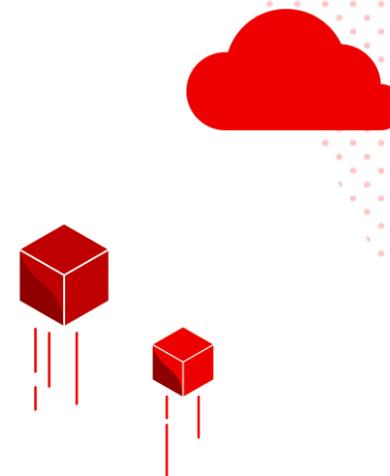




Agustín Gallego

Director de Administración Pública y Mercado Empresarial,
Intel



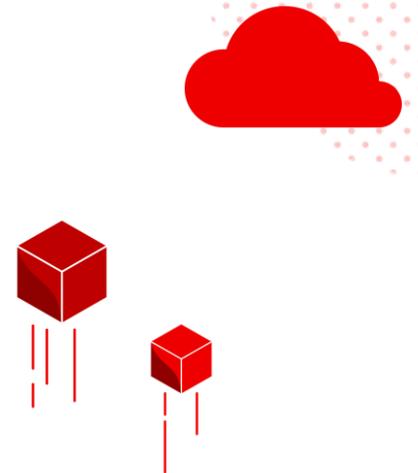


AI Everywhere

Revolutionizing AI in the Hybrid Cloud

Agustín Gallego

Director de Administraciones Públicas y
Mercado Empresarial





Over **25** Years of Relationship

Fun fact:

Intel is #1

Linux Kernel

Corporate Contributor

since 2007¹



Llevar la IA a todas partes



\$

IA está transformando nuestro trabajo y nuestra vida cotidiana

Del reconocimiento facial al aprendizaje personalizado. La IA está aquí para mejorar la vida de todos los habitantes del planeta

Data Encryption
Facial Recognition
Personalized Learning
AI Based Rendering
Video Conference

Purchase Recommendation
Code Generation
Robotics Vision
Inventory Management
Recommendation Systems

IA está evolucionando muy rápidamente

Tecnologías de datos subyacentes...



Graph Databases



Data Lakehouse



Data Fabrics



Synthetic Data

58%

de los CEOs de grandes empresas públicas están invirtiendo activamente en IA²

Más de 75%

De los datos gestionados por las empresas se crearán y procesarán fuera del centro de datos o la nube para el 2025³

\$300B

El gasto mundial en GenAI superará los \$300B en 2026¹

50%

de los despliegues en el Edge incluirán IA en 2026⁵

80%

de los PCs van a ser PCs IA en 2028⁶
de las empresas usarán Gen AI en 2026⁷

IA tan disruptiva como Internet

Se prevé que la **Generative AI** añada \$4.4T de valor a la economía mundial en 2040⁴

AI inferencing aumenta los costes de computación y supera el ritmo de la Ley de Moore

Crecimiento de modelos de gran tamaño (**large model sizes**) con más de 1T+ de parámetros

Crecimiento de modelos más pequeños y ágiles (**smaller, nimbler models**) con ~10B de parámetros

1. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#key-insights>

2. <https://chiefexecutive.net/the-rise-of-the-ai-ceo/>

3. Gartner®, Hyperscalers Stretching to the Digital Edge, July 2023. GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All right reserved.

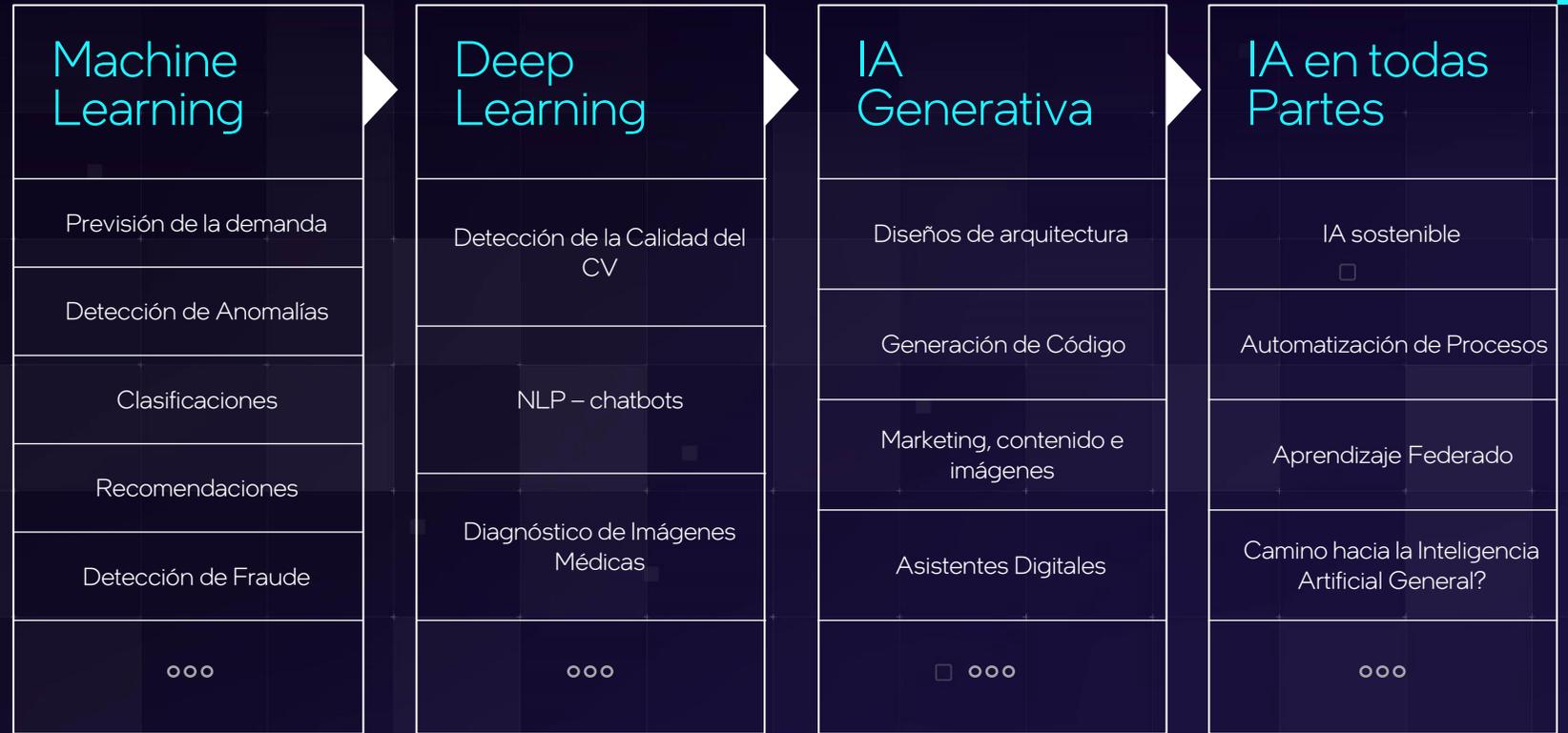
4. [Worldwide Artificial Intelligence Spending Guide \(IDC\)](#)

