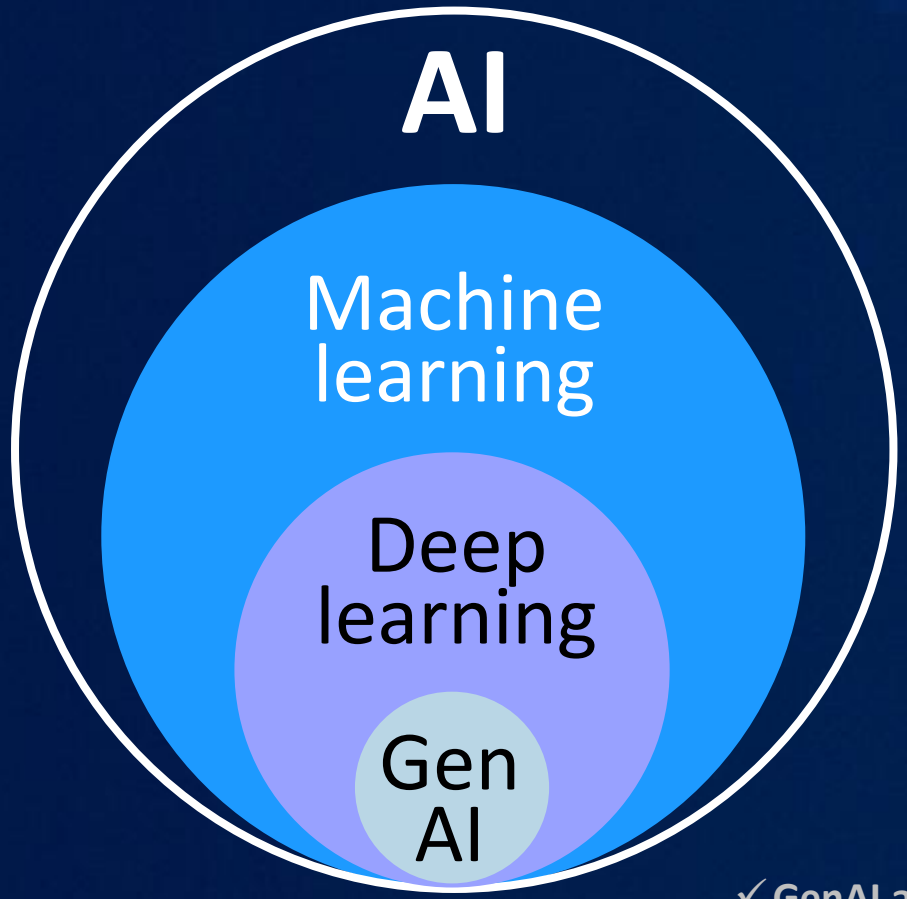
1.55x Perf / Server

Image Construction
Vision Transformer on 6760P

2.09x Perf / Server

*Estimated over 4 years. See [9T223, 9T222, 7T223, 7T221] intel.com/processorclaims: Intel Xeon 6. Results may vary.
This offering is not approved or endorsed by OpenCFD Limited, producer and distributor of the OpenFOAM software via www.openfoam.com, and owner of the OPENFOAM® and OpenCFD® trademark.

Many Types of “Intelligence”



Regression

Classification

Clustering

Decision Trees

Data Generation

- ✓ Practical to reverse engineer
- ✓ Tabular/limited dataset
- ✓ Good enough accuracy
- ✓ Fully-explainable

Image Processing

Natural Language Processing

Recommender Systems

- ✓ Difficult problem to reverse engineer
- ✓ Large, uniform dataset
- ✓ Highest accuracy

Subset of AI that focuses on creating new, original content

- ✓ **GenAI algorithms** use advanced techniques like deep learning and neural networks to produce realistic and coherent outputs

