JUNIPER AI DATA CENTER DEPLOYMENT SERVICES DATASHEET

Overview

Data centers built to run AI workloads have requirements distinctly different from non-AI data centers. If network inefficiencies delay job completion times (JCTs), the time and resource costs to train your AI models multiply. Juniper AI Data Center (AIDC) Deployment Services help you avoid that problem.

By delivering leading-edge deployment and production optimization expertise and assistance, our AIDC Deployment Services help you achieve enhanced network performance, smoother operations, and unlock the full potential of your AI workloads. We tailor these services to your specific AI training model to accelerate time to value and maximize ROI.

Description

Juniper AIDC Deployment Services provide a turnkey solution to deploy, monitor, and optimize key network performance metrics for your AIDC implementation. The services are based on Juniper Validated Designs (JVDs) for AIDC deployments using Ethernet, ensuring top-quality deployments that are compatible with GPU types from NVIDIA (fixed services) and all vendors (custom service). The services include one each of up to three types of greenfield networks: frontend compute, backend compute, and backend storage with Juniper AIDC-qualified devices.

The services have two phases: deployment and production optimization. The deployment phase ends with network acceptance validation followed by a 30-day production optimization phase.

The services are available with or without Apstra. For customers with Juniper Apstra, AI Data Center Deployment with Apstra creates blueprints for all AIDC fabrics and then auto-provisions all devices. This saves time and eliminates potential operator errors during the deployment phase by automating the high-level design (HLD) and low-level design (LLD) processes. AI Data Center Deployment—the standard deployment service without Apstra—requires the manual creation of an HLD and LLD prior to actual deployment.

Both services begin with a workshop run by the Juniper project manager and consultant engineer who will collaborate with the customer to develop a mutual understanding of the overall requirements, deployment plan, and outcomes for the production optimization phase. This leads to the design and actual deployment of the fabrics.

Once all devices are deployed, the GPU network interface cards will be configured and validated to use Ethernet by Juniper. The final step in the deployment phase is acceptance testing and validation, in which Juniper will run continuous collective communications library functions to achieve high bandwidth and low latency across the GPU fabrics.

After the networks have been turned over to the customer, Juniper experts continue to engage for the 30-day production optimization phase. During the 30 days, Juniper closely monitors key parameters and Ethernet tuning through multiple iterations of the customer's training model cycles, including deployment of trained models into the frontend inference network. This phase provides the customer with the key advanced analytics and final adjustments to Ethernet interfaces that minimize JCT at maximum bandwidth.
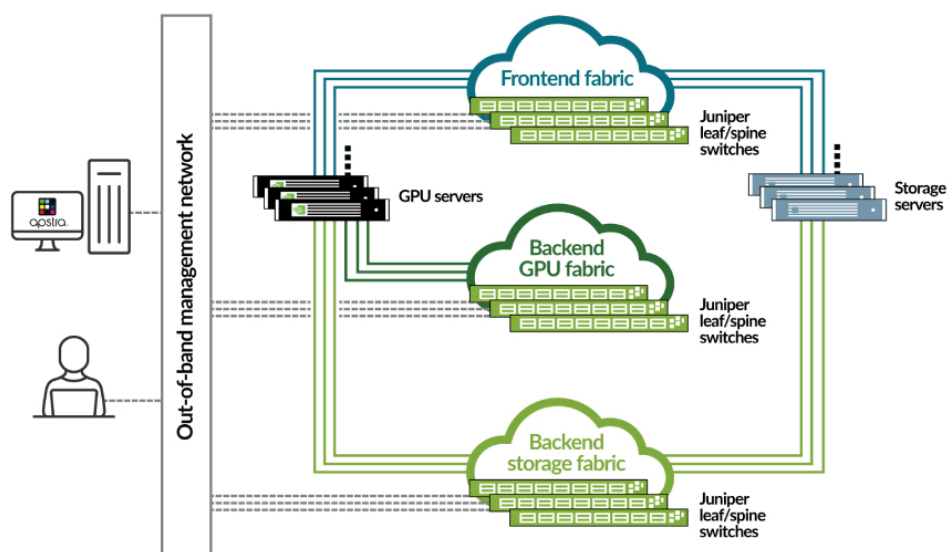
Features and Benefits

Table 1: Juniper fixed AI Data Center Deployment Services features and benefits

Key Features	Description	Benefit(s)
Solution workshop	Pre-deployment workshop with the customer to review all customer input data and personalize the reference design, as well as agree on the network performance outcomes of the production optimization phase	Deployment tailored to customer requirements. Greater oversight of entire process with agreement on objectives and metrics for deployment and production phases
Apstra platform deployment and three fabrics provisioning (with Apstra only)	Deploy the Apstra server. Configure blueprints for up to three network fabrics. Implement those fabrics with the newly installed network devices	Validated design. Rapid, simplified deployment for all devices in each fabric. Real-time visibility into the pre- and post-deployment fabrics
High-level design (HLD) and low-level design (LLD) (without Apstra only)	Create HLD and then LLD and have customer approve for implementation	Get visibility to and provide approval of the detailed design to help ensure its alignment with your requirements before deployment
Network implementation plan execution (without Apstra only)	Following the creation of the network implementation plan, full deployment of network fabrics (up to three) per the approved LLD by Juniper experts incorporating best practices	Predictable rollout with reduced risk and minimal disruption
GPU NIC configuration and validation	Configure and validate the ConnectX GPU NIC cards for Ethernet operation (instead of default Infiniband) across the GPU cluster fabric	Ensure operation on lower cost Ethernet fabrics with engineers experienced in NIC reconfiguration
Acceptance validation testing	Run AI fabric stress testing using NVIDIA Collective Communications Library (NCCL). Monitor and tune the Ethernet fabric for zero packet loss and maximum bandwidth utilization	Increased confidence in the production readiness of the network and its ability to run at maximum speed and with minimal loss. Achieve the outcomes set during the solution workshop
Production optimization	Once the Ethernet fabrics are in production, run multiple iterations of the customer's training model. Collect GPU performance and utilization data, review it with the customer, and provide additional tuning recommendations to optimize JCT	Ensure optimal performance and JCTs for training model in production
Knowledge transfer workshop	Review of network fabrics deployed and advanced operations of Juniper Apstra	Enable the operations team to sustain optimal network performance and JCTs efficiently

How to Order

Juniper AI Data Center Deployment Services are available globally. For details, please contact your local Juniper account team, local Juniper partner, Juniper field sales manager, or assigned Juniper service business manager.

For additional details such as scope, deliverables, eligibility, and exclusions, please refer to the corresponding Service Description Document: https://support.juniper.net/support/guidelines/

About Juniper Networks

Juniper Networks believes that connectivity is not the same as experiencing a great connection. Juniper's AI-Native Networking Platform is built from the ground up to leverage AI to deliver exceptional, highly secure, and sustainable user experiences from the edge to the data center and cloud. Additional information can be found at juniper.net or connect with Juniper on X (formerly Twitter), LinkedIn, and Facebook.

Corporate and Sales Headquarters

Juniper Networks, Inc.

1133 Innovation Way

Sunnyvale, CA 94089 USA

Phone: 888.JUNIPER (888.586.4737)

or +1.408.745.2000

www.juniper.net

APAC and EMEA Headquarters

Juniper Networks International B.V.

Boeing Avenue 240 1119 PZ Schiphol-Rijk

Amsterdam, The Netherlands

Phone: +31.207.125.700