5. Gartner®, Building an Edge Computing Strategy, Thomas Bittman, 12 April 2023. GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All right reserved.

6. Source: Boston Consulting Group

7. Gartner [news release](#) – Oct. 10, 2023.

El rápido crecimiento de la IA



¿Qué necesita la IA?

Datos, computación, red, memoria y algoritmos

Velocidad de entrenamiento y ajuste fino en clústeres Informáticos grandes y ágiles (de Semanas a días a horas)

Despliegue responsable e inferencia en cualquier lugar y en todos los dispositivos (milisegundos)

¿Por qué es un reto la IA?

Complejidad

Rápido crecimiento del número de métodos, capacidades, tipos y tamaños de datos, así como los requisitos de infraestructura para ejecutar la IA

Costes

Aumento de los costes debido a la mayor demanda de computación a medida que la IA se adopta y consume de forma generalizada

Puesta en práctica

Se necesitan muchos pasos y habilidades para llevar la IA de la prueba de concepto a la producción en un proceso escalable y sostenible.

Seguridad y privacidad de los datos

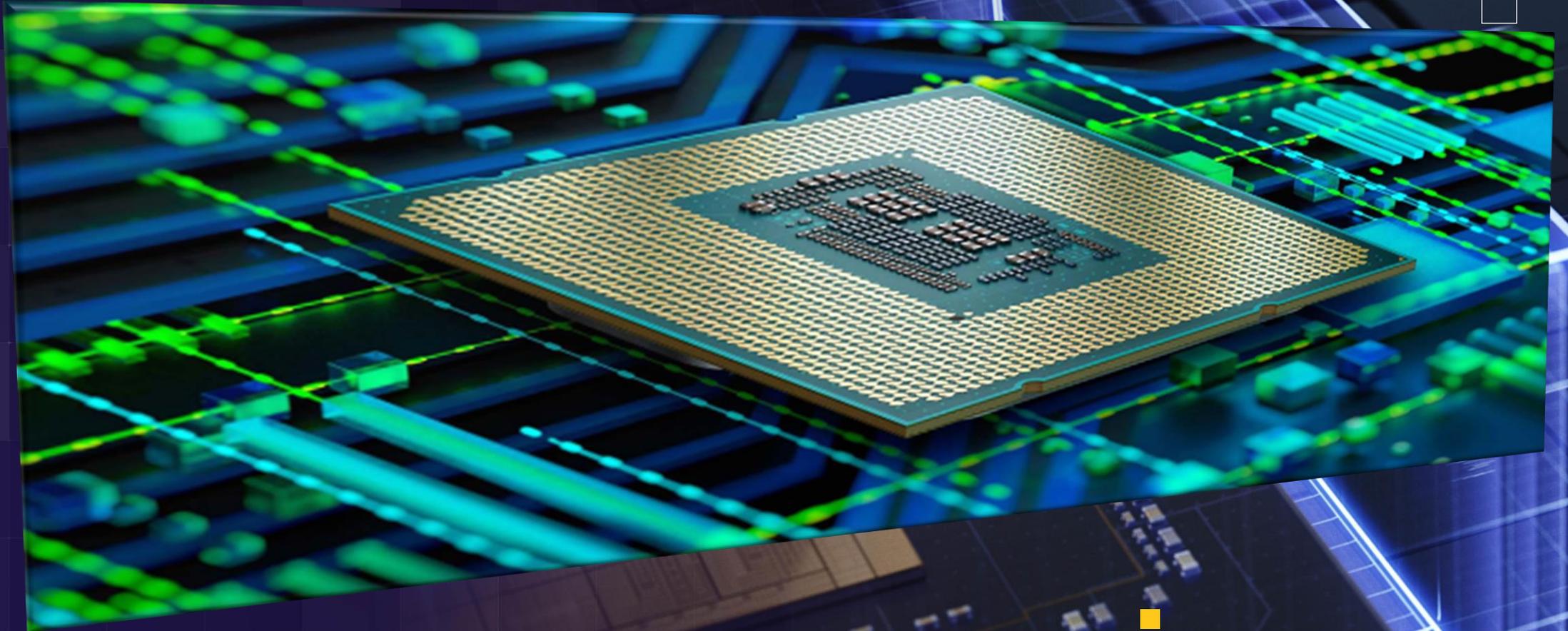
Activación de datos sensibles o regulados a escala mundial sin dejar de ser seguros y conformes a las normas

Impacto Humano

Garantizar que la tecnología de IA avance de forma responsable, ética y equitativa con un enfoque integral que reduzca los riesgos, mejore las vidas y optimice los beneficios.