Intel® Xeon® 6 Processors for AI

Hewlett Packard Enterprise

intel

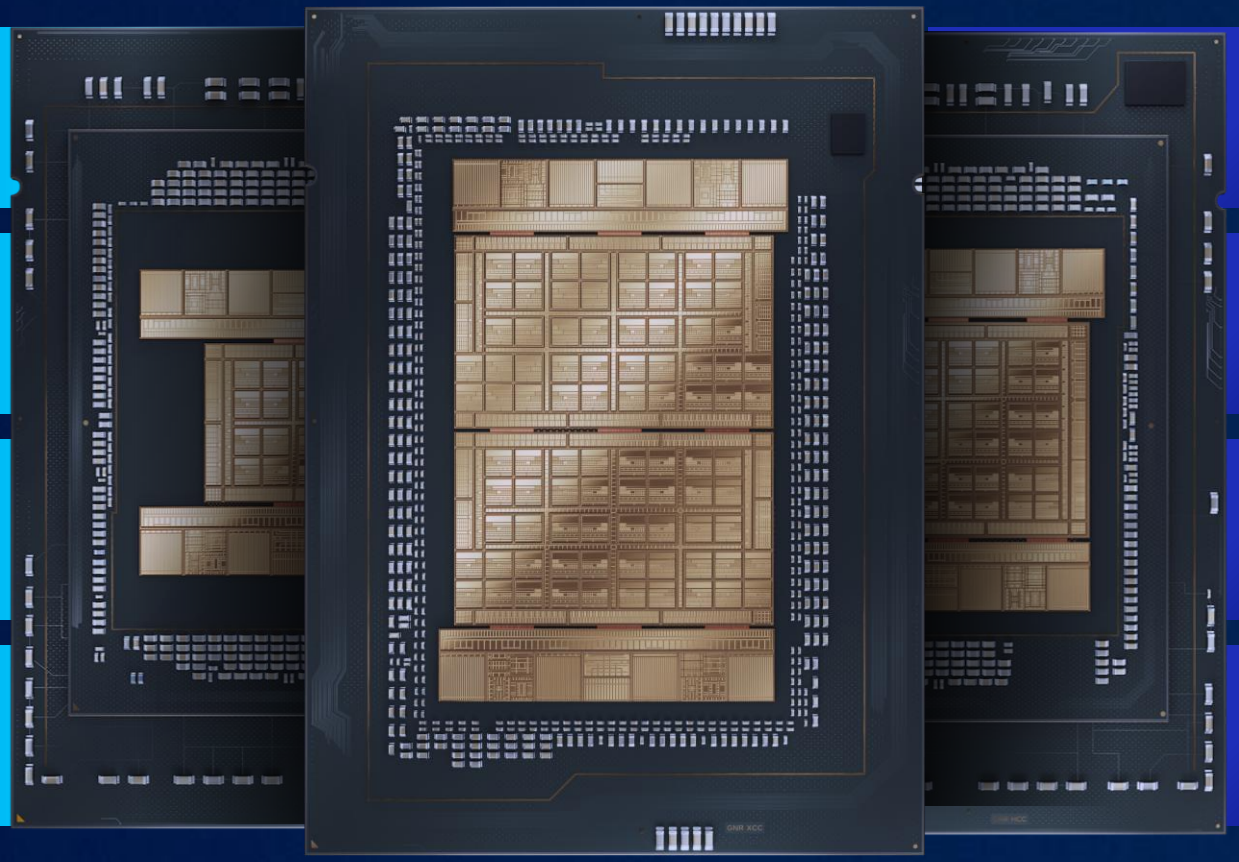
World's Best CPU for AI

Up to 128 P-cores
on 6900-series
up to 86 P-cores on 6500/6700-series

More bandwidth & cache
MRDIMM memory support
Up to 504MB low latency LLC

AI accelerators built-in
Intel® AMX, Intel® AVX-512,
and Intel® AVX-2

Comprehensive SW suite
AI development across classical ML
and small GenAI models



The Most Deployed Host CPU

Superior I/O performance
up to 192 PCIe 5.0 lanes

High Single Threaded Performance
With Intel's latest generation P-core

Top Tier Memory Support
30% higher memory B/W with MRDIMMs
Expandability with CXL 2.0

Ready for Deployment
DC-MHS & NVIDIA MGX™
form factors supported

Best CPU: See ISC 2024 section of intel.com/performanceindex for workloads and configurations. Your results may vary. Intel technologies may require enabled hardware, software or service activation.
Most Deployed Source: IDC Server Tracker report, based on 1H'24 system volume.
*NVIDIA logo and MGX are trademarks of NVIDIA and/or its subsidiaries

