Juniper® Validated Design

AI Data Center Network with Juniper Apstra, NVIDIA GPUs, and WEKA Storage

sol-overview-JVD-AICLUSTERDC-AIML-02-08

Executive Summary

Building infrastructure for AI services presents unique challenges, even as many traditional data center networking principles still apply. The use of high-performance Graphics Processing Units (GPUs) for AI training and inference, coupled with the need to process massive amounts of data, puts a strain on the network infrastructure. The network must be able to handle large traffic flows, while minimizing latency and packet loss, maximize connectivity, and as a result, enable acceptable workload completion times. To meet these requirements, it is necessary to deploy the appropriate hardware and software features and along with the proper configuration.

Juniper Networks solutions stand out in supporting AI networks by not only meeting AI workload requirements but also offering a scalable, standards-based, automated approach for various cluster sizes. Juniper's solution assures the fastest innovation, maximizes design flexibility, and prevents vendor lock-in for back-end, front-end, and storage AI networks. Advanced JVD testing by Juniper subject matter experts provides optimum designs to deliver high performance, increased reliability, and simplified troubleshooting.

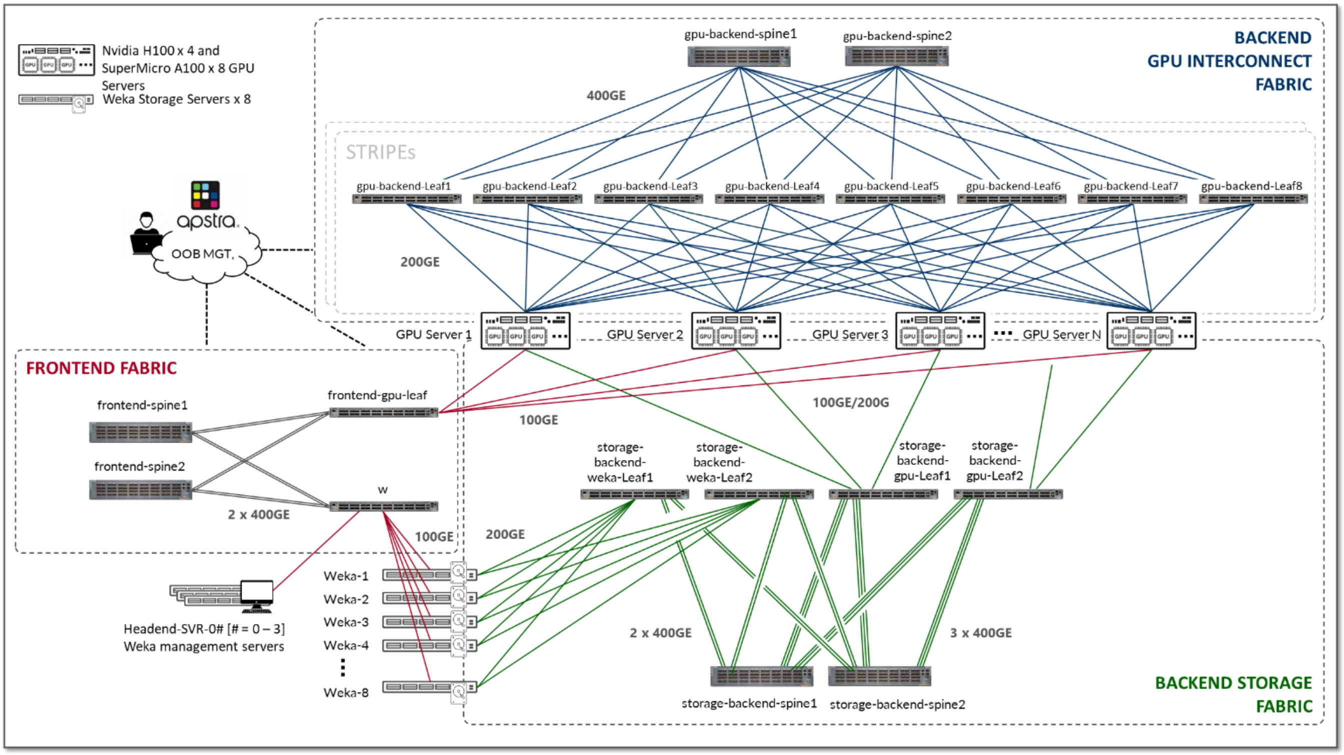
Solution Overview

The **AI Data Center Network Design with Juniper Apstra, NVIDIA GPUs, and WEKA storage solution** outlines best practices and validated solution components and configuration guidelines to successfully deploy a network infrastructure for an AI cluster of NVIDIA GPUs and WEKA storage. The solution is based on Juniper Network QFX Series high-performance switches and PTX Series Routers and leverages the intent-based capabilities of Juniper Apstra to automate and validate the deployment. Customers can orchestrate the deployment systematically without requiring previous in-depth knowledge of the products and technology.

The solution consists of 3 blocks:

- **Front-end Fabric**: provides access to the AI systems from the tools and applications that allow users to interact with the system, initiate training or inference workloads, and visualize the results.
- **GPU Interconnect Back-end Fabric**: provides the infrastructure for the GPUs performing the AI/ML tasks to communicate with each other.
- **Storage Back-end Fabric**: provides the infrastructure for storage devices to be accessible from the GPU servers.

Figure 1: AI Data Center Network Design with Juniper Apstra, NVIDIA GPUs, and WEKA storage solution



Packet loss and latency can delay AI jobs and worsen the user experience. A key goal in designing the back-end for an AI/ML cluster is to create a network that prevents data loss. This network should allow GPUs to communicate and access storage without interference. It should also maximize throughput and minimize latency for AI/ML traffic flows.

The back-end fabrics of this solution were designed with these goals in mind. The **Storage back-end fabric** follows a 3-stage IP Clos architecture. The GPU Interconnect Back-end fabric follows a 3-stage IP clos architecture and NVIDIA's Rail Optimized Stripe Architecture.

A **Rail Optimized Stripe Architecture** provides efficient data transfer between GPUs, especially during computationally intensive tasks such as AI/ML Large Language Models (LLM) training workloads, where seamless data transfer is necessary to complete the tasks within a reasonable timeframe. A Rail Optimized topology aims to maximize performance by providing minimal bandwidth contention, minimal latency, and minimal network interference, ensuring that data can be transmitted efficiently and reliably across the network.

To ensure uninterrupted AI/ML traffic flows, the network devices are configured to handle congestion using DCQCN (Data Center Quantized Congestion Notification), also referred to as RCM (RoCEv2 Congestion Management). DCQCN has become the industry-standard end-to-end congestion control method for RDMA over Converged Ethernet (RoCEv2) traffic. DCQCN combines two different mechanisms for flow and congestion control:

- Priority-Based Flow Control (PFC), and
- Explicit Congestion Notification (ECN).

By employing these techniques, DCQCN can balance between reducing traffic rates and stopping traffic flow at the source, alleviating congestion without resorting to packet drops.

Additionally, the solution takes into consideration the effect of load distribution across the different paths between the spine and leaf nodes, which due to the traffic characteristics may impede optimal link utilization. The default hash algorithms using Layer 3 header information often favor certain links due to limited flow variation. As a result, some links will become overutilized leading to congestion and potential packet loss. To address this issue, the design also includes **Dynamic Load Balancing (DLB)** on the leaf nodes.

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
Fax: +1.408.745.2100
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
Fax: +31.207.125.701