Bringing AI everywhere

Expansión de la Inteligencia Artificial



Los semiconductores es la tecnología subyacente para acelerar la IA



Red Hat

intel.

intel.

Siliconomy

Una Economía en evolución impulsada por la magia del silicio, en la que los semiconductores son esenciales para mantener y posibilitar las economías modernas



\$1 trillion
semiconductor industry by 2030

\$8 Trillion
Global Tech Economy

Siliconomy: Algunos datos de referencia

Artificial intelligence, real demand

Demand for processing power could grow the market by \$500 billion in the next 10 years, largely due to demand for AI-related chips.

7%

percentage of semiconductor demand driven by AI-related chips in 2017

20%

expected percentage of semiconductor demand driven by AI-related chips by 2025

Shifting into overdrive

Cars are more computers than computers. Today's modern car requires as many as 3,000 semiconductors. And that could double with the advent of electric and autonomous vehicles.



1970s sports car

8 semiconductors



Average non-electric car

1,000 semiconductors

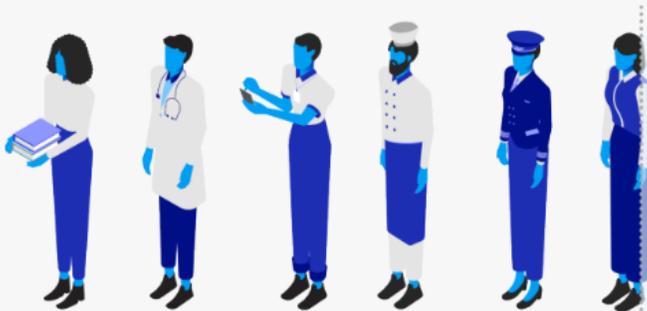
1 =

Siliconomy job



5.7

jobs supported by the siliconomy



Today's modern car

3,000 semiconductors



Electric car

5,000 – 7,000 semiconductors

Los semiconductores emprenden una ruta global de producción



Red Hat

intel.

Una cadena de suministro global

La fabricación de obleas es un proceso disperso por todo el mundo, con países implicados directamente en la cadena de suministro y en la de apoyo (incluidos equipos, herramientas de grabado y limpieza)

1 chip ... countries

BBC NEWS MUNDO

Qué son el galio y el germanio, los metales con los que China contraataca en su guerra de los microchips contra Occidente



Estados Unidos y China están enfrascados en una guerra de recursos y tecnología.



Bolsa, mercados y cotizaciones

El metal clave para Occidente que se dispara un 100% por el 'bloqueo' de China: el problema del Galio

- China controla el 98% de la producción mundial y limita las exportaciones
- El dominio del mercado impide entrar a otros 'jugadores' y las firmas piden precios 'públicos'
- El galio se utiliza para semiconductores, smartphones, placas solares... etc



LBX

HAZ TU MUNDO
EXTRAORDINARIO
HÍBRIDO



OREGON



ARIZONA



NUEVO MEXICO



OHIO



COSTA RICA



IRLANDA



ALEMANIA



POLONIA



ISRAEL



CHENGDU



VIETNAM



MALASIA

Home / Tech / Intel to make world's largest AI chip factory in the US, spend over \$100 billion across four states

Intel to make world's largest AI chip factory in the US, spend over \$100 billion across four states

FP Staff • March 21, 2024, 17:23:01 IST



Intel CEO Pat Gelsinger wants to build the largest AI chip manufacturing site in the world near Columbus, Ohio. The AI chip factory will be the centrepiece of Intel's ambitious five-year spending plan to expand in the US



Intel wants to build the world's largest AI chip making factory in Ohio, and plans to have it operational by 2027. Image Credit: Intel