Elevating Xeon for AI - A two prong approach

The best CPU for General AI

ML workloads, Small & Mid Model
Inference (<20B) & Small Model Fine-tuning

- P-core perf/thread & core count
- Built-in AI Accelerators (AMX)
- High Bandwidth & Cache
- SW Enablement

- **1.38x** Perf Llama2-7B chat-hf
- **1.53x** Perf on ResNet-50
- **1.27x** Perf on BERT Large
- **1.21x** Perf on DLRM

2S AMD EPYC 9755 (128 cores) vs. 2S Intel Xeon 6787P (86 cores)
Llama2-7B is 2S AMD EPYC 9965 (192 cores) vs. 2S Intel Xeon 6980P (128 cores)

Most deployed Host CPU for Large Scale AI

Foundational LLM Training, Model Fine-tuning,
Large Model Inference

- **Superior I/O Performance**
- **Single Thread Performance**
- **Core Clock Frequency**
- **Top Tier Memory Support**

- **20% more*** PCIe lanes gen on gen
- **High Per Core Performance**
- **High** Clock and Turbo Frequency
- **30% Higher^** bandwidth (MRDIMM)

See [7A220 – 7A224] intel.com/processorclaims: Intel Xeon 6. Results may vary. * Xeon 6: 96 PCIe gen5 , 5th Gen Xeon: 80 PCIe gen5. ^See configuration [13d] in backup

HPE ProLiant Compute Gen12

Hewlett Packard Enterprise

intel

Available now

AI optimized

4U, 2P
HPE ProLiant Compute DL380a Gen12
Ten NVIDIA RTX PRO™ 6000 or H200 NVL
Intel Xeon 6 processor

2U,2P
HPE ProLiant Compute DL384 Gen12
NVIDIA GH200 NVL2
Intel Xeon 6 processor

TCO/Power optimized

NEW!

1U, 1P
HPE ProLiant Compute DL320 Gen12
Intel Xeon 6 processor

NEW!

2U,1P
HPE ProLiant Compute DL340 Gen12
Intel Xeon 6 processor

TCO/Power optimized

NEW!

1U, 2P
HPE ProLiant Compute DL360 Gen12
Intel Xeon 6 processor

NEW!

2U, 2P
HPE ProLiant Compute DL380 Gen12
Intel Xeon 6 processor

SMB/Edge optimized

NEW!

2P Tower
HPE ProLiant Compute ML350 Gen12
Intel Xeon 6 processor

Available Summer 2025

Blade optimized

HPE Synergy 480 Gen12
Intel Xeon 6 processor

Big Data optimized

4U, 4P
HPE ProLiant Compute DL580 Gen12
Intel Xeon 6 processor

And more to come!

Hewlett Packard Enterprise

intel®

Fernado Salado

Intel Xeon-based HPE ProLiant systems supporting HPE Morpheus VM Essentials

Hewlett Packard Enterprise

intel

Platform type

Intel 1 RU Server	HPE ProLiant DL320 Gen 12	HPE ProLiant DL360 Gen 10+	HPE ProLiant DL360 Gen 11	HPE ProLiant DL360 Gen 12	
Intel 2 RU Server	HPE ProLiant DL340 Gen12	HPE ProLiant DL380 Gen 10+	HPE ProLiant DL380 Gen 11	HPE ProLiant DL380a Gen 11	HPE ProLiant DL380 Gen 12
Intel Single Slot Blad	HPE Synergy 480 Gen 11 (local Storage, iSCSI, FC)				

