CyberFlood DPU and SmartNIC Assessment Solution

Benchmarking with AI/RDMA workloads

The rapid growth of AI/ML, cloud computing, and hyperscale data centers demands low-latency, high-bandwidth, and efficient resource use. As network functions move from dedicated hardware to software-driven virtualization, general-purpose CPUs struggle with tasks they weren't designed for, affecting performance, power efficiency, and cost-effectiveness.

Data Processing Units (DPUs) and SmartNICs offload critical infrastructure functions, enhancing performance and reducing CPU overhead. This improvement boosts the performance of complex AI/ML, storage, and advanced security workloads, enabling more efficient and scalable data centers.

To ensure data center operators realize the efficiency gains of DPUs, they must benchmark and tune configuration parameters to avoid bottlenecks, inefficient resource utilization, and unpredictable performance. This proactive approach ensures that data centers can handle increasing demands and remain cost-effective while delivering high performance.

Solution Overview

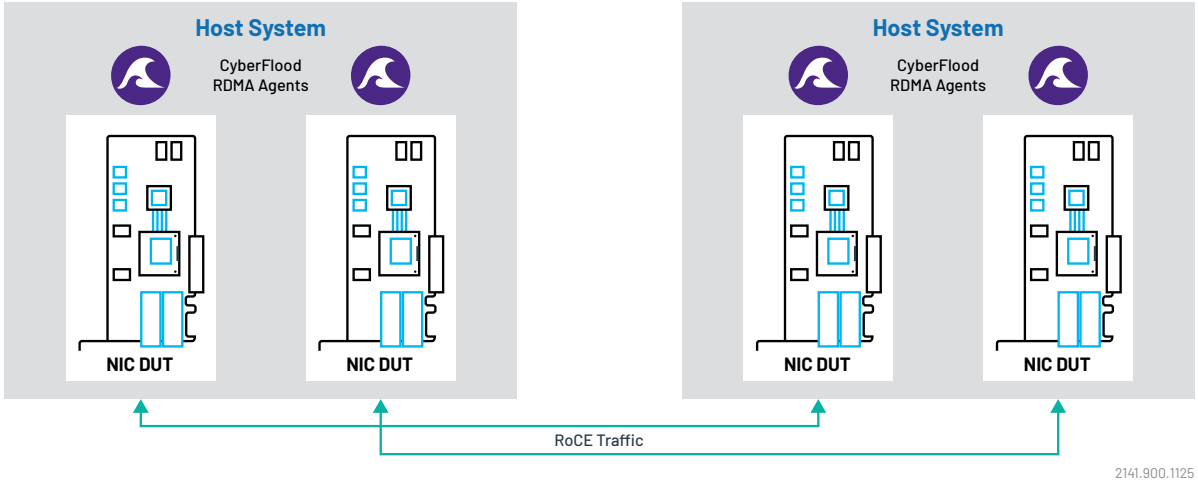
CyberFlood delivers stateful RDMA emulation to test and validate DPUs and SmartNICs with real-world AI and storage network workloads, ensuring scalability, low latency, and maximum throughput for next-generation AI applications.

Key Features & Benefits

- Emulates stateful RDMA over IPv4 and IPv6 transport to assess RDMA offload performance.
- Compare both RDMA over Converged Ethernet (RoCE) or InfiniBand (IB) transports.
- Performs RDMA READ and WRITE operations with Reliable Connection (RC) service type.
- Define QP Memory Regions (Single or Global Regions).
- Generate N:1 (in cast), and N:M, traffic patterns for congestion control validation.
- Provides user-configurable Q-Pair numbers and block sizes.
- Provides a flexible 'Job' function that combines multiple RDMA contexts and sequences RDMA operations to emulate complex AI workloads across multiple compute nodes, validating job completion times.
- Provides detailed real-time statistics and reporting, including the ability to expose statistics and counters specific to the NIC.
- Provides comprehensive REST API support for automation.

Lightweight CyberFlood software agents installed on the host compute establishes Q-Pair connections and performs RDMA READ and WRITE operations in Reliable Connection (RC) mode. CyberFlood provides complete flexibility for users to define the data transfer size, RDMA message size, and number of Q-pairs, and define custom jobs or workloads. In an AI cluster, a job or workload involves multiple GPUs working together to share and compute results. The 'Job' function in CyberFlood allows users to define an action list that combines multiple RDMA contexts. This makes it possible to emulate complex GPU cluster communication patterns and measure job completion times—rather than relying solely on message latency statistics. CyberFlood DPU assessment solution enables users to:

- Benchmark the performance of DPUs and SmartNICs with varying block sizes, MTUs, Q-pairs, and complex multi-node topologies
- Tune DPU and SmartNIC parameters such as Q-pairs, buffers, and flow control mechanisms to optimize configurations to suit the deployment environment to achieve the best performance
- Validate scaled scenarios with complex N:1, and N:M topologies to assess the impact of congestion on overall performance both with and without congestion control
- Ensure interoperability in complex multi-vendor switching and NIC environments
- Validate the server efficiency gains realized with RDMA offload capabilities to justify adoption and investment



Test traffic originates from the CyberFlood RDAM Agents residing on the host and terminates in another CyberFlood