Un alto coste de fabricación conduce a una consolidación de la Industria

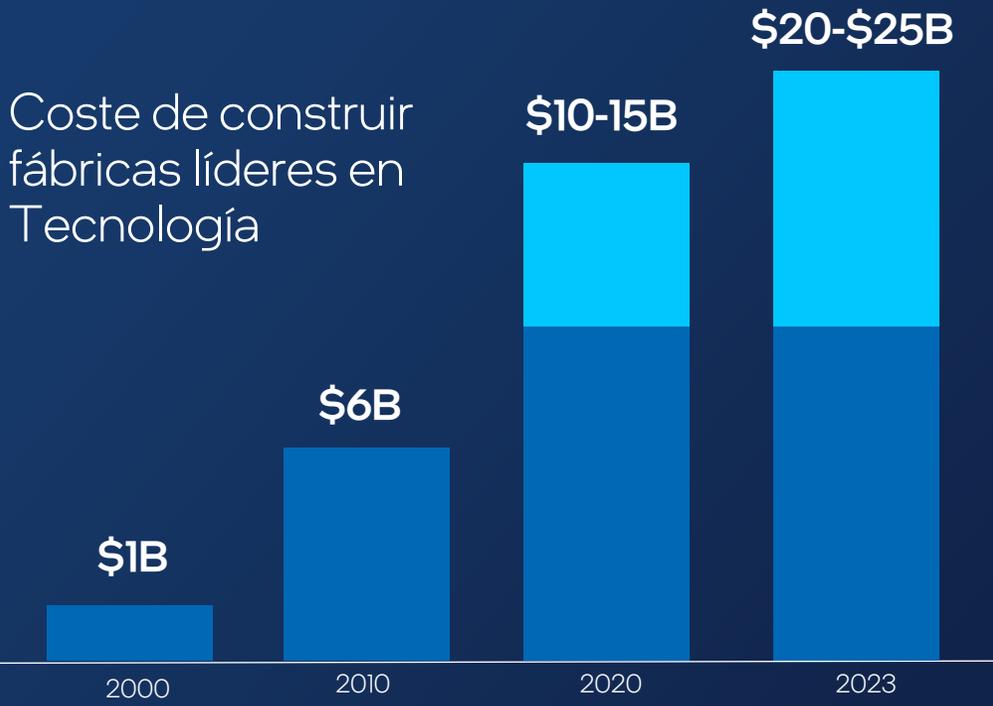
Sharp
Cypress
Sony
Infineon
TI
Toshiba
Freescall
SMIC
Renesas
Fujitsu
Panasonic
UMC
ST-M
IBM
AMD
Samsung
TSMC
Intel
2004-06 (90nm)

Panasonic
UMC
ST-M
IBM
GloFo
Samsung
TSMC
Intel
2010-12 (32nm/28nm)

IBM
GloFo
Samsung
TSMC
Intel
2012-14 (22nm/20nm/14nm)

2017-20 (10nm/7nm)

Coste de construir fábricas líderes en Tecnología



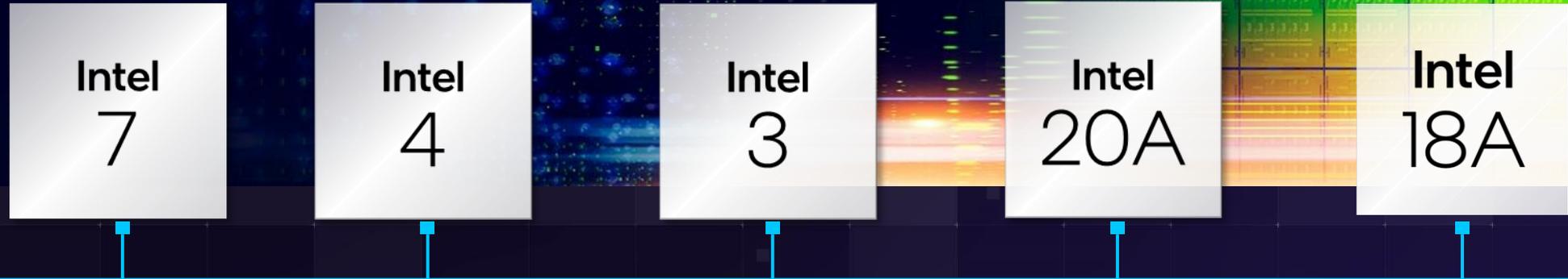
INVERSIONES EN EQUIPOS DE FABRICACIÓN

ASML envía a Intel el equipo de fotolitografía más avanzado del mundo: es un mazazo para las ambiciones de China

- El fabricante de chips estadounidense será el primer cliente de ASML en recibir un Twinscan EXE:5000
- Se trata de una máquina de litografía UVE de alta apertura de última generación
- Las sanciones comerciales impiden a China acceder a esta tecnología de vanguardia

El camino hacia el liderazgo tecnológico

Creación de un marco coherente y una visión más precisa de los nodos de proceso. Un modelo de desarrollo más rápido permite la Innovación en la ejecución y 5 nodos en 4 años