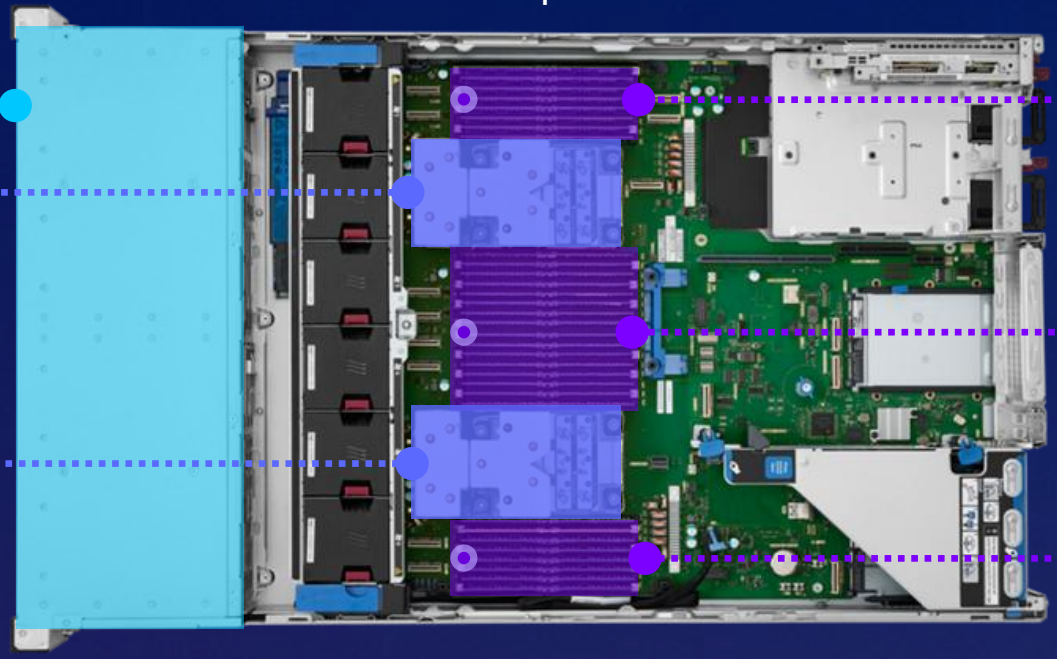
HPE ProLiant Gen 12 with Intel Xeon 6

System Performance Measurements and Configuration

HPE ProLiant Compute DL380 Gen 12

HPE Morpheus VM Essentials Software

Intel Xeon 6 Processors



16 x 128GB DDR5 6400MT/s DIMMs per server

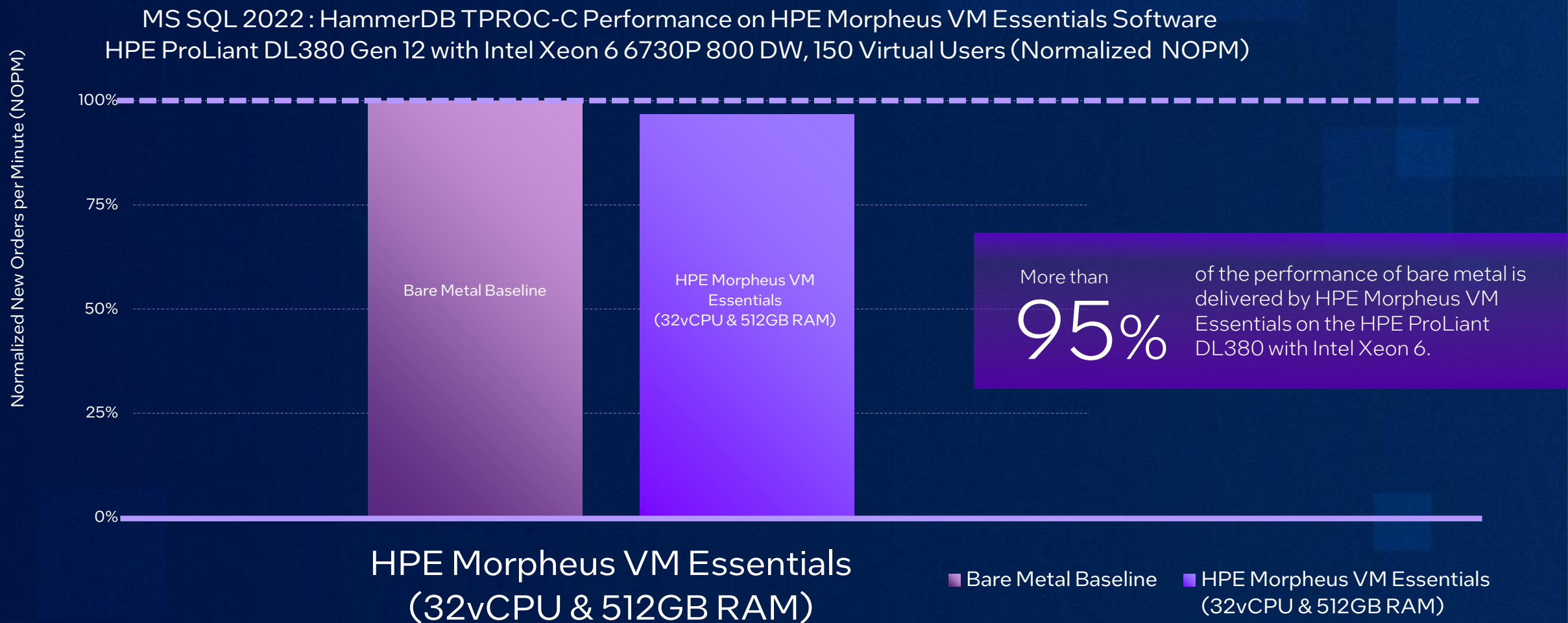
MS SQL 2022: HammerDB TPROC-C Workload

- Metric: Performance on VM Essentials vs. Bare Metal
- System: HPE ProLiant DL380 Gen 12, 2x Intel Xeon 6 6730P (32c)
- Parameters: 800 DW, 150 Virtual Users (Normalized NOPM)

GenAI Inference Serving with vLLM

- Metric: Number of Concurrent Prompts
- System: HPE ProLiant DL380 Gen 12, 1x Intel Xeon 6 6767P (64c)
- Workload: Llama 3.1 8B Instruct Inference Serving

