

Cableado estructurado e Inteligencia Artificial, la infraestructura de red que la hace posible



Guillermo Ortegon
Territory Account Manager
Panduit

¡Bienvenidos!



Guillermo Ortegón A.

Territory Account Manager
Panduit Corp.
Especialista en Gerencia de Proyectos
ATS - Uptime Institute
M. +57 310 763 1845
guillermo.ortegon@panduit.com
[in guillermo-ortegon](https://www.linkedin.com/in/guillermo-ortegon/)

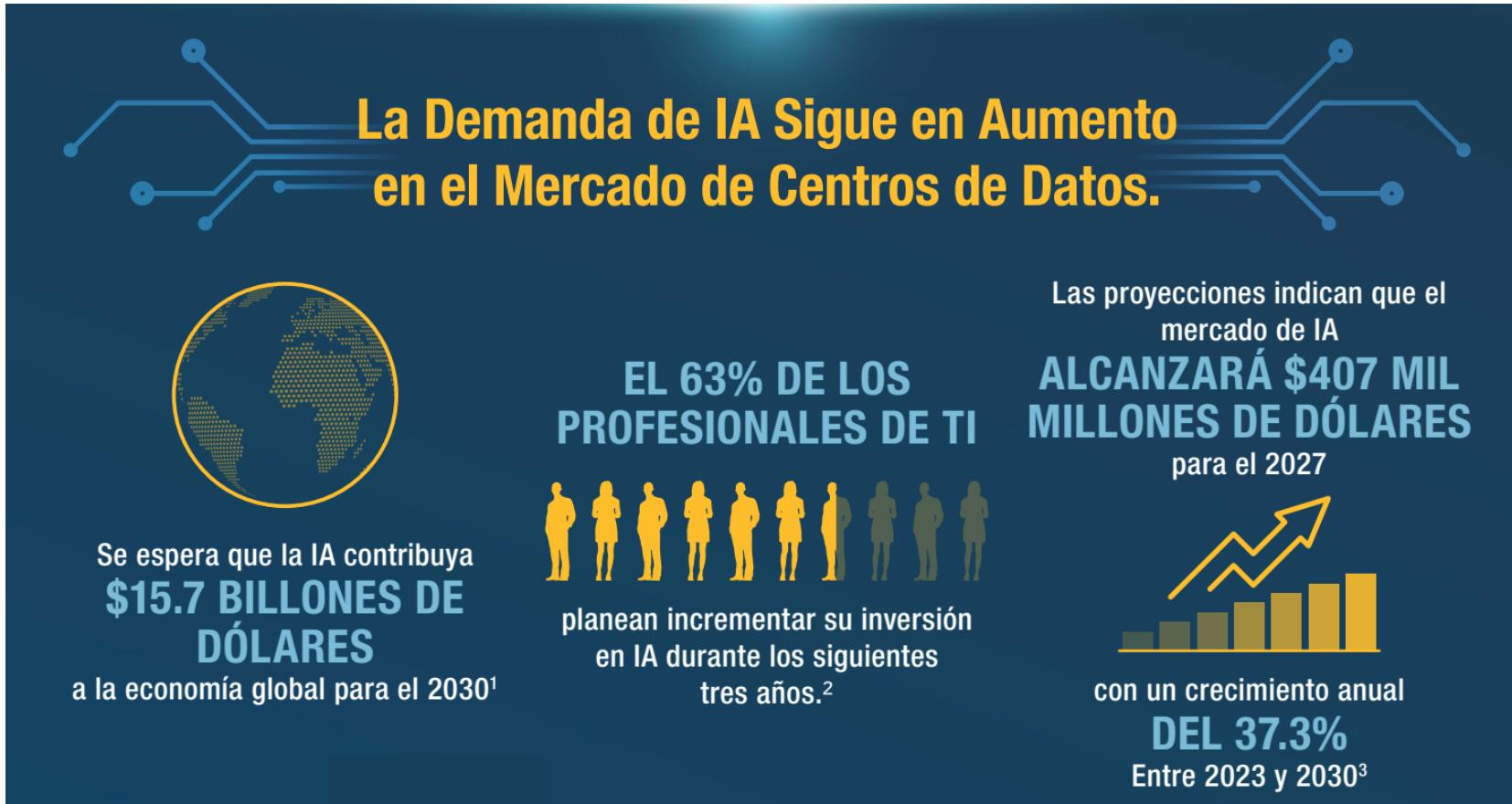


Cableado estructurado e Inteligencia Artificial

La infraestructura de red
que la hace posible



La revolución de la IA. La nueva realidad



1. Fuente: PwC.

2. Fuente: McKinsey Global Institute

3. Fuente: Forbes

La revolución de la IA. La nueva realidad

La Fibra Óptica Permite Crear Redes de Alta Velocidad y Baja Latencia.

 Se requiere **4 VECES MÁS FIBRA** para la IA que para el cómputo normal

 **SÓLO EL 21% DE LAS EMPRESAS** declaran que su red tiene una latencia óptima para sus necesidades de cargas de trabajo de IA⁵

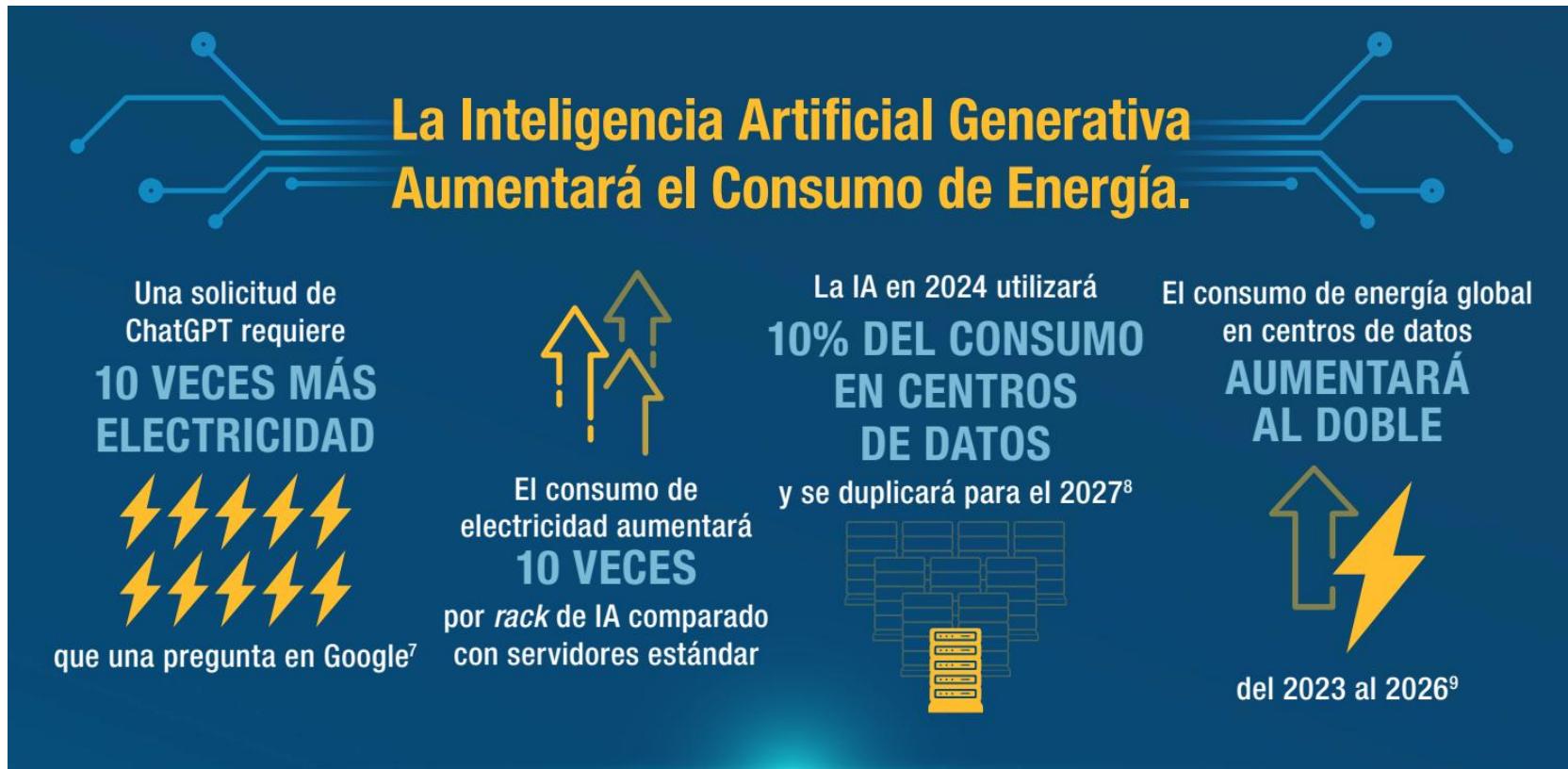
¿SABÍAS QUÉ 

Nvidia espera entregar 100 000 plataformas para servidores de IA en este año y 1.5 millones para 2027⁶?