RibbonFET

Nuevas arquitecturas de puerta de los transistores

PowerVia

Nueva red trasera de suministro de energía



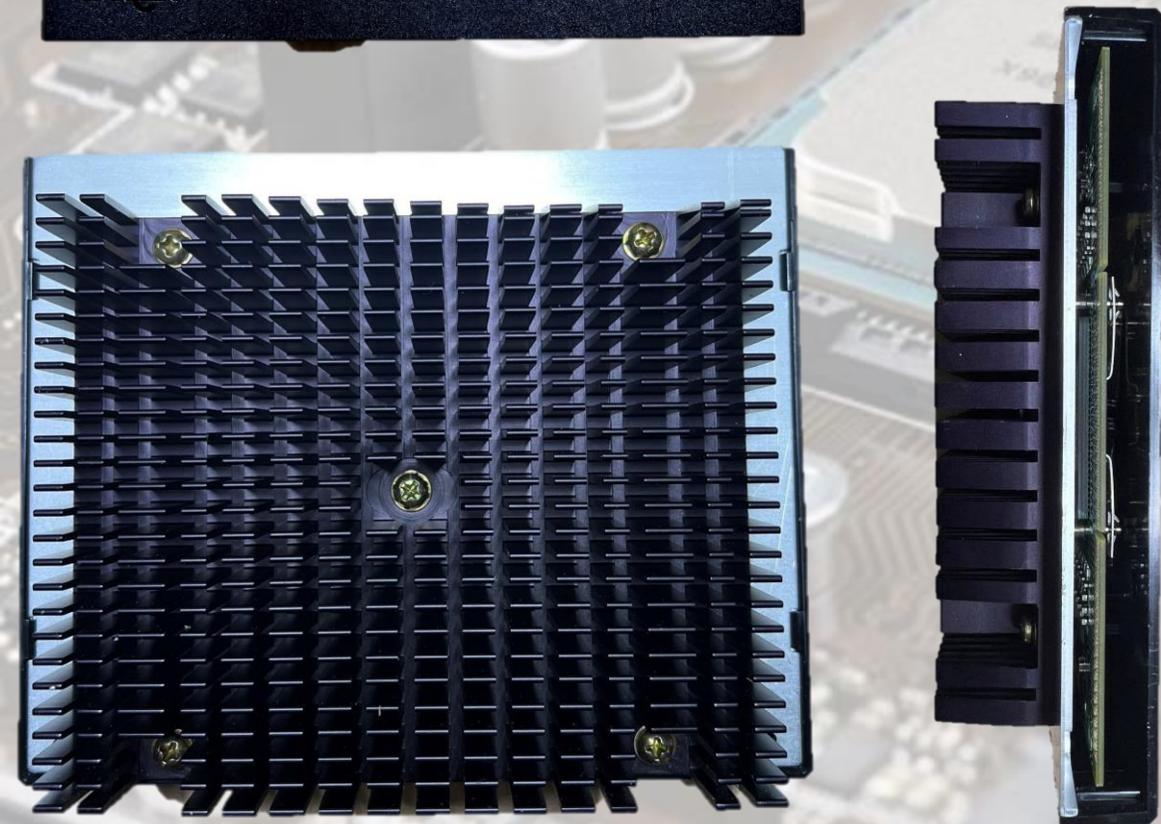
Un poco más de historia reciente



Año 1997: Pentium II Xeon
Nodo: 350nm
Frecuencia: 450MHz
Número de transistores: 7,5 Millones
Cores: 1



Año 2023: Intel Xeon 5ª Gen
Nodo: Intel 7
Frecuencia: 3,5Ghz
Número de transistores: 100.000 Millones
Cores: 64



Intel Xeon 6
Nodo: Intel 3nm
Cores: 144 y 248 por procesador

LIDERAZGO EN FABRICACIÓN: IDM 2.0

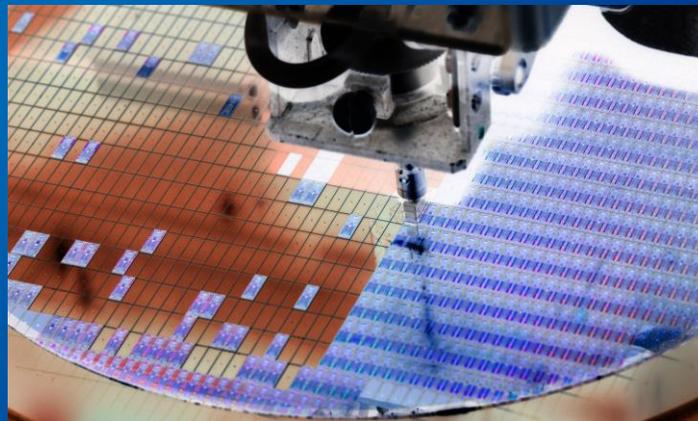
Liderazgo en producto, resiliencia en la cadena de suministro

Red de Fábricas Internas



La red global de fábricas internas de Intel para la fabricación a escala

Fábricas Externas



Uso extendido de la capacidad de fábrica de terceros

Intel "Foundry"



Creación de la unidad de fabricación para terceros. "IFS: Intel Foundry Services"

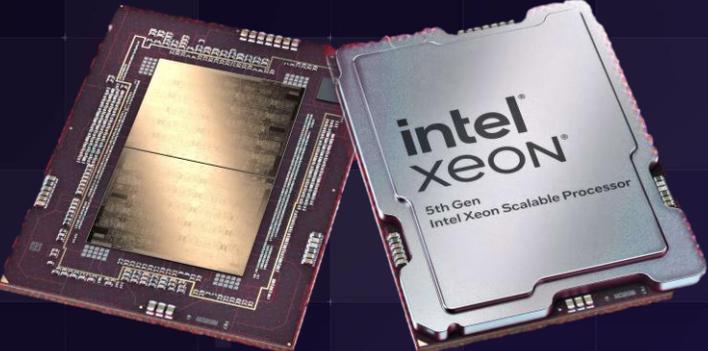
Tecnología líder de Intel en proceso, "packaging" y propiedad Intelectual