HPE Morpheus VM Essentials Software Performance

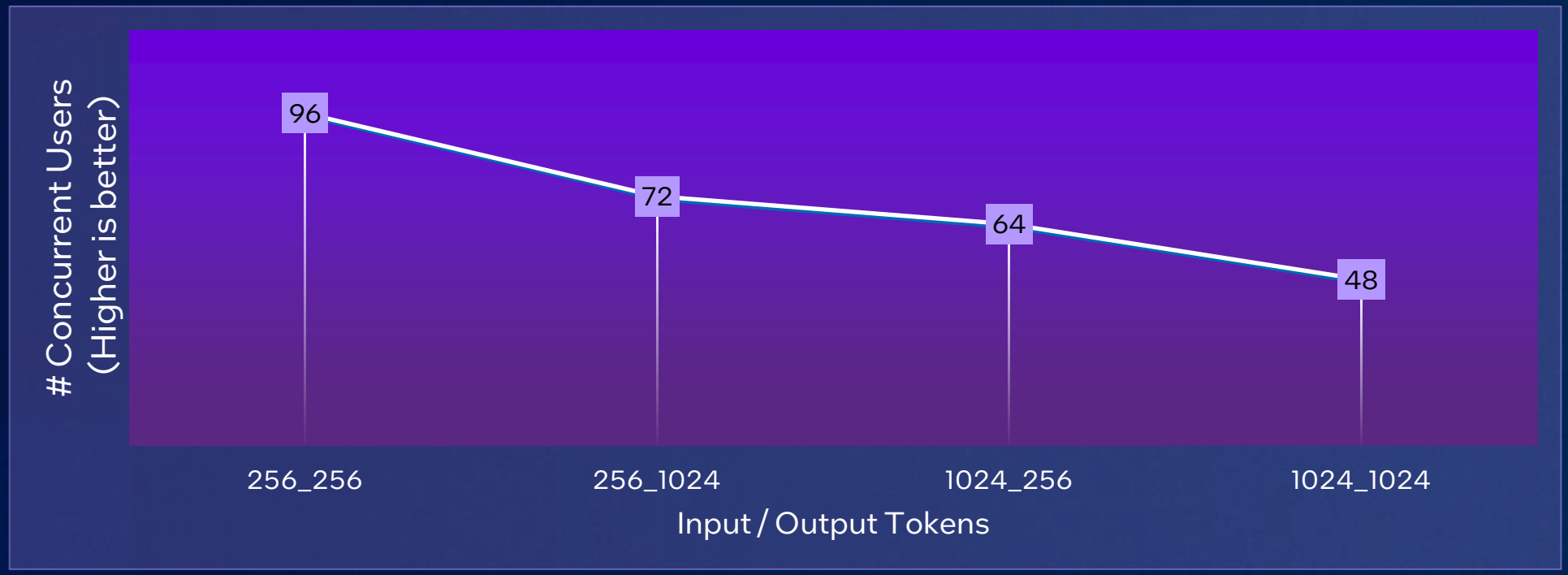


HPE Morpheus VM Essentials Software Performance

vLLM based on Llama 3.1 8B model hosted on HPE Morpheus VM Essentials

vLLM Serving on VM Essentials: Conversational Chatbot with LLAMA3.1-8B-Instruct
HPE ProLiant DL380 Gen 12 with Intel Xeon 6 1 x 6767P

Maximizing CONCURRENT USERS at <150MS TPOT SLA



HPE ProLiant DL380 with Xeon 6 1x6767P delivers max of 96 concurrent prompts