5. Fuente: Cisco.

6. Fuente: Forbes

La revolución de la IA. La nueva realidad



7. Fuente: Goldman Sachs

8. Fuente: New York Times

9. Data Center Dynamics

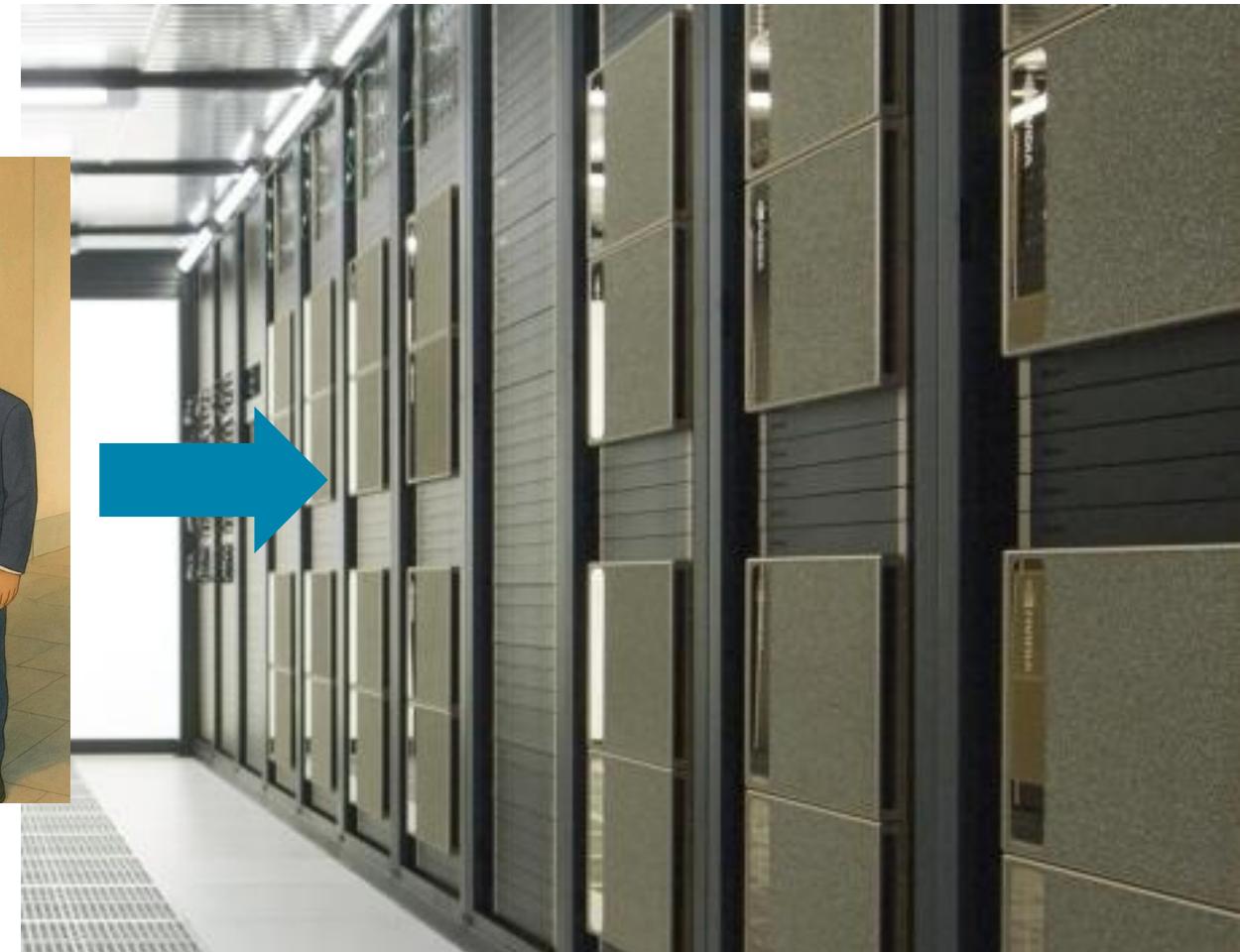
La revolución de la IA. La nueva realidad



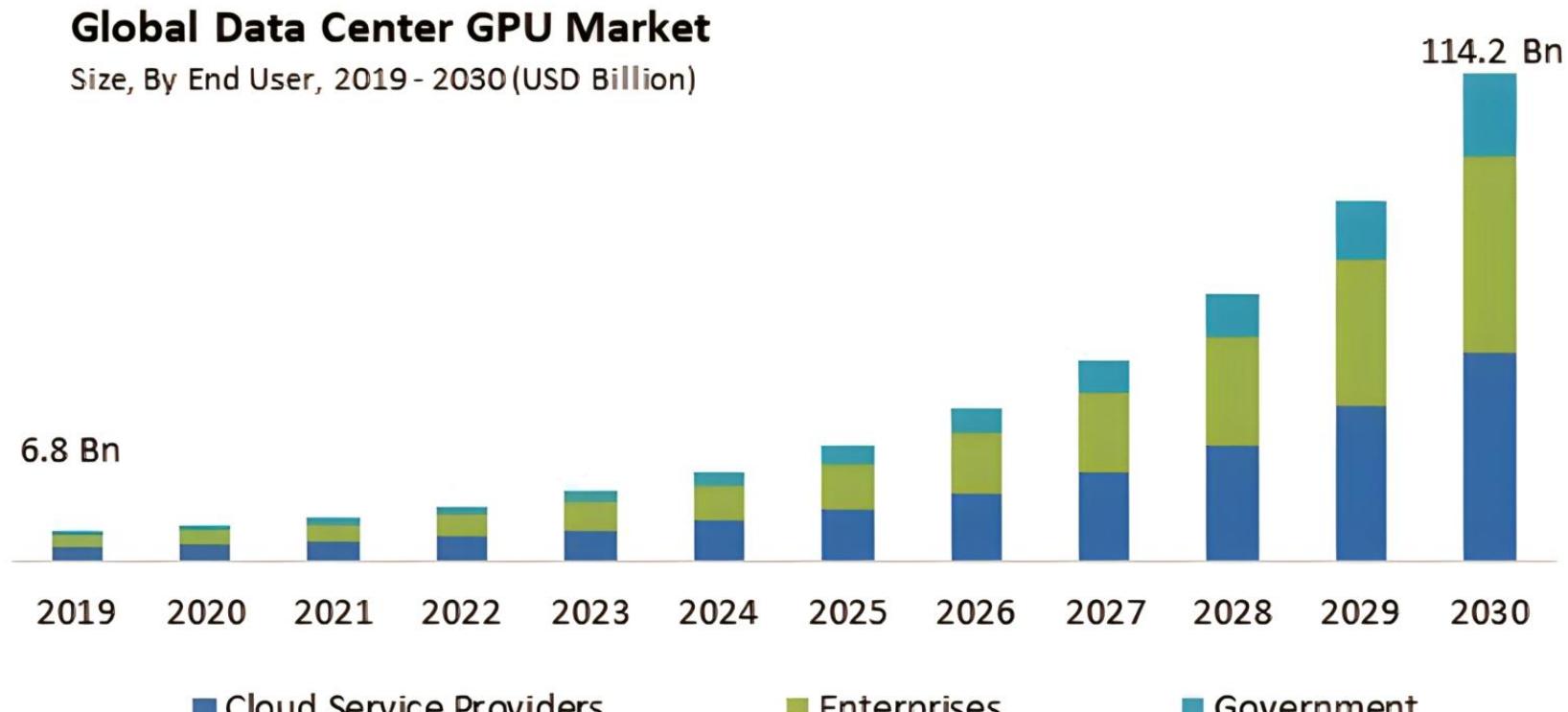
10. Fuente: DataBase

11. Fuente: Forbes

La adopción de la IA. La nueva realidad



Crecimiento esperado en ventas de GPU para IA



Source: www.kbvresearch.com

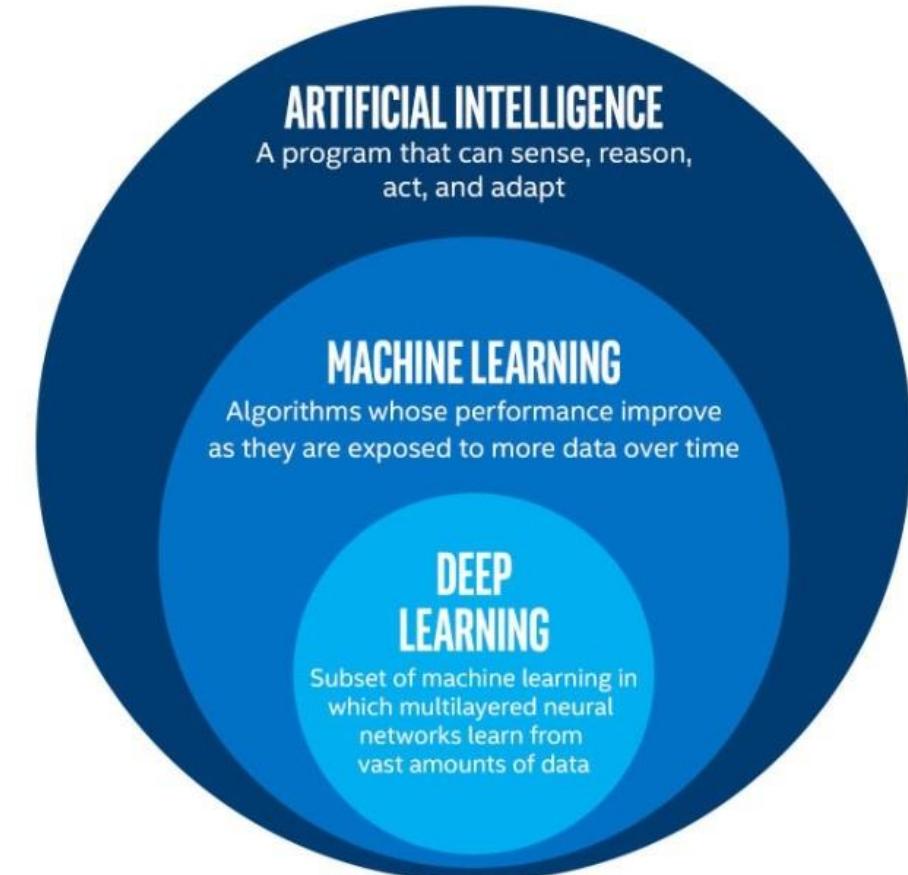
Los tipos de IA

Inteligencia artificial generativa / GenIA: IA generativa que crea nuevo contenido (ChatGPT, Gemini, Copilot).

IA predictiva: Utiliza tendencias para predecir resultados futuros (Fabricación, Finanzas).

Machine learning / Aprendizaje automático (ML): Forma de IA que utiliza un conjunto de datos más amplio y puede aprender con la experiencia.