Computación para cargas de Inteligencia Artificial

GPU's



CPU's



CPU's



FPGA's



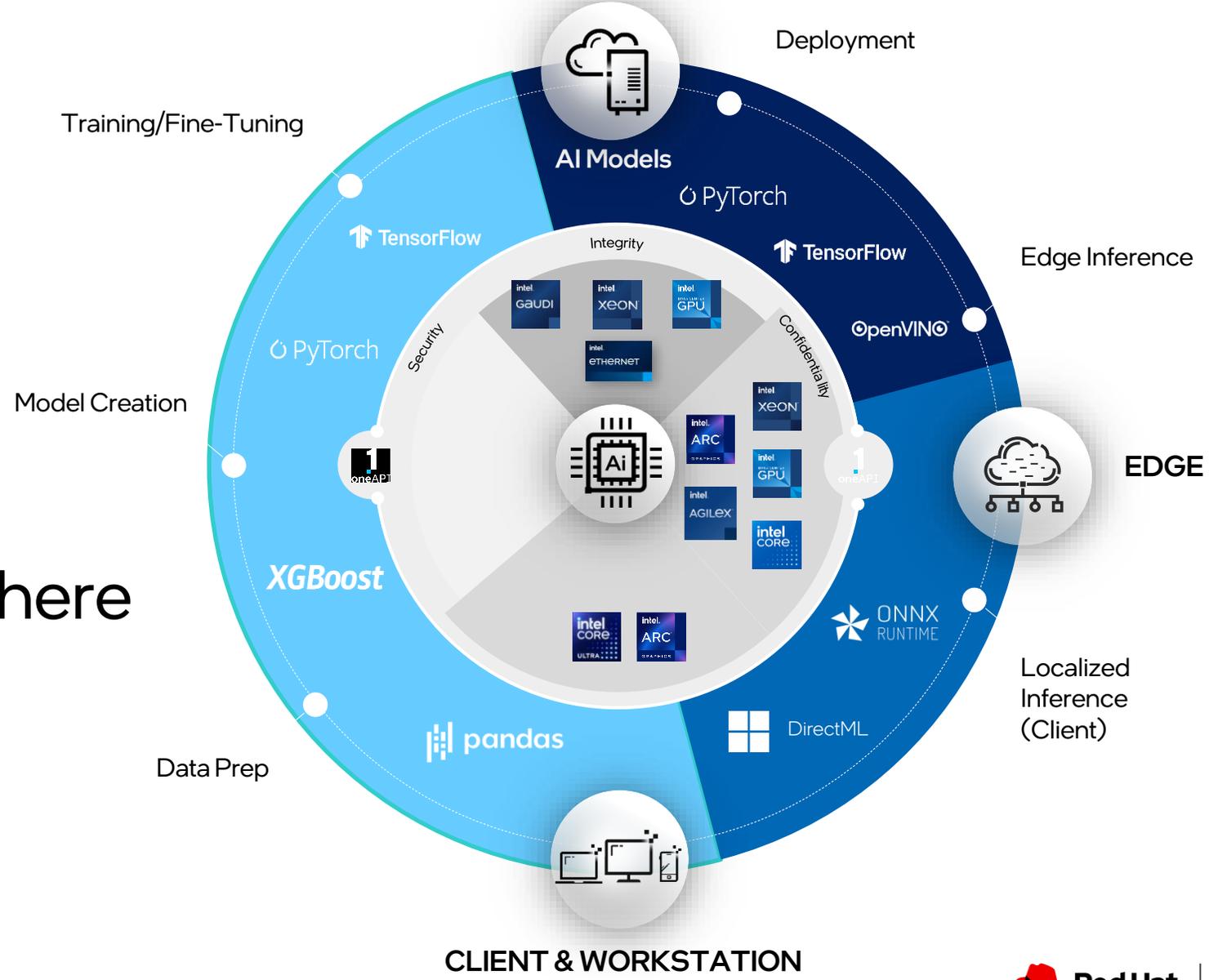
Aceleradores Gaudi 2 y Gaudi 3



NPU's



CLOUD & ENTERPRISE

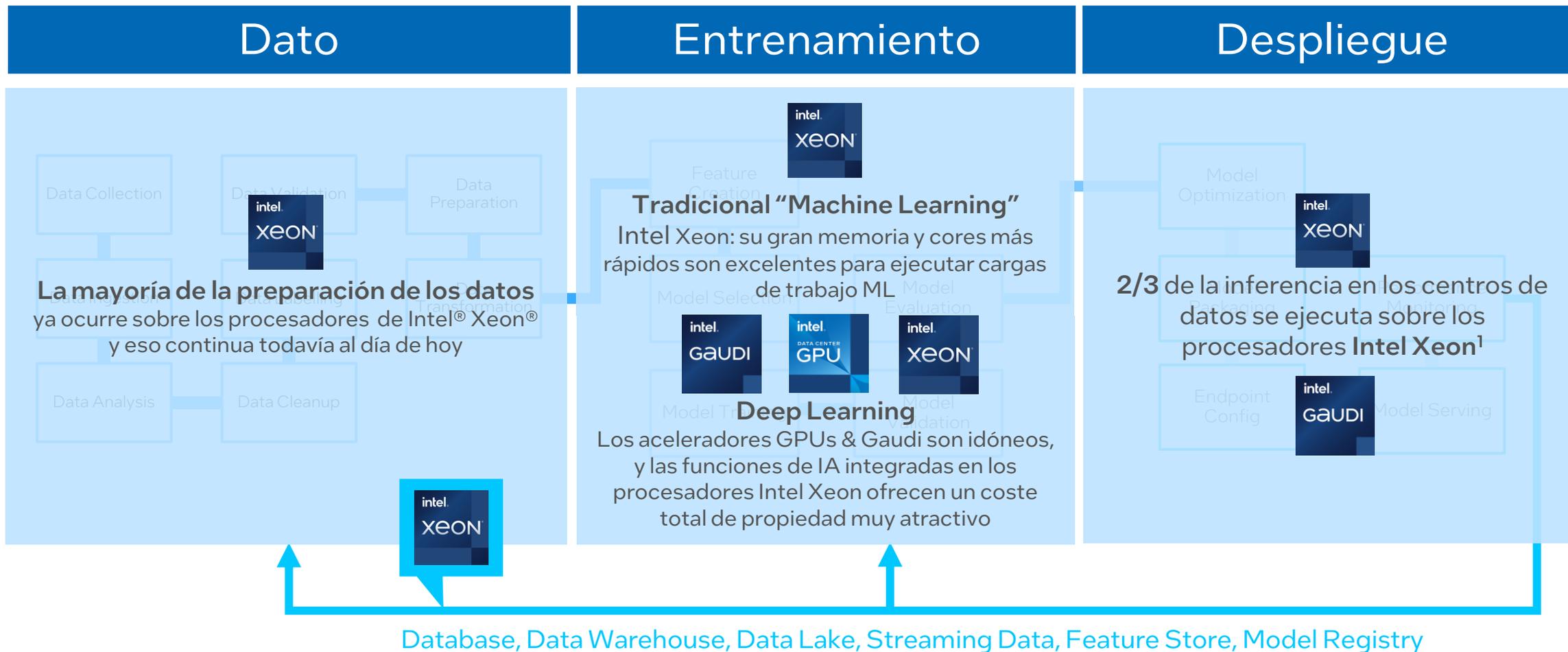


