HPE JUNIPER networking

Descripción general de la solución

Redes para centros de datos de IA abiertas y basadas en Ethernet

Redes para la IA de extremo a extremo, multiproveedor y basadas en la intención para lograr velocidad, eficiencia y ahorro

HPE

Desafío

Implementar clústeres de IA es complejo, requiere mucho tiempo y es costoso. Las empresas que invierten en aplicaciones de IA tienen recursos limitados para implementar centros de datos de IA y optimizar recursos de GPU costosos sin dependencia del proveedor.

Solución

La solución para centros de datos de IA de HPE Juniper Networking es una forma rápida y flexible de implementar clústeres de almacenamiento, inferencia y entrenamiento de IA de alto rendimiento, y la más sencilla de manejar con recursos de TI limitados.

Ventajas

— Las operaciones simplificadas de Día 0/1/2+ utilizan menos recursos y ahorran tiempo y dinero.

— Flexibilidad abierta para diseñar redes independientes de la GPU utilizando tecnologías y productos probados que evitan la dependencia del proveedor con un rendimiento comparable al de InfiniBand.

— Las soluciones validadas de extremo a extremo llave en mano garantizan confianza en la elección de productos con tiempos de implementación acelerados.

Las tecnologías y los diseños de los centros de datos tradicionales no satisfacen los exigentes requisitos de rendimiento, capacidad y latencia que las cargas de trabajo de IA imponen a la infraestructura. Los centros de datos de IA, caracterizados por un entrenamiento especializado de back-end y diseños de estructura de inferencia de front-end, requieren soluciones de vanguardia en computación, almacenamiento y redes. En medio de estos avances, se ha producido un aumento en la demanda de unidades de procesamiento de gráficos (GPU) potentes y una dependencia de redes InfiniBand patentadas y de una sola fuente. Lamentablemente, esto ha provocado un aumento de los precios y cuellos de botella en la cadena de suministro. Reducir costes, a la vez que se impulsa la innovación y la flexibilidad, es un imperativo para las empresas. Al reconocer la necesidad de optimizar el rendimiento de la GPU y minimizar los costes, las empresas recurren cada vez más a Ethernet como la alternativa de red abierta preferida para los centros de datos de IA.

El desafío

En múltiples sectores, la IA está abandonando los laboratorios de ingeniería para entrar en las sala de juntas en forma de activo, y los arquitectos de la IA y sus homólogos de la TI se encuentran, de manera repentina, explorando nuevos territorios. Para aprovechar al máximo el valor de la IA, primero deben gestionar las complejidades de su implementación y lidiar con las realidades de sus costes y plazos de entrega.

Con pocas opciones prácticas de GPU y redes InfiniBand patentadas, los retrasos en la cadena de suministro y el bajo inventario han aumentado los costes de una infraestructura de IA costosa y poco común. Una sola GPU, por ejemplo, cuesta más de 30 000 USD y los servidores de IA que utilizan estos chips pueden costar más de 400 000 USD. Con un precio tan alto, incluso los clústeres de GPU pequeños pueden costar millones, mientras que los clústeres de GPU grandes pueden costar cientos de millones de dólares.

Como el motor que impulsa la IA, las redes del centro de datos desempeñan un papel fundamental en la interconexión y maximización de las GPU. Reducir el tiempo de finalización de tareas (JCT, por sus siglas en inglés), es decir, el tiempo que lleva completar el entrenamiento de la IA, es la clave para lograr una mayor velocidad y ahorrar costes.

La adopción exitosa de la IA también depende de la rápida respuesta del mercado a la demanda. Para aliviar la dependencia de soluciones de un solo proveedor, el sector ha tomado medidas importantes para fomentar un mercado abierto y competitivo con diversidad de GPU y la tecnología L2 con mayor implementación en el mundo: Ethernet.

Soluciones para centros de datos de IA de Juniper

Las soluciones para centros de datos de IA de Juniper ofrecen una forma rápida y flexible de implementar clústeres de almacenamiento, inferencia y entrenamiento de IA de alto rendimiento, y la más sencilla de operar con recursos de TI limitados. Con operaciones únicas basadas en la intención, un asistente de red virtual nativo de IA y diseños validados, HPE Juniper Networking simplifica el diseño, la implementación y la resolución de problemas de las redes para el centro de datos de IA. La solución también ofrece una flexibilidad insuperable a los clientes, evitando la dependencia del proveedor con diversidad de silicio, gestión de conmutadores de múltiples proveedores y un compromiso con estructuras Ethernet abiertas y basadas en estándares.