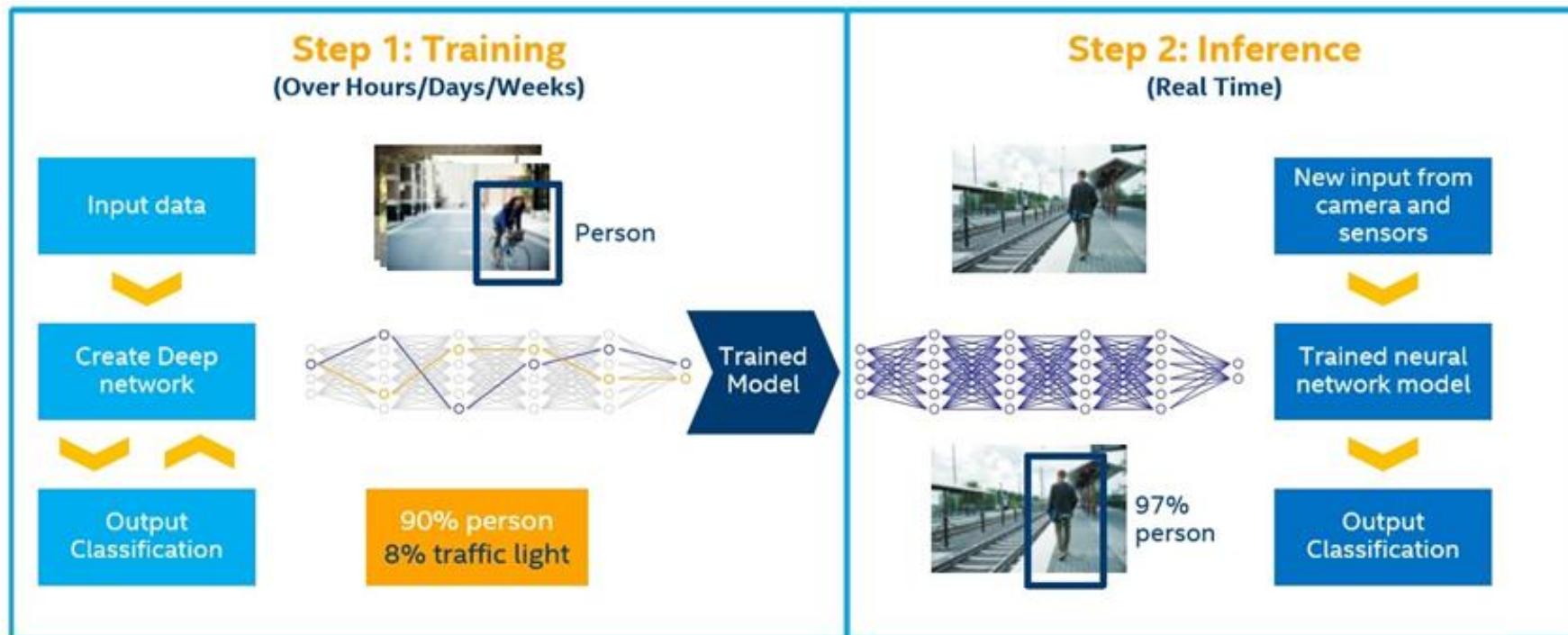
Deep learning / Aprendizaje profundo (DL): Forma de IA que utiliza una red neuronal masiva para interpretar con mayor profundidad los resultados potenciales y es autodirigida para aprender por sí misma.



Del Data Center Hiperescala al Edge

“La mayoría de las implementaciones de IA de hoy se realizan en la nube híbrida, pero en el futuro, habrán más implementaciones Edge, predominantemente para inferencia y aplicaciones de borde...
Nos esperan tiempos muy emocionantes.”

Baron Fung
Dell’Oro Group



El futuro del Entrenamiento vs. Inferencia

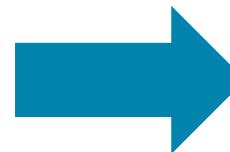
Entrenamiento: Un desarrollador alimenta el modelo con un conjunto de datos seleccionados para que pueda aprender todo lo necesario sobre el tipo de datos que analizará.

Inferencia: El modelo genera resultados/predicciones basados en datos en tiempo real para producir resultados prácticos.

- Los racks de GPU se diseñarán específicamente para cada aplicación: entrenamiento o inferencia.
- Se necesitan ambos para un sistema de IA completo. El entrenamiento es la fase previa antes de implementar la inferencia.
- La inferencia estará en el borde.

“Habrá una **rápida transición del entrenamiento a la inferencia**: para 2030, solo el 10% de la demanda computacional se destinará al entrenamiento y el resto a la inferencia”.

Jeff Clarke
DellWorld 2024



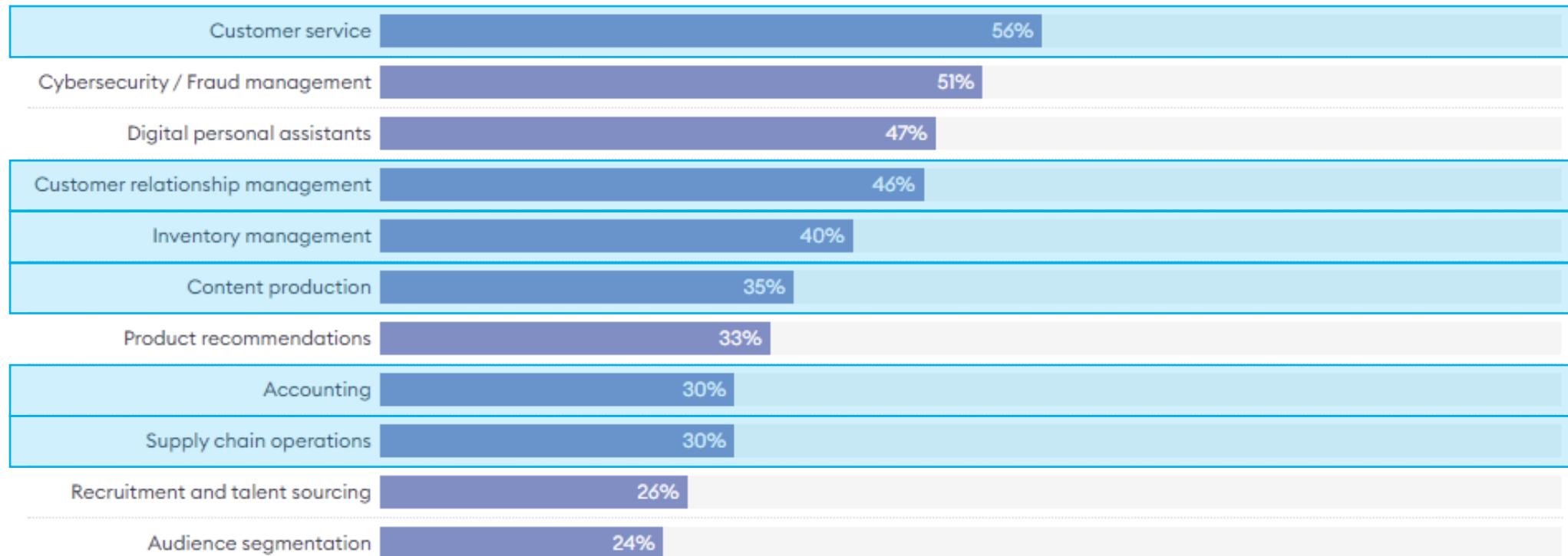
Aplicaciones:

- Vehículos autónomos.
- Análisis predictivo.
- Modelos financieros.
- Soluciones de ciberseguridad.

La tendencia de IA por tipo de negocio

Top Ways Business Owners Use Artificial Intelligence

Forbes Advisor surveyed business owners to find out how they currently use or plan to use AI within their business



Los impulsores en infraestructura para la IA

Infraestructura de red

Conectividad en fibra óptica

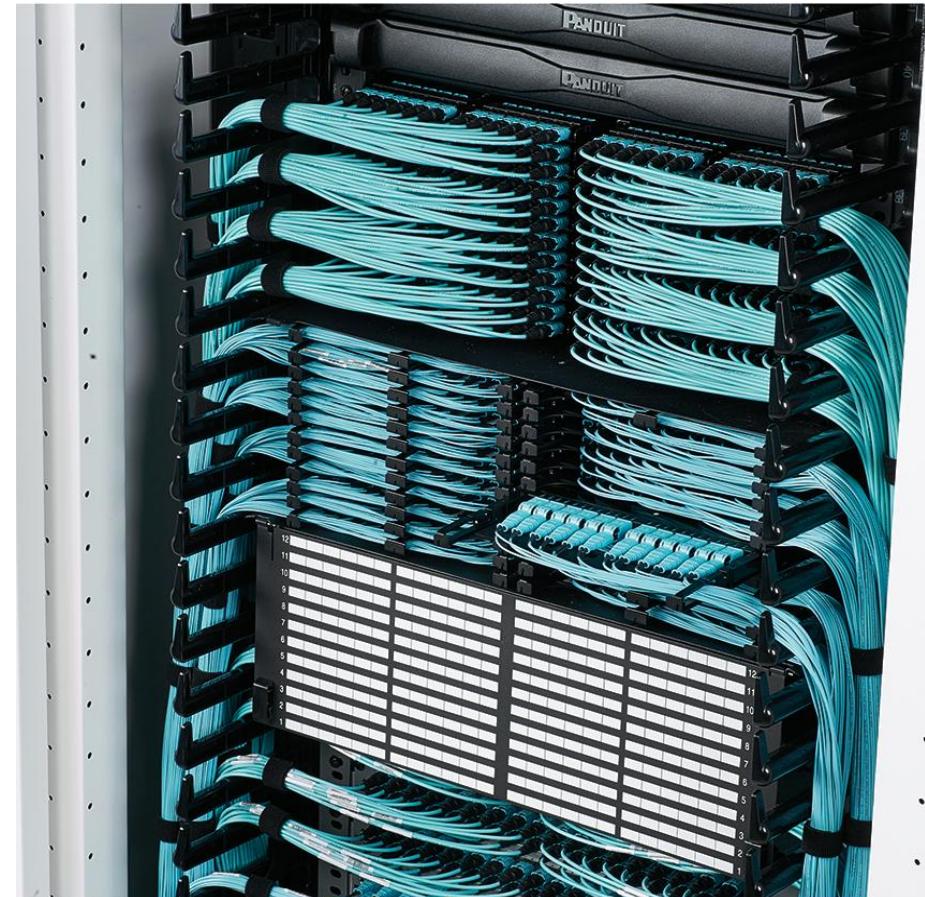
- **Baja Latencia.**
- **Alta densidad preconectorizada** que supera los DC actuales.
- Velocidades de **100, 200 y 400 Gbps.** con una proyección de medio plazo de 800 Gbps. y hasta 1.6 Tbps.
- Conectores **MPO de 8 y 16 fibras** para procesamiento, almacenamiento y comunicaciones.

Conectividad en Cobre

- Categoría 6A para **gestión y telemetría (BMS y DCIM).**

Enrutamiento

- **Canalizaciones de fibra especializadas de mayor capacidad.**
- Múltiples niveles.



Los impulsores en infraestructura para la IA

Espacio. POD (Point of Delivery)

Gabinetes

- Tipo y dimensiones para alta capacidad (48RU - 52RU).
- Anchos estándar de 80 cm.
- Para clientes empresariales se utilizaran 1 o 2 servidores GPU por rack.
- Carga estática >1.5 TON.
- Alta capacidad de circulación de aire >80%.
- Gabinetes que soporten Liquid Cooling directo al chip DCLC.

Confinamiento de pasillos

- Alta eficiencia y reducción de energía (PUE).
- Autosostenido para crecimiento modular que soporte diferentes tipos/marcas de gabinetes.
- Contención para enfriamiento basado en aire sigue siendo el sistema mas usado en DC.



Los impulsores en infraestructura para la IA

Energía.

- Fuentes de suministro fuera de la red, incluidas las **energías renovables** e incluso nucleares.
- Alimentación ininterrumpida basada en Litio (UPS).
- Múltiples PDU de alta capacidad, 120 KW >8 PDU / gabinete.
 - Aumento de la corriente y las **temperaturas de operación (protecciones avanzadas hidráulico-magnéticas)**.
 - Tipo y número de tomas de corriente gestionables o administrables.

Enfriamiento.

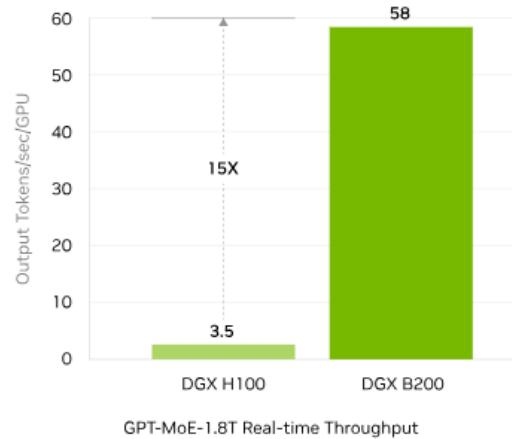
- Sistemas **Rear Door Heat Exchanger RDHX**.
- Refrigeración líquida directa al chip **DCLC (>60 kW/rack)**.
- Avance en tecnologías de enfriamiento por inmersión.



IA GPU, capacidad de procesamiento

Next-Generation Performance Powered by DGX B200

Real Time Large Language Model Inference



Supercharged AI Training Performance



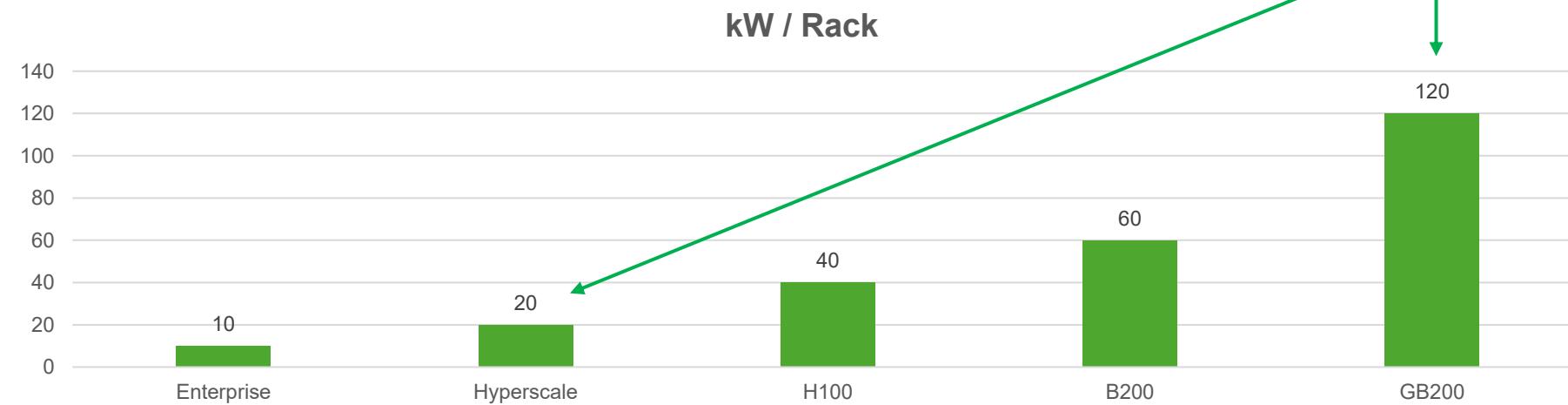
Versiones recientes de GPU pasan a 3X el desempeño en entrenamiento y hasta 15X en inferencia.

Esto demanda conexiones de alta velocidad, de baja perdida y baja latencia.

IA GPU, requerimientos de alimentación

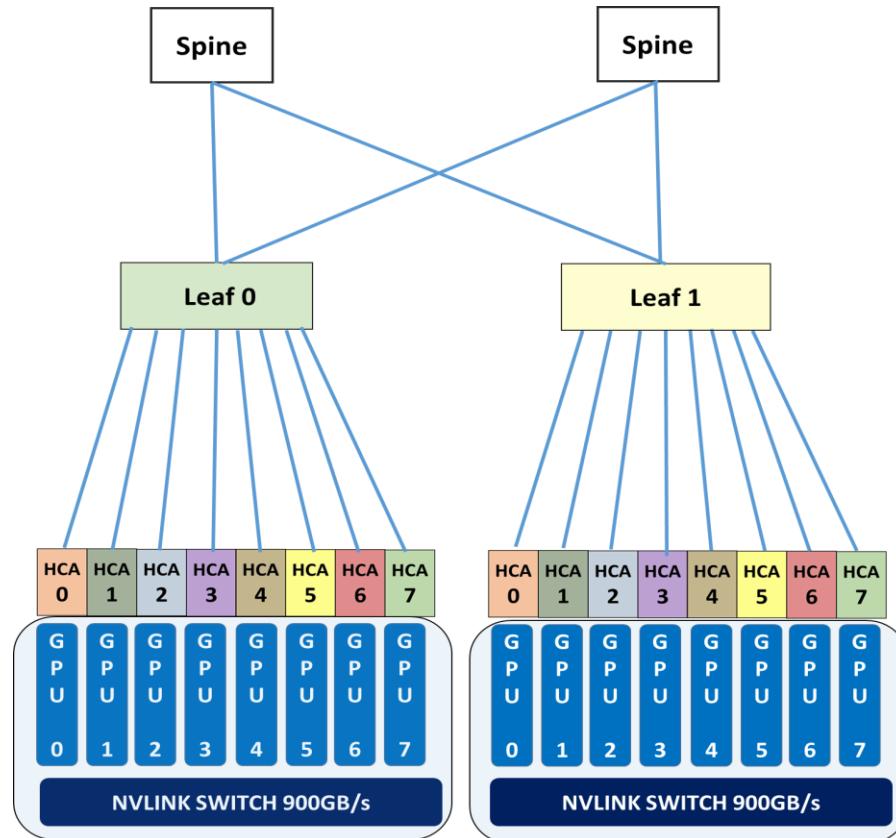
Nvidia GPU	Launch Date	Name	Power
V100	2018	Volta	300W
A100	2020	Ampere	400W
H100	2022	Hopper	700W
B200	2024	Blackwell	1000W
GB200*	2024	Grace Blackwell	2700W
R100**	2026	Rubin	?

El uso de energía de la IA es de **6 a 10** veces mayor que un hiperescala.

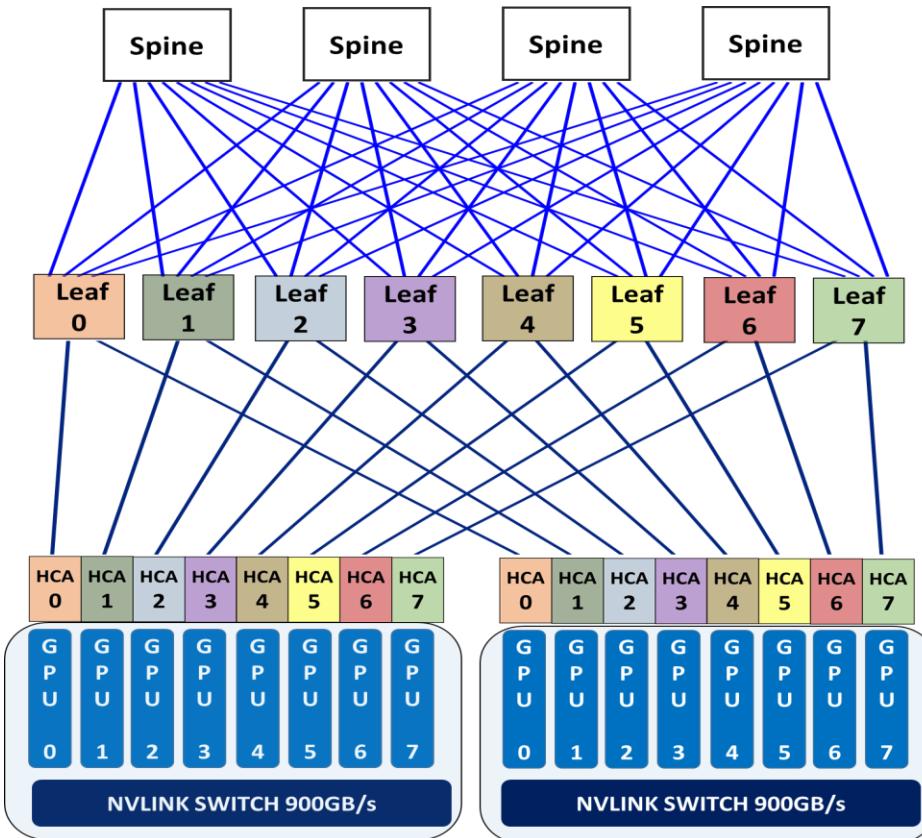


Arquitectura de conectividad en fibra óptica

Traditional Fat Tree Spine-Leaf

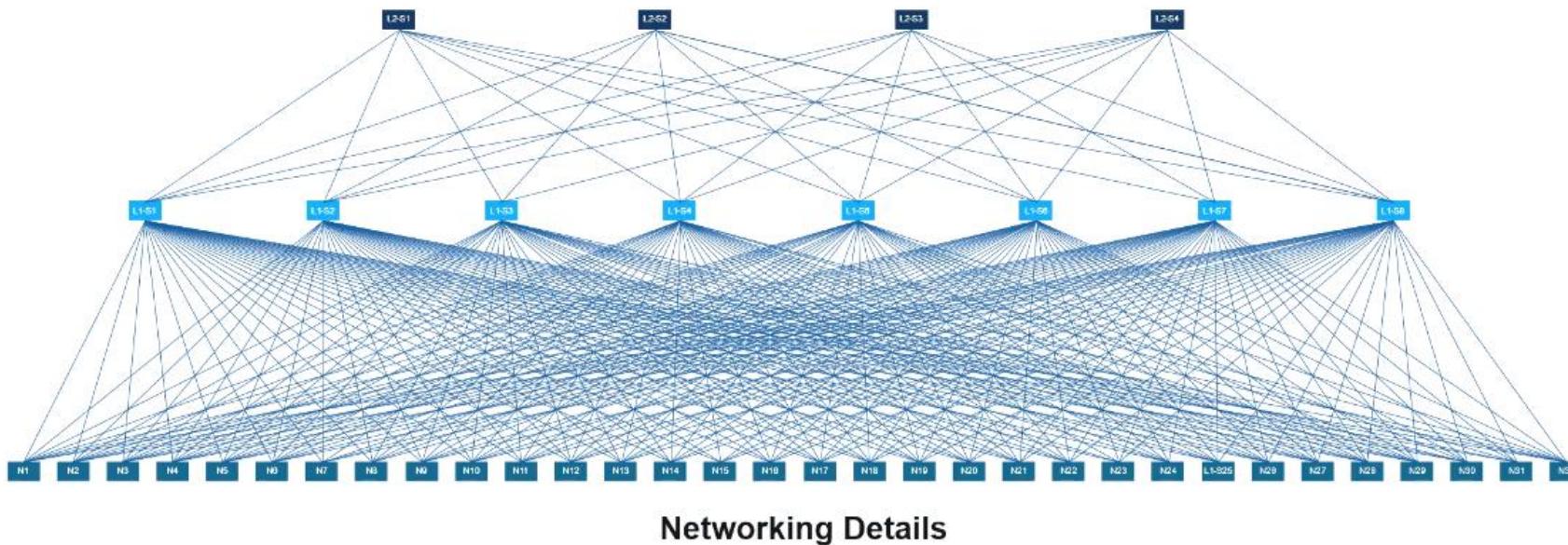


Rail Optimized H100



Arquitectura de conectividad SuperPOD

32-Node (SU) Rail- Optimized Network Topology

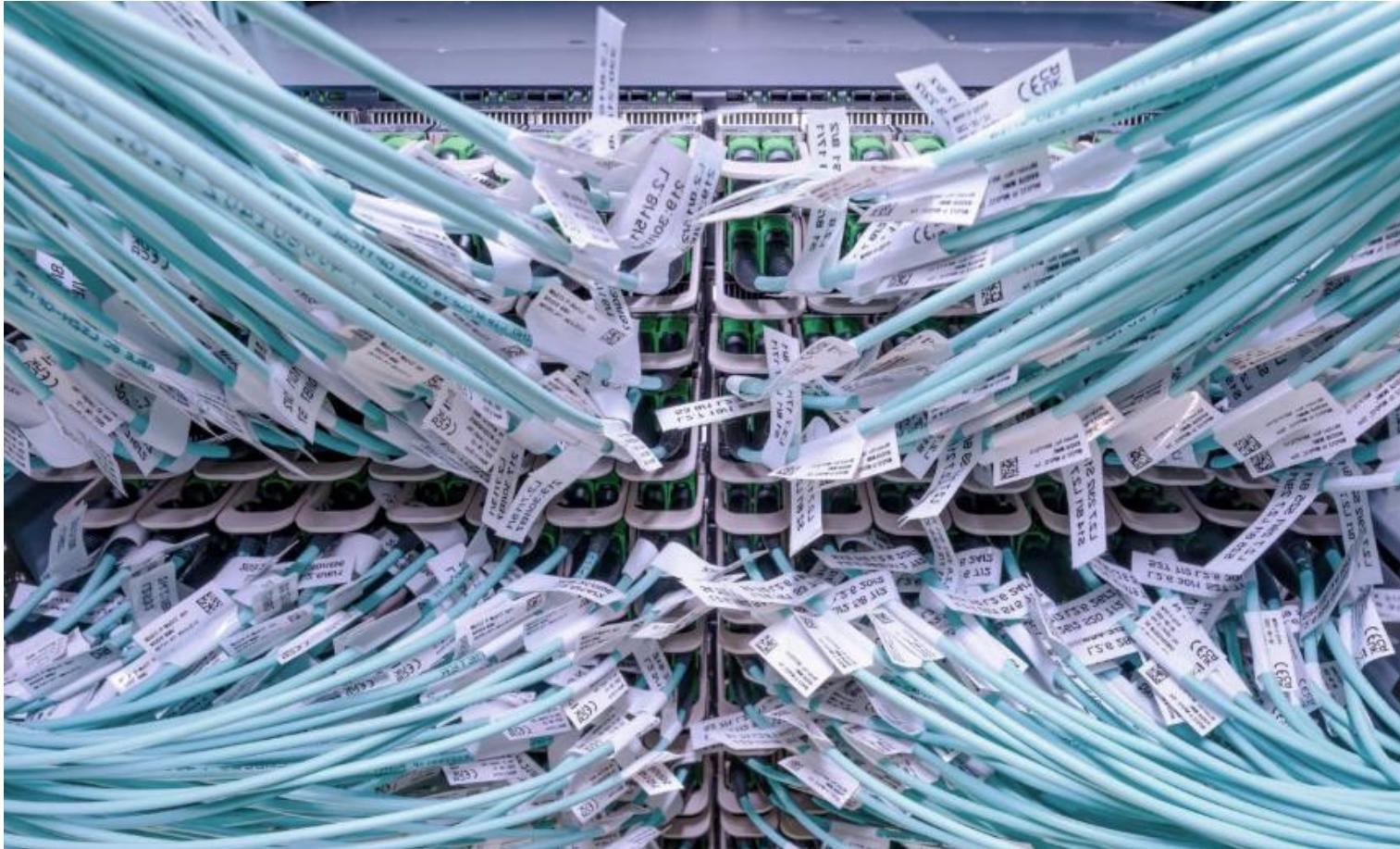


Networking Details

- 12x NVIDIA QM9700 64-Port InfiniBand NDR, 32 OSFP ports switches
 - Back-end compute non-blocking network
 - 8x Leaf switches (L1)
 - 4x Spine switches (L2)

Mas de 10.000 fibras
en un solo sistema
DGX SuperPod.

Conexiones en un gabinete de IA



Cableado estructurado de alta densidad



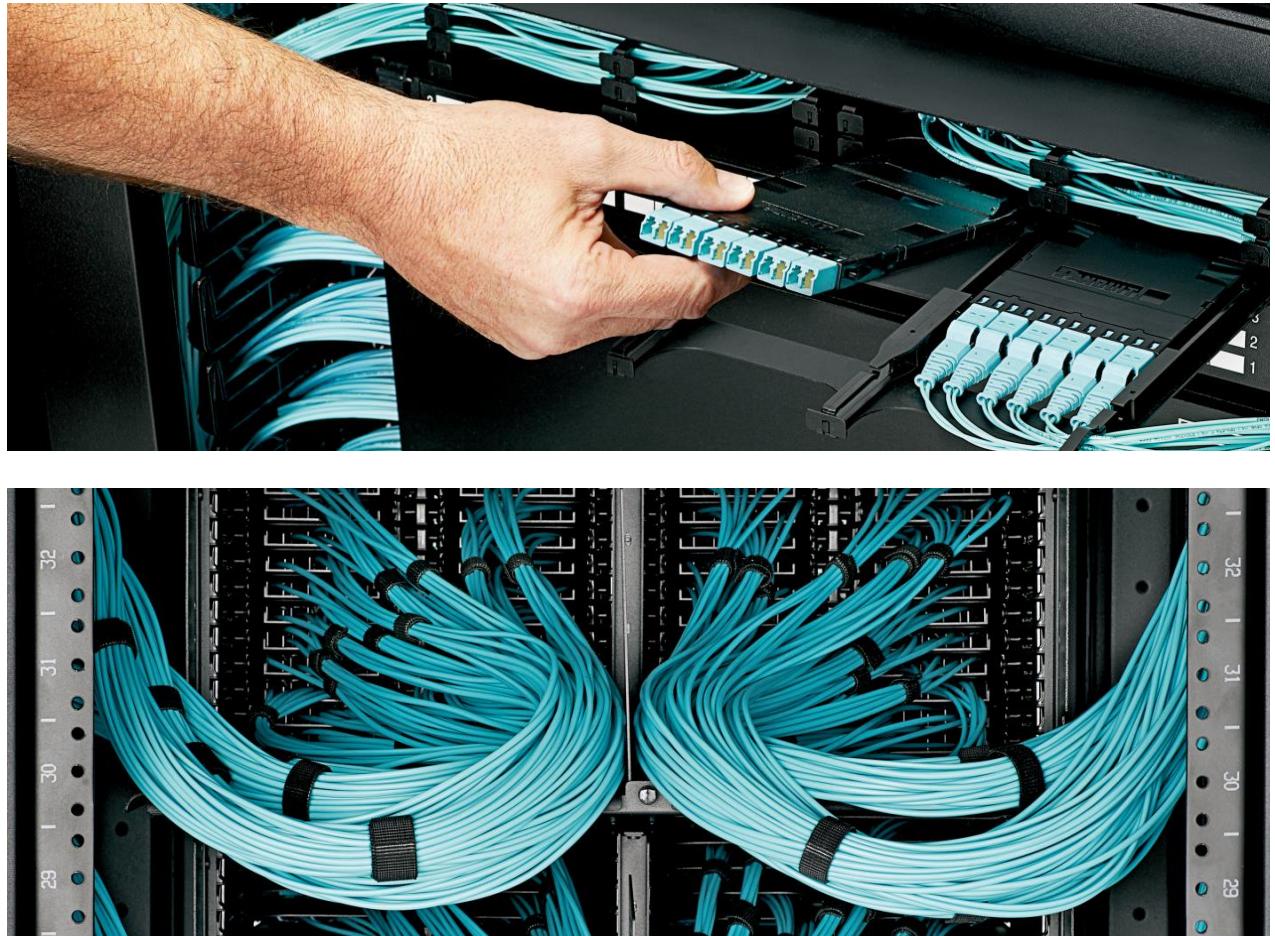
Cableado estructurado para IA

La adopción global de IA está creciendo a una tasa anual compuesta (TCAC) del 40 % (1).

En los centros de datos de IA, la densidad de fibra es de cuatro a ocho veces mayor que en los centros de datos tradicionales.

Los Data center para IA estarán diseñados para soportar velocidades de hasta 1.6 Terabits por segundo.

(1) Fuente: Grand View Research.



Las ventajas del **cableado estructurado en IA**

- Sin efectos negativos en la latencia.
- Modular y escalable.
- Enrutamiento más sencillo.
- Reduce el tiempo de implementación.
- Control de holguras de cables.
- Mejora la gestión de las conexiones.
- Ideal para altas densidades.



Las ventajas del cableado estructurado en IA

	Cableado estructurado	Cableado punto a punto
Baja latencia	Bueno Los conectores de fibra no afectan la latencia	Bueno Menos componentes
Alta densidad de cables	Bueno Consolida varios hilos en pocos troncales, reduce el uso de canalizaciones hasta en un 70%	Pobre Mayor número de cables, Mas congestión
Facilidad de instalación	Bueno Simplifica el tendido de enlaces de fibra antes de instalar los gabinetes	Medio Solo un cable del servidor al switch, aunque a largo plazo se dificulta la administración
Gestión de holguras	Bueno Troncales y cables de parcheo con longitudes estándar con control de holguras	Pobre Inconvenientes al asegurar la longitud más holguras

Las cuatro redes para IA

Compute

Back-end network

East / West Network

- InfiniBand or Ethernet
- Highest Speed (400G)
- GPU to GPU
- Strict 1:1 ratio

Management

In-Band

Font-End Network

North / South Network

- Ethernet
- High speed (200G)
- Admin to cluster
- Loose 1:1 ratio

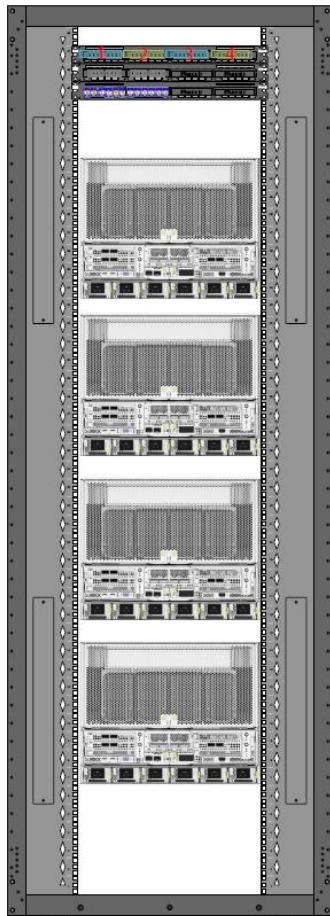
Out-of-Band

- Copper (1G / 10G)
- Admin redundancy
- Oversubscribed

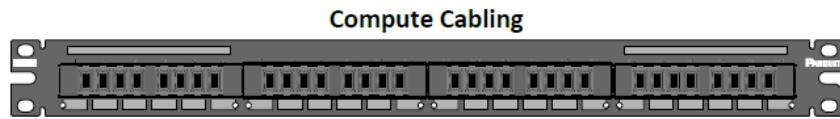
Storage

- InfiniBand or Ethernet
- High Speed (200G)
- GPU to storage
- Strict 1:1 ratio

Componentes del cableado estructurado en IA



Server Cabinet



(32) MPO interconnects from Nodes to Patch Panel and (32) from Patch Panel to Leaf switch
(1) QPP24BL Patch Panel with (4) FQMAP85BL 8-MPO FAPs

Each H100 node needs (1) eight fiber MPO cable to connect to each of the (8) Leaf switches for a total of (8) MPOs per node. (4) nodes per rack needs a total of (32) MPO cables per rack

Cabling Option:
Use (4) 64 fiber trunks with (8) MPOs instead of (32) 8 fiber MPO interconnects from Patch Panel to Leaf switch to reduce pathway congestion



(8) MPO interconnects for Storage and (8) MPO interconnects for In-band management
Each H100 node needs (4) Network / Storage MPO cables: (2) Storage and (2) In-band management
Total of (16) MPO cables per rack only use two of the four Patch Panel slots

Cabling Option:
Can combine the copper jacks in this patch panel if the port count is 12 or fewer

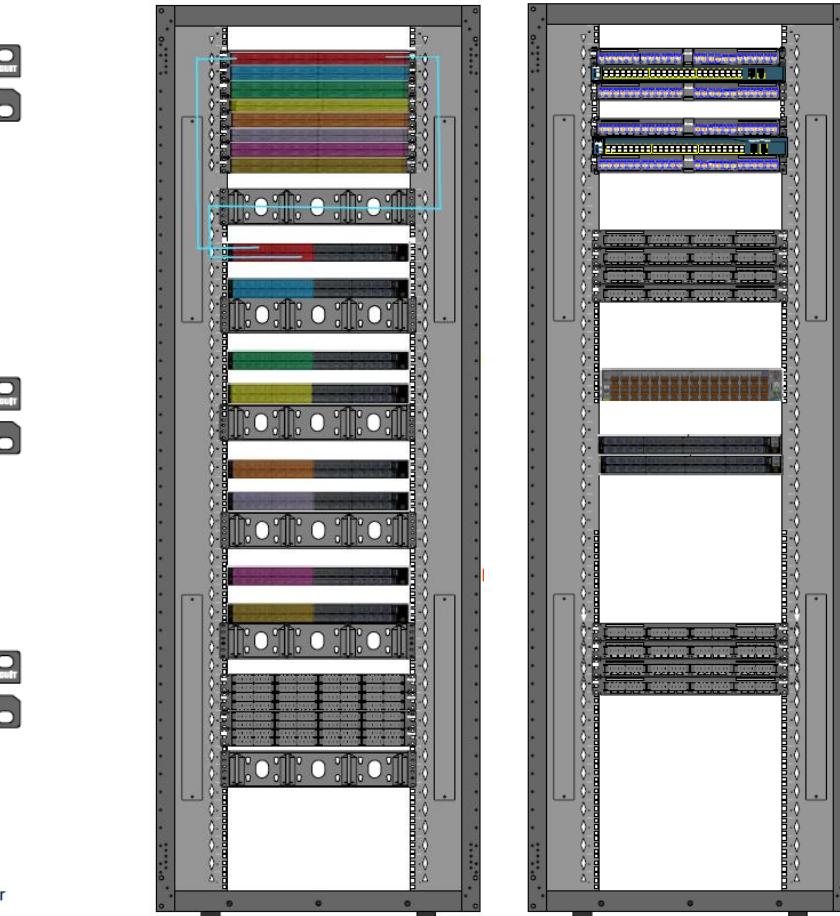
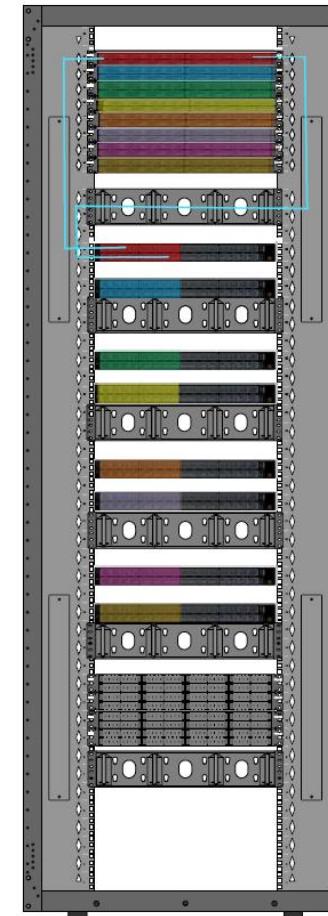


(7-12) Cat6 patch cords up to Patch Panel and from Patch Panel to OBM switch
(1) CFAPPBL1 Patch Panel with (2) FMP6 plates, (12) CC688BU couplers and (2) FAPB blank plates

Each H100 node has (1-2) 1G ports

(3-4) PDU Sensor ports and (1-4) other sensor ports

Cabling Options:
- QPP24BL with (2) 6-jack pre-terminated Quicknet Trunk per slot - High performance, large bundle
- CFAPPBL1 with (12) field terminated CJ6X88TGBU jacks - Highest performance but requires on-site labor



Network Cabinets
Separate Compute from Management/Storage

Alta densidad para la red de computo

Cable preconectorizado
MPO - MPO Base 8



Adaptador de perfil delgado
MPO a MPO



Administrador o panel
de 1 RU



Los sistemas basados en bandejas de alta densidad soportan hasta 576 fibras (72 puertos MPO) por RU. Son la opción conveniente para racks con servidores de 4 GPU como NVL72.

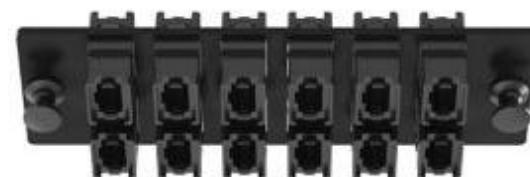
Aplicación convencional para gestión en banda

Cable preconectorizado
MPO - MPO Base 8



+

Adaptador
MPO - MPO



+

Administrador de fibra
de 1 RU



Los sistemas basados en bandejas soportan hasta 576 fibras (72 MPO) por RU, usando diferentes tipos de paneles adaptadores MPO de baja y media densidad. Son soluciones escalables para racks con servidores de 8 GPU como H100 y B200.

Aplicación modular para red de almacenamiento

Cable preconectorizado
MPO - MPO Base 8



Adaptador modular
MPO a MPO



Panel de parcheo de
1 RU



Los sistemas basados en paneles de conexión soportan hasta 512 fibras (64 puertos MPO) por RU. Son soluciones escalables para racks con servidores de 8 GPU como H100 y B200.

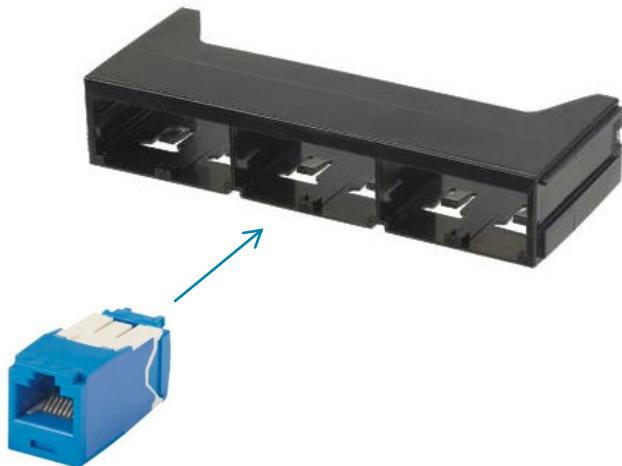
Aplicación modular para red fuera de banda

Cable UTP
Categoría 6A / 6



+

Adaptador
Cassette a RJ-45



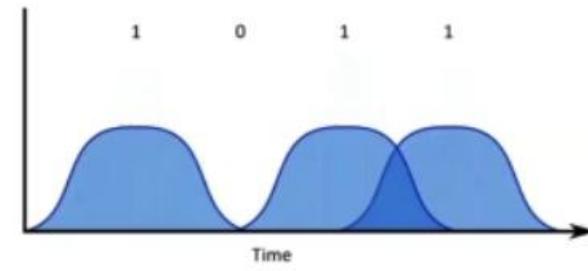
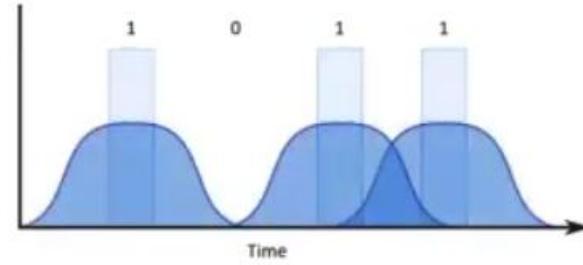
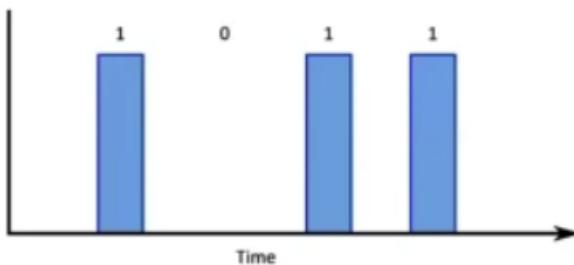
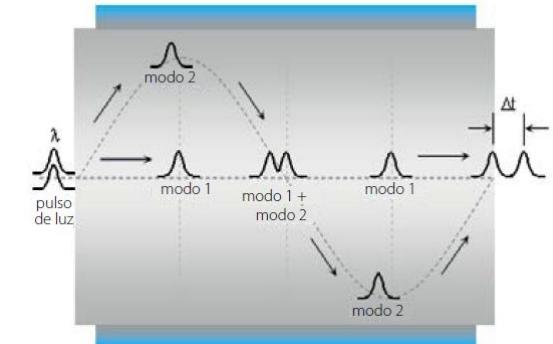
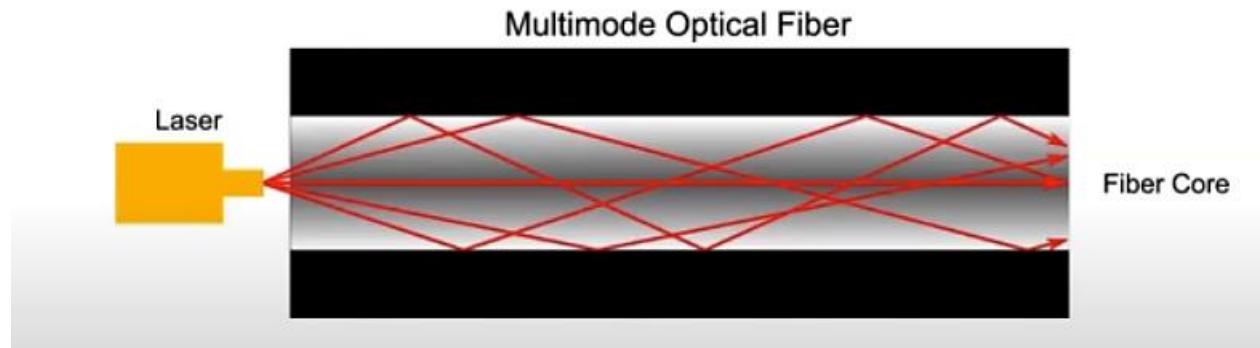
+

Panel de parcheo de
1 RU



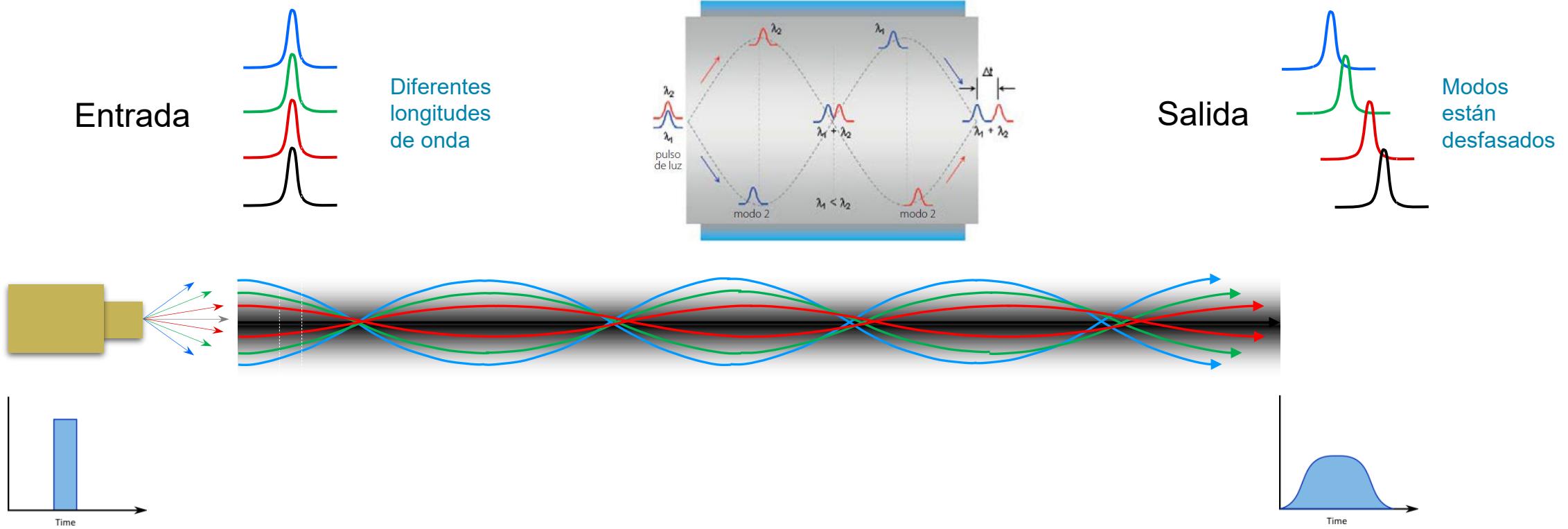
Los sistemas basados en paneles de conexión soportan hasta 512 fibras (64 puertos MPO) por RU. Son soluciones escalables para racks con servidores de 8 GPU como H100 y B200.

Dispersión modal para FO Multimodo



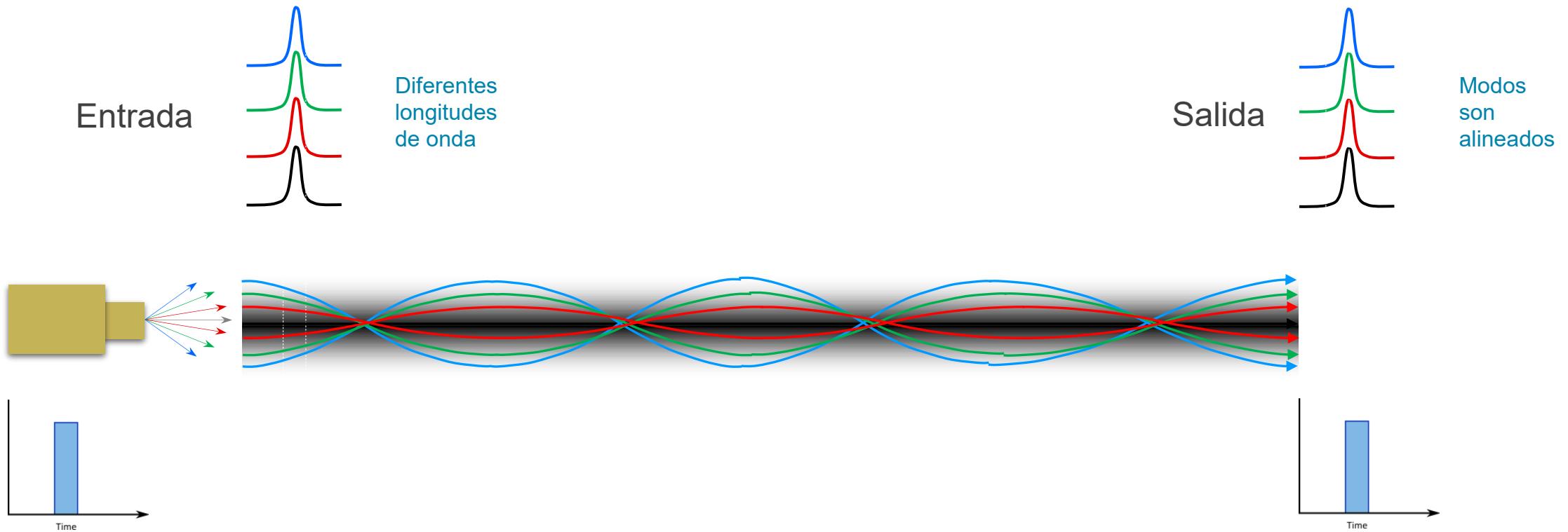
La mayoría de las fibras ópticas actuales para Enterprise y Data center corrigen este fenómeno para transmisiones de baja velocidad, sin embargo, cuando se aumentan, entran a considerarse otros fenómenos que afectan el desempeño de la luz en el medio.

Dispersión cromática en FO Multimodo



Al igual que un prisma, la luz se esparce en múltiples colores, los cuales, al tener diferentes longitudes de onda, causan distorsiones significativas, tiempos de tránsito desiguales con la consecuente perdida de paquetes y retransmisiones.

Dispersión cromática corregida



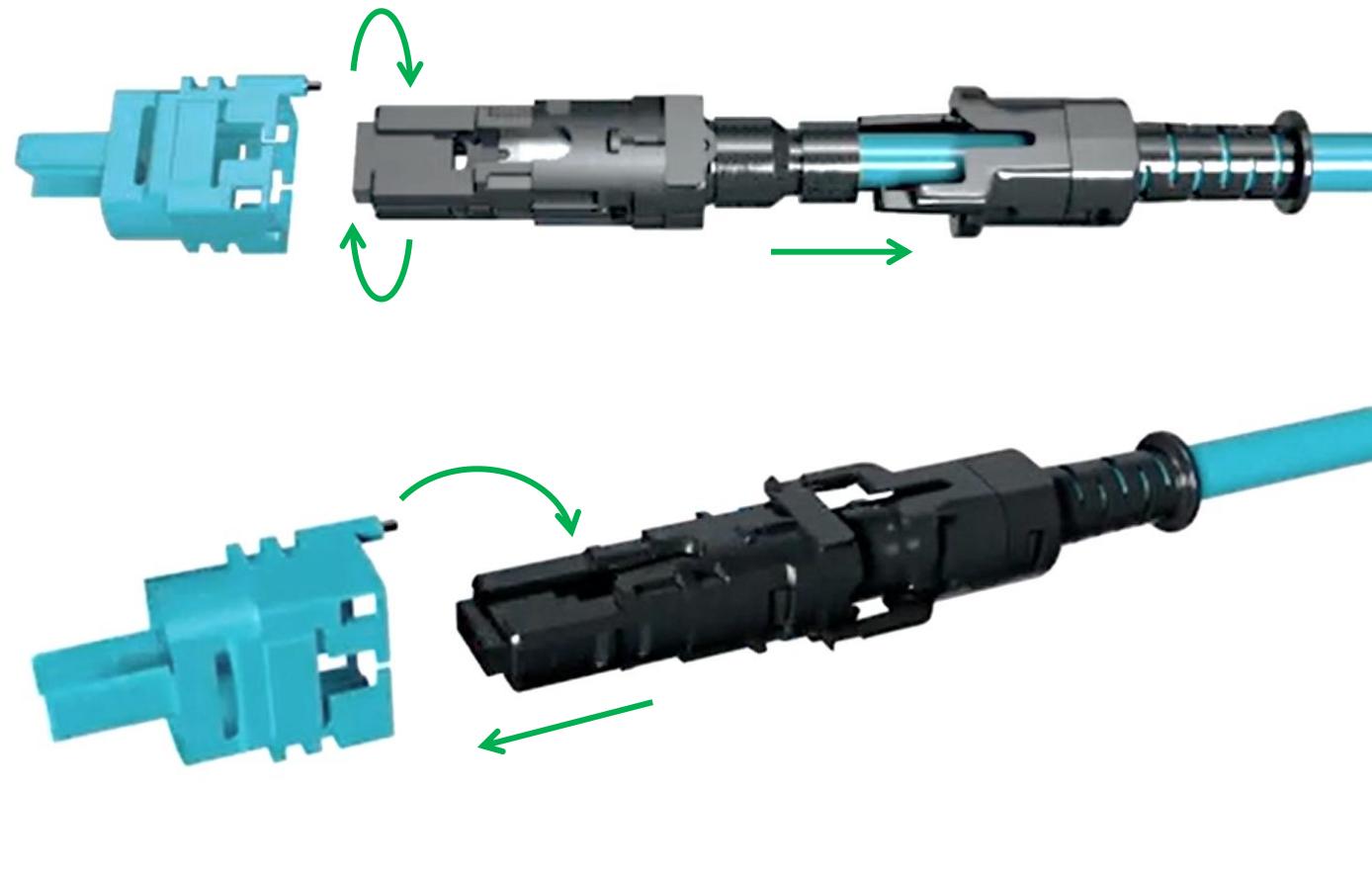
La corrección de la dispersión cromática reduce las retransmisiones y la perdida de paquetes, mientras aumenta el desempeño y la distancia que alcanza la fibra óptica multimodo.

Conexiones Base 8 y 16 con conexiones MPO

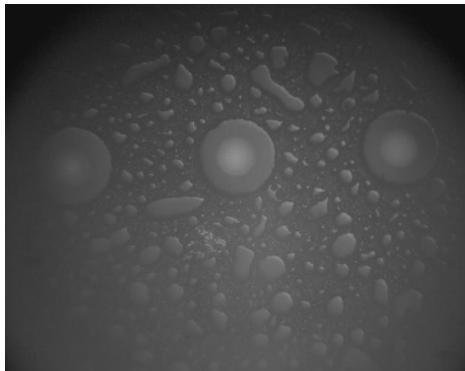
	AI	Hyperscale	Enterprise
Servers	GPU	CPU	CPU
Server Links	GPU to GPU & Leaf Switch	CPU to Leaf	CPU to ToR
Server Connectivity	MPO APC	Duplex LC	Duplex LC
Switch Links	Leaf to Spine	Leaf to Spine	TOR to Access to Aggregation to Core
Switch Data Rate	800G	800G	100G
Protocol	Infiniband / Ethernet	Ethernet	Ethernet
Fiber	Multimode	Singlemode	Multimode
Cabinet Loading	3K – 4K lbs dynamic	3K lbs dynamic	2.5K lbs dynamic
Power / Rack	40 - 120 KW	10 - 20 KW	5-15 KW
PDU	3-phase 415v 60-100A	3-phase 208v 60A	Single-phase 120v 15A
Power / Site	200 - 1,000 MW	50 -300 MW	1-5MW
Cooling	Liquid / Air	Air + some liquid	Air
Location	Easy to get power	Near large cities	On-premise

Conexiones Base 8 y 16 con cambio de polaridad y genero

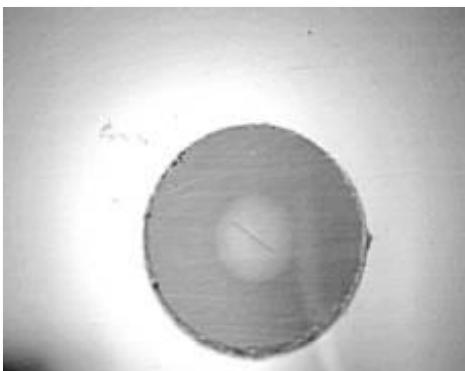
Prevea la infraestructura para soportar **100 - 200 - 400 - 800 Gb**. Cambie la polaridad y género de los conectores en sitio, **eliminando nuevos cables** particulares, reservar amplios stocks o esperar el aprovisionamiento.



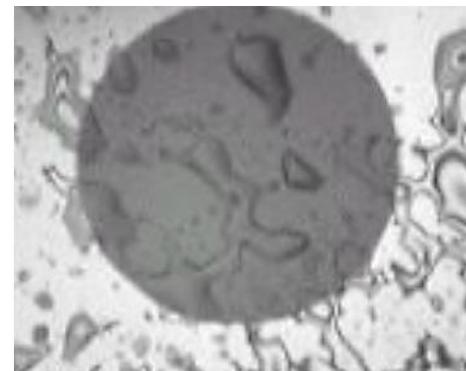
La limpieza para el desempeño de la fibra



Cara del extremo MPO



Rayón en el núcleo



Huellas digitales



Contaminación suelta

La limpieza es clave

El 85% de los problemas relacionados con la fibra se deben a falta de limpieza en los diferentes puntos de conexión (conectores de cables, cassettes, adaptadores y equipos).



Solución para limpieza



Toallitas limpiadoras



Hisopos



Limiador de conectores MPO hembra y macho



Limiador de fibra para cassette MPO

¿Su infraestructura es amiga o enemiga?

La infraestructura de red es para su negocio: ¿Alto rendimiento o una selección en función del costo?.



Lo que necesita...



Corto ciclo de vida de la inversión



...lo que instala.

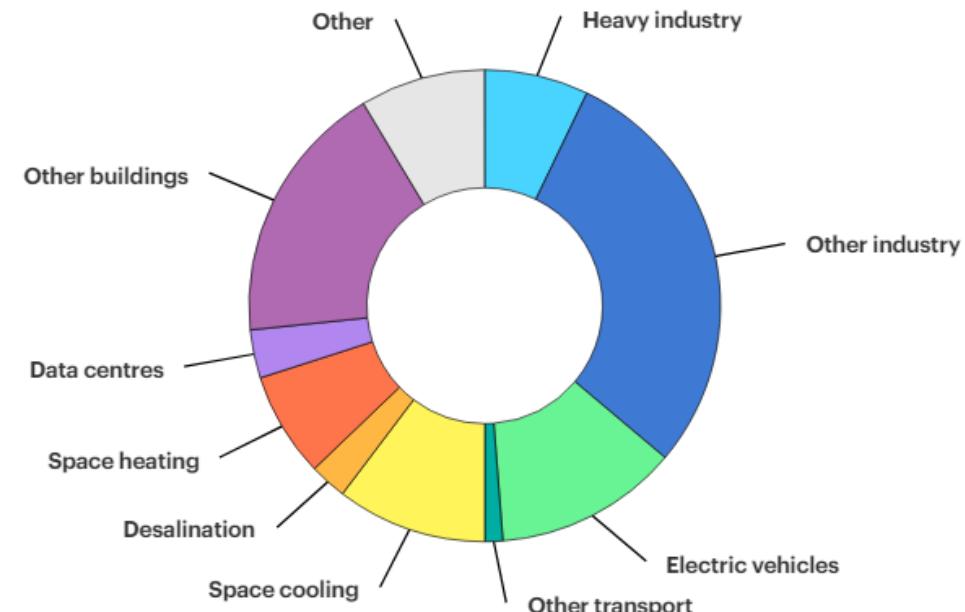
La sostenibilidad, el desafío de la IA

Según la EIA, hoy los Data Center no son los grandes consumidores a nivel global, pero con la IA, multiplicarán por 10 la demanda por sitio, de 5-10 MVA a 100 MVA.

En NA, China y Europa, los DC representan entre el 2% - 4% del consumo total, sin embargo, se concentran espacialmente impactando las comunidades cercanas. En cinco estados de USA los DC ya ha superado el 10% del consumo eléctrico, y en Irlanda, el 20%.

Mientras, el cambio climático ha llevado a priorizar la reducción de la huella de carbono buscando cero emisiones (carbono neutro) de manera directa o vía bonos.

El Data Center debe considerar todas las oportunidades de ser más eficiente, en el uso de recursos (energía, agua, combustibles), el acondicionamiento térmico y conectividad.



Crecimiento mundial de la demanda final de electricidad por uso en el escenario de políticas establecidas, 2023-2030



Panduit World Headquarters. Tinley Park ILL. USA
LEED Gold USGBC

La sostenibilidad está en nuestro ADN

La sostenibilidad está transformando todas las industrias, es parte fundamental de las **Políticas de Responsabilidad Social Empresarial** y las directivas de reducción de huella de carbono de las organizaciones y gobiernos.

Por esto, **la sostenibilidad no es un plus, es requisito básico para el desarrollo de proyectos para edificios, infraestructura, data center e industria.**

Muchos hablan de sostenibilidad por estos días. **En Panduit, hacemos más que solo hablar,** ofrecemos soluciones totales responsables y sostenibles con la sociedad y el medio ambiente.

¡Estamos conectados!



Vladimir Linares

Technical Systems Engineer
Panduit Corp.
RCDD BICSI - CCRD ICREA
vladimir.linares@panduit.com
[vladimir-linares-rcdd-ccrd-5234a0134](tel:vladimir-linares-rcdd-ccrd-5234a0134)



Guillermo Ortegón A.

Territory Account Manager
Panduit Corp.
ATS - Uptime Institute
guillermo.ortegon@panduit.com
[in guillermo-ortegon](https://www.linkedin.com/in/guillermo-ortegon/)

A modern conference room with large windows on one side, a long rectangular table in the foreground, and a whiteboard on an easel in the background. The room is well-lit and appears to be set up for a meeting or presentation.

**Gracias por
acompañarnos.**