Bringing AI Everywhere

AI Continuum

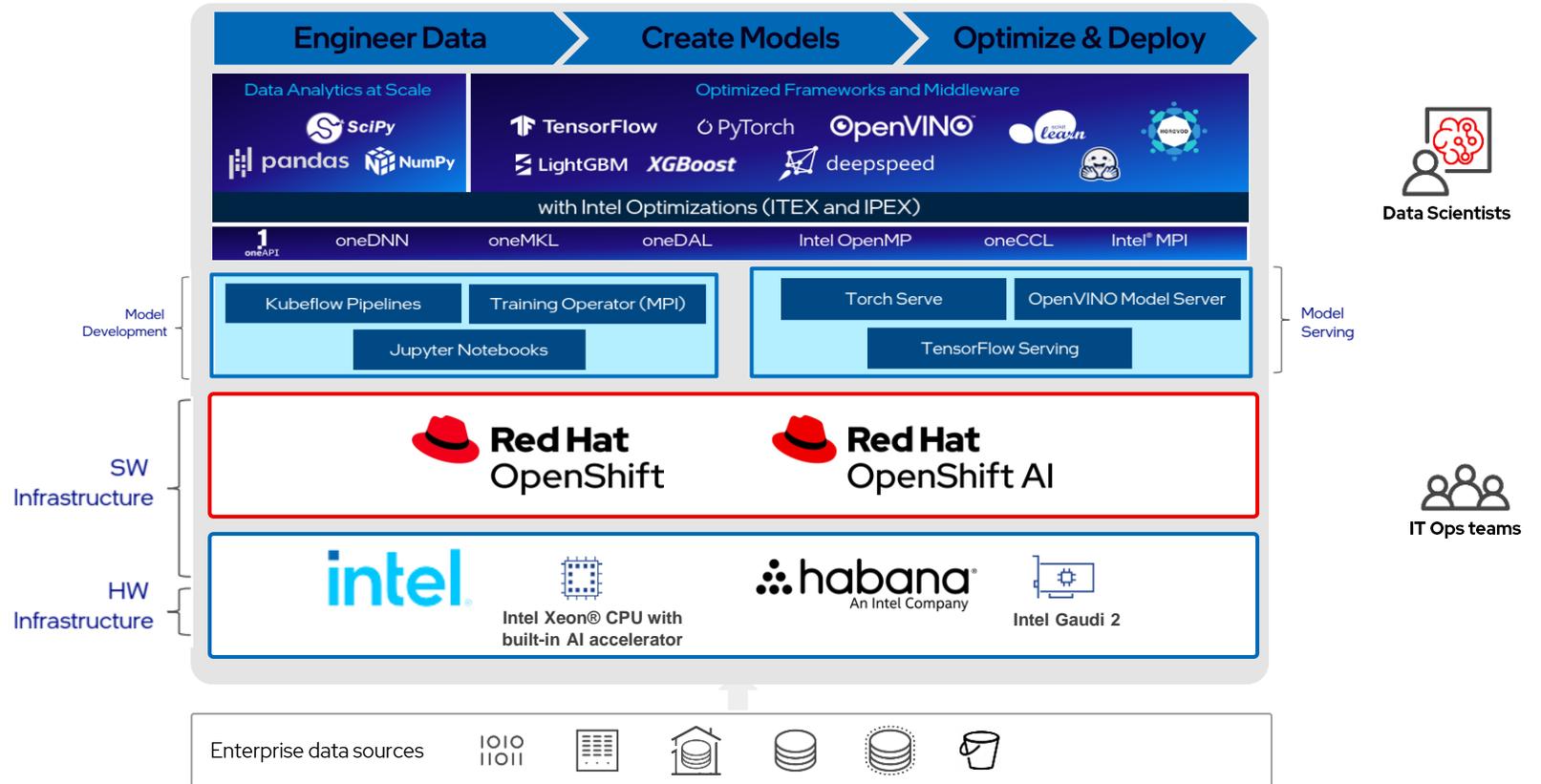
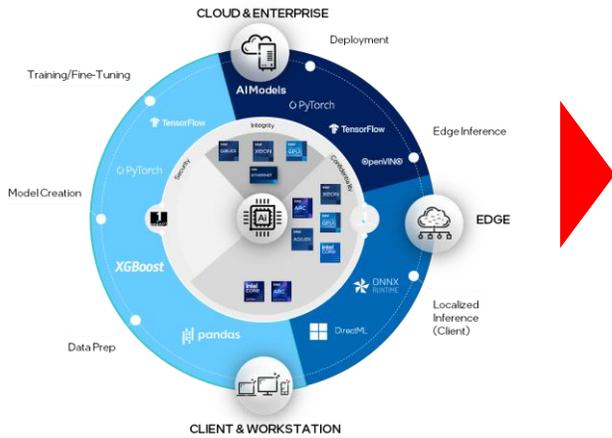
Note: Intel® Core™ Ultra processors integrate NPU low power inference engine from 15th Gen processors onwards.