Operaciones más sencillas

Las operaciones simplificadas de Día 0/1/2+ utilizan menos recursos y ahorran tiempo y dinero. El software HPE Juniper Networking ofrece una solución de automatización del centro de datos de múltiples proveedores con redes basadas en la intención, que acelera el tiempo de implementación y solución de problemas al evitar los desafíos de la configuración manual. Por ejemplo, las funciones de equilibrio de carga avanzadas, notificación de congestión explícita (ECN) y control de flujo prioritario (PFC) esenciales para redes de IA de alto rendimiento y sin pérdidas se pueden aplicar en toda la estructura mediante Juniper Apstra Data Center Director sin escribir un solo comando CLI. La visualización del flujo de red en tiempo real de Apstra Data Center Director identifica los puntos críticos de congestión del centro de datos con IA con mapas de calor y análisis basados en la intención para una rápida resolución de problemas.

Como única fuente de verdad con garantía de circuito cerrado, Apstra Data Center Director ofrece análisis potentes para predecir y solucionar problemas de manera proactiva y evitar disrupciones del servicio. Con proyectos basados en plantillas, Apstra Data Center Director ofrece fiabilidad, coherencia y repetibilidad que se traducen en una reducción del OpEx del 90 % e implementaciones que se realizan a una velocidad un 85 % mayor.

A través de la integración de Mist AI con Apstra Data Center Director, HPE Juniper Networking lleva AIOps al centro de datos, para proporcionar automatización e información adicionales que permiten optimizar las experiencias del centro de datos. Marvis Virtual Network Assistant para el centro de datos extiende la rica telemetría de transmisión y la supervisión en tiempo real de Apstra Data Center Director al panel de Marvis para ofrecer solución de problemas y análisis proactivos en todos los dominios operativos. Esta integración es un primer paso fundamental para proporcionar información procesable impulsada por la IA en el futuro para centros de datos de múltiples proveedores. Los operadores de campus y sucursales que usan Marvis ahora pueden ver los problemas del centro de datos en el mismo panel.