with vLLM Llama3.1-8B-Instruct Inference Serving on HPE VME virtualized stack

Intel® Xeon® 6 | 1-socket platform with 136 PCIe lanes

Delivers enhanced performance and lower TCO for today and tomorrow's data center demands

Increased PCIe lanes

Up to 136 lanes for greater lane volume for peripherals & devices

Improved I/O performance

Removal of power and latency penalties in a single socket

Enhanced performance and power efficiency

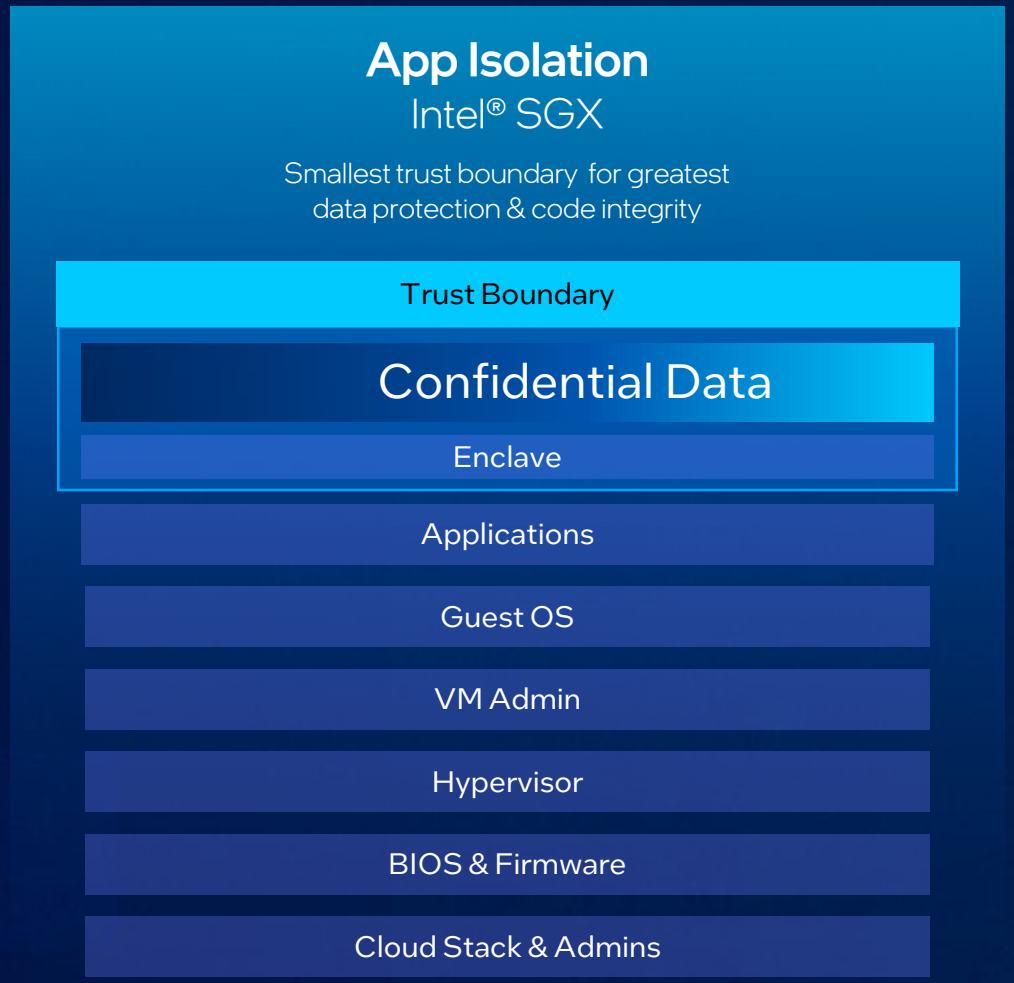
Greater core density, MRDIMMs and Intel® Accelerators

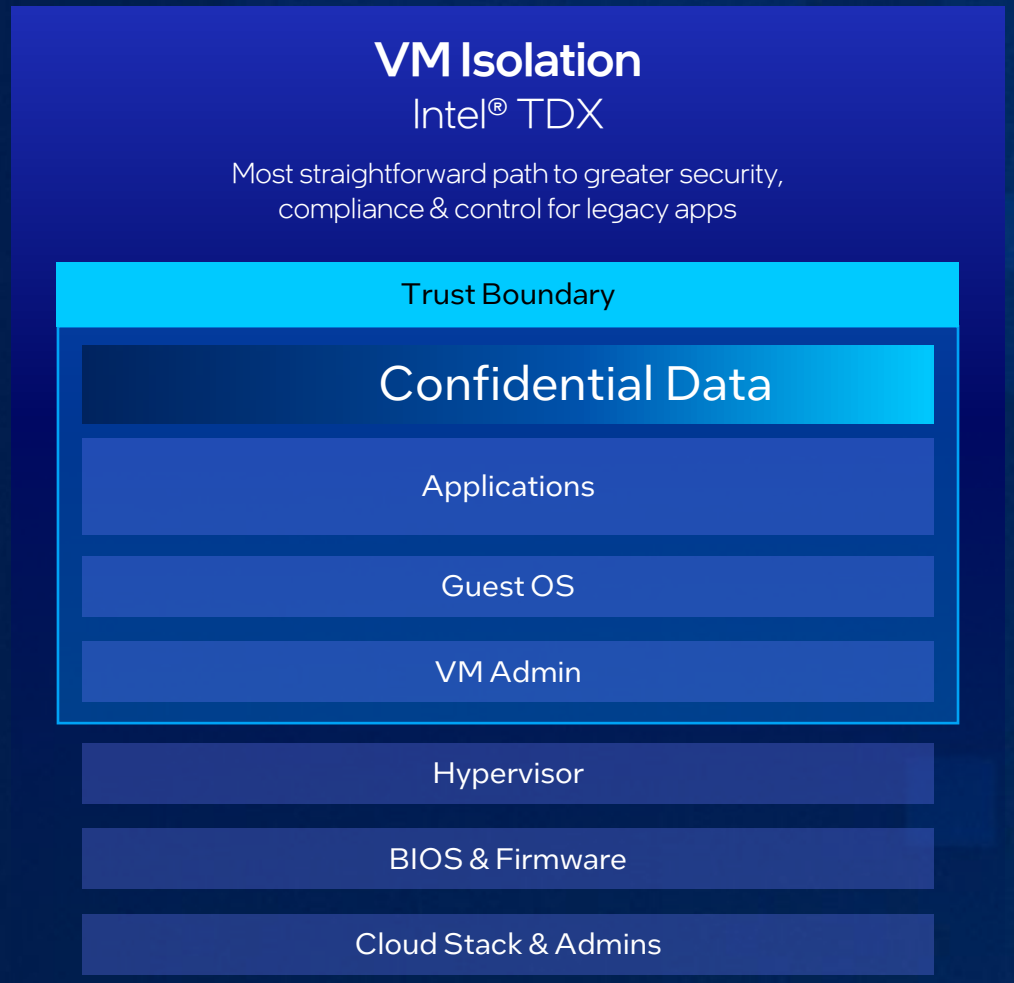
Cost optimization

Opportunity for server consolidation for improved total cost of ownership

The Most Comprehensive Portfolio

for confidential computing





Intel® Software Guard Extensions (Intel® SGX), Intel® Trust Domain Extensions (Intel® TDX)