El camino hacia la AI se ejecuta sobre Intel



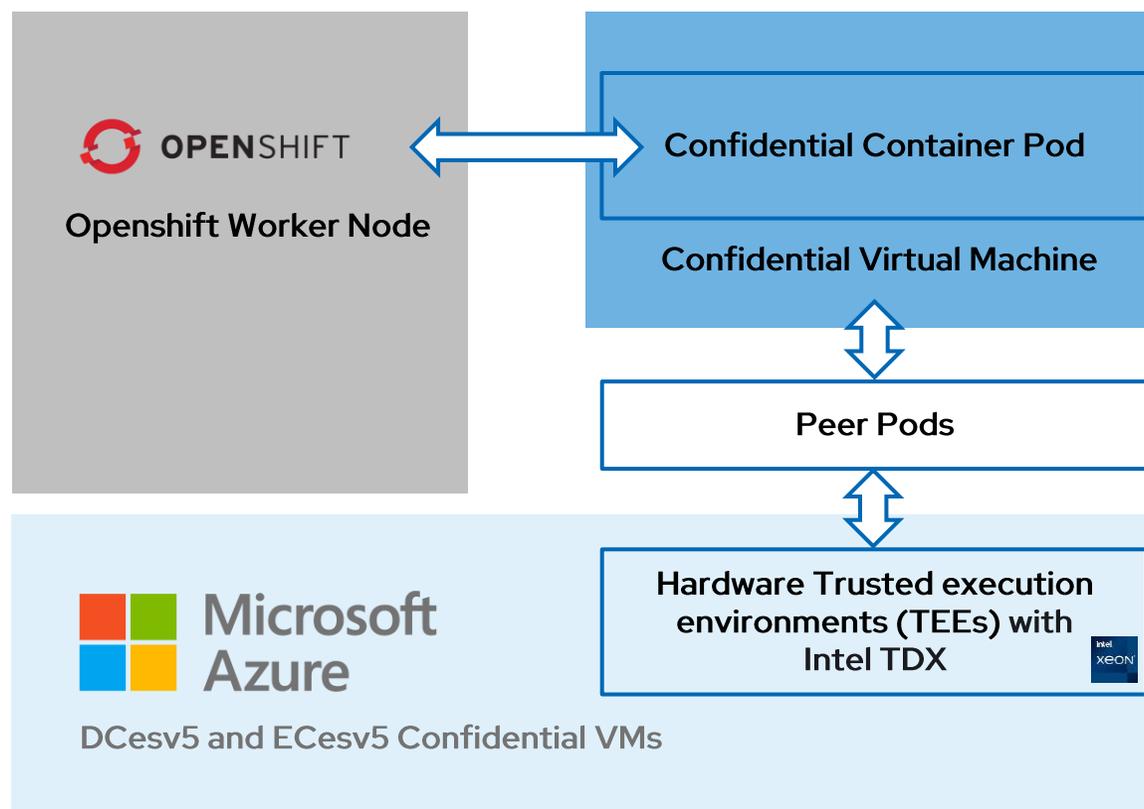
¹ Based on Intel market modeling of the worldwide installed base of data center servers running AI Inference workloads as of December 2022.

Construyendo tu stack de AI



Confidential Containers: Seguridad de cargas de trabajo de IA en la nube pública

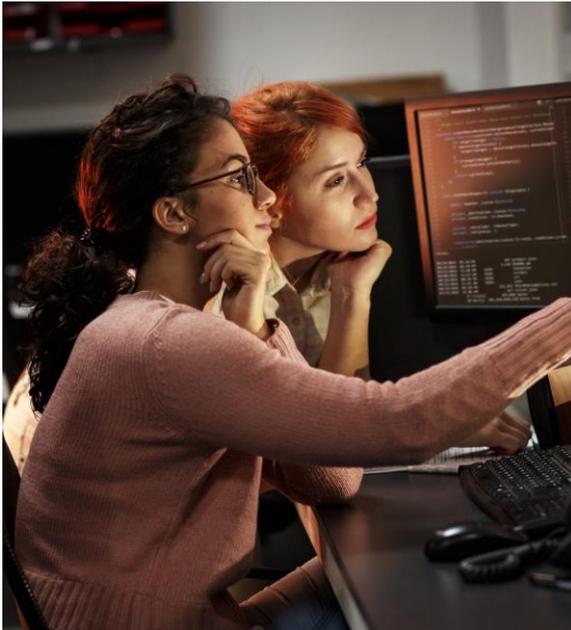
Con Intel Trusted Domain Extensions (TDX)



- ▶ La propiedad intelectual de alto valor, como los modelos de aprendizaje automático (*machine learning models*), requieren seguridad adicional.
- ▶ Confidential Containers para cargas de trabajo de IA, permiten procesar de manera segura datos sensibles en el cloud
- ▶ Intel TDX provee Virtual Machines aisladas a nivel HW diseñadas para proteger los contenedores (datos y aplicaciones) de accesos no autorizados.

Resumen

¡La Inteligencia Artificial está evolucionando muy rápidamente, pero Intel te ofrece la Tecnología que necesitas!



- ▶ Intel posee un amplio portafolio de HW optimizado para tu viaje a la IA
- ▶ Está totalmente integrado con OpenShift y OpenShift AI, que te ofrece rendimiento en aplicaciones y productividad para desarrolladores
- ▶ Para aprender más, únete al programa para desarrolladores de IA de Intel y Red Hat y comenzar a trabajar inmediatamente en IA con las plataformas Cloud configuradas para Red Hat OpenShift y OpenShift AI que te ofrecen los fabricantes.