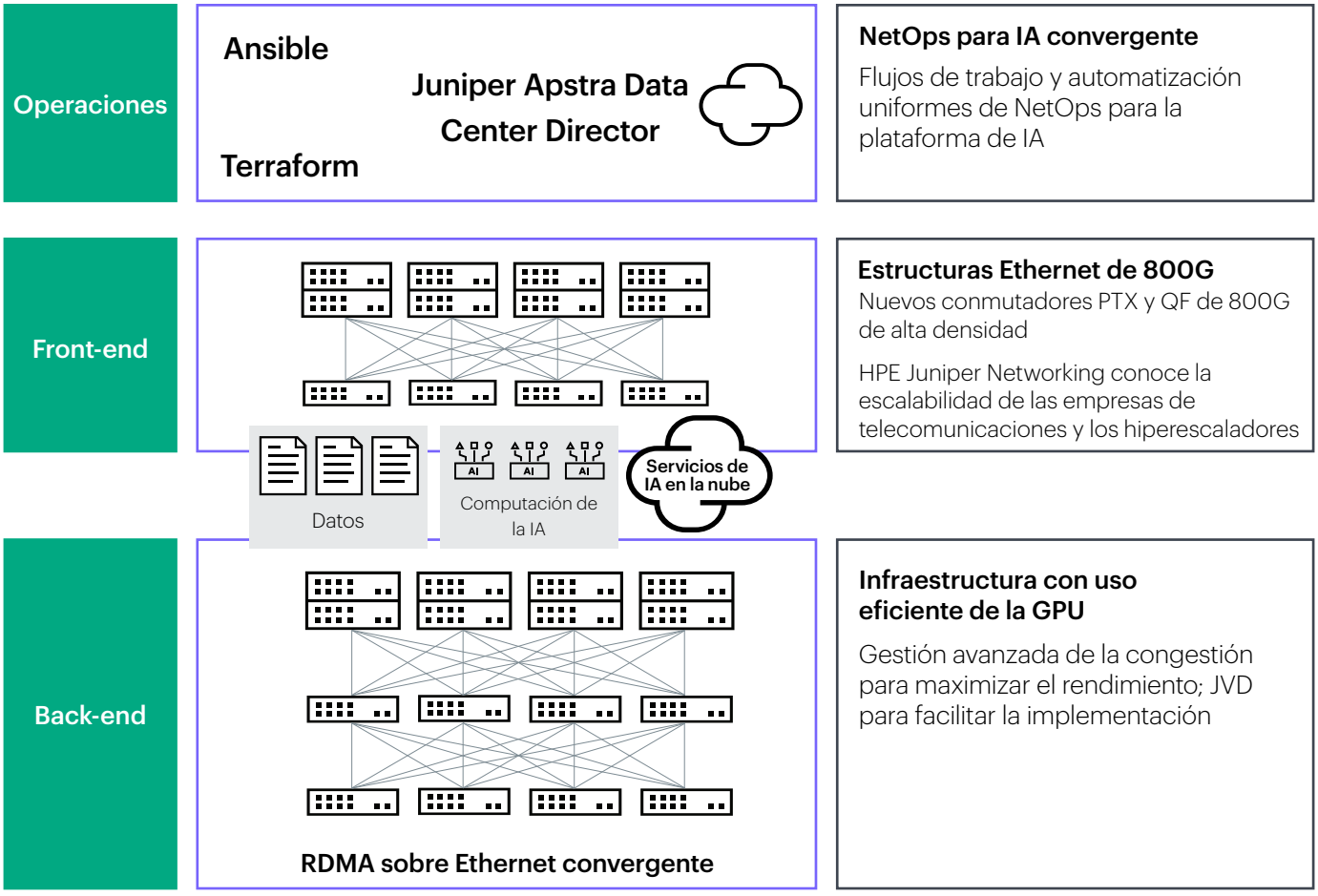

Figura 1. Soluciones de redes para centros de datos de IA de Juniper Networks

Flexibilidad abierta

Las soluciones Ethernet abiertas y flexibles te permiten utilizar tecnologías y productos probados que evitan la dependencia del proveedor. Las nuevas estructuras para centros de datos leaf-spine de 800 GbE de grado elevado y optimizados para la IA de HPE Juniper Networking ofrecen capacidades avanzadas de gestión de tráfico para garantizar rendimiento de alto ancho de banda, baja latencia, sin pérdidas y escalable a través de Ethernet. Con una pasarela hacia los 1,6 terabits y soporte de múltiples proveedores para sistemas independientes de la GPU, Ethernet reduce costes, permite una innovación más rápida, maximiza la flexibilidad del diseño y evita problemas en la cadena de suministro. La compatibilidad con el acceso directo a memoria remota a través de Ethernet de convergencia v2 (RoCEv2) convierte a las estructuras Ethernet de HPE Juniper Networking en la única solución multiproveedor para la gestión y automatización de estructuras de centros de datos optimizadas para el entrenamiento de la IA de back-end y los modelos de inferencia de front-end.

Soluciones llave en mano

Las soluciones validadas de extremo a extremo y llave en mano garantizan confianza en la elección de productos para agilizar las implementaciones de centros de datos. Cuando las empresas tienen a su disposición soluciones de red de extremo a extremo, pueden construir centros de datos de IA de alto rendimiento con flexibilidad y facilidad. La gama de enrutadores de la serie PTX de Juniper para centros de datos se ha ampliado con el nuevo enrutador fijo PTX10002-36QDD de 800 GbE de alta densidad y nuevas tarjetas de línea de 800 GbE para el chasis PTX10000. Basado en el silicio personalizado Express 5, el chasis PTX10000 más grande admite ahora hasta 576 puertos de 800 GbE para arquitecturas spine y superspine de grado elevado. Además, una nueva plataforma fija QFX5240 de 800 GbE y alta densidad basada en el último ASIC Broadcom Tomahawk 5 ofrece diversidad de silicio para centros de datos de IA de alto rendimiento, escalables y energéticamente eficientes.

Los conmutadores de la serie QFX y los enrutadores de la serie PTX de Juniper ofrecen un rendimiento, una escalabilidad y un control de la congestión excepcionales, lo que garantiza un funcionamiento máximo de los clústeres de IA. En las pruebas de referencia MLPerf™, el Ethernet optimizado para la IA de HPE Juniper Networking registro unos tiempos de finalización de tareas para cargas de trabajo de entrenamiento al mismo nivel que InfiniBand. De este modo, los clientes tienen a su disposición una alternativa de estructura abierta y de alto rendimiento que evita la dependencia del proveedor. Para simplificar las implementaciones de IA, HPE Juniper Networking está presentando nuevos diseños validados por Juniper (JVD, por sus siglas en inglés). Los JVD para la IA permiten estructuras CLOS prescritas, optimizadas para rieles o de multicapa para ofrecer una solución para centros de datos completa, que incluye conmutación, operaciones y seguridad. Los JVD, rigurosamente probados, validados y documentados por profesionales del laboratorio de HPE Juniper Networking, se pueden utilizar como diseños listos para usar o como un conjunto de directrices para reducir el riesgo y dimensionar y presupuestar adecuadamente cada clúster de IA.

Características y ventajas

Operaciones simplificadas: utiliza menos recursos y ahorra tiempo y dinero con operaciones simplificadas de Día 0/1/2+. Las redes basadas en la intención aceleran el tiempo de implementación y solución de problemas al evitar los desafíos de la configuración manual (en cualquier hardware del proveedor).

— Fuente única de verdad con garantía de circuito cerrado: predice y soluciona problemas de forma proactiva para evitar disrupciones del servicio

— AIOps for Data Center ofrece automatización e información adicionales para optimizar las experiencias del centro de datos

Flexibilidad abierta: diseña tu red utilizando tecnologías y productos probados que eviten la dependencia del solo proveedor.

— Las capacidades avanzadas de gestión de tráfico garantizan un rendimiento de alto ancho de banda, baja latencia, sin pérdidas y escalable a través de Ethernet

— El aprovechamiento de Ethernet reduce los costes, permite una innovación más rápida, maximiza la flexibilidad del diseño y evita los problemas en la cadena de suministro

— La solución multiproveedor para la gestión y automatización de estructuras de centros de datos maximiza la flexibilidad del diseño

— Interoperable con todas las GPU, estructuras y conmutadores

Soluciones llave en mano: las soluciones validadas de extremo a extremo proporcionan confianza en la elección de productos y aceleran los tiempos de implementación.

— La diversidad del silicio impulsa la escala, el rendimiento y la flexibilidad. Compatibilidad con Broadcom e interconexión de silicio personalizada para miles de GPU a las velocidades más altas (800 Gbps) y admite todas las configuraciones de estructuras de centros de datos

— Portfolio de seguridad del centro de datos para una protección de extremo a extremo

— Los JVD ofrecen una solución completa de centro de datos, que incluye conmutación, operaciones y seguridad

— El amplio portfolio de conmutadores de HPE Juniper Networking (con diversidad de silicio) impulsa la escala y el rendimiento

— Portfolio integral de seguridad para el centro de datos con prevención de amenazas unificada para garantizar protección del centro de datos de confianza cero

Componentes de la solución

Apstra Data Center Director: la solución de automatización del centro de datos de HPE Juniper Networking permite a los operadores automatizar todo el ciclo de vida de la red del centro de datos en entornos de múltiples proveedores, desde el diseño hasta las operaciones continuas. Apstra Data Center Director proporciona validación constante, una fuente única de verdad, análisis potentes e identificación y resolución rápidas de la causa raíz.

Conmutadores de la serie QFX: nuestra línea de conmutadores ofrece un rendimiento y una escalabilidad superiores, una pila de enrutamiento integral, la programabilidad abierta del sistema operativo Junos® y un conjunto amplio de capacidades de estructura IP y EVPN-VXLAN.

Enrutadores serie PTX: los enrutadores de la serie PTX forman la base de las arquitecturas de núcleo y WAN más grandes del mundo. Equipados con la última generación de ASIC de la familia Express, los enrutadores están optimizados para la nube para permitir migraciones fluidas a 800 GbE. Ofrecen simultáneamente la escala, la flexibilidad y la protección de la inversión necesarias para satisfacer las necesidades del mercado actual.

Asistente de red virtual Marvis para centro de datos: Marvis ofrece un asistente de red virtual nativo de la IA para el centro de datos, que ofrece información excepcional durante todo el ciclo de vida del centro de datos en cualquier hardware de cualquier proveedor. Además, Marvis proporciona visibilidad y seguridad de extremo a extremo en todos los dominios empresariales, desde el campus y la sucursal hasta el centro de datos.

Evolución del sistema operativo Junos: este sistema operativo de red unificado de extremo a extremo proporciona fiabilidad, agilidad y programabilidad abierta para implementaciones exitosas a escala de la nube.

Operaciones simplificadas, flexibilidad abierta y redes llave en mano

Prácticamente todas las empresas están definiendo una estrategia sobre cómo utilizar mejor la infraestructura del centro de datos para ejecutar iniciativas de IA corporativa. Las soluciones para centros de datos de IA de Juniper constituyen una forma rápida y sencilla de ofrecer redes escalables y de alto rendimiento para el entrenamiento y la inferencia de la inteligencia artificial. Con operaciones basadas en la intención y multiproveedor, un asistente de red virtual nativo de la IA y diseños validados por Juniper, HPE Juniper Networking elimina la complejidad de las redes para centros de datos de IA.

Más información en

HPE.com/Juniper-AI-Data-Center

Visita HPE.com

Iniciar chat ahora

© Copyright 2025 Hewlett Packard Enterprise Development LP. La información que contiene este documento está sujeta a cambios sin previo aviso. Las únicas garantías de los productos y servicios de Hewlett Packard Enterprise figuran en las declaraciones expresas de garantía incluidas en los mismos. Nada de lo que aquí se indica debe interpretarse como una garantía adicional. Hewlett Packard Enterprise no se responsabilizará de los errores u omisiones técnicos o editoriales que pudiera contener el presente documento.

MLPERF™ es una marca comercial y marca de servicio de MLCommons Association en Estados Unidos y otros países. Todas las marcas de terceros son propiedad de sus respectivos titulares.

a00150898ESE

HEWLETT PACKARD ENTERPRISE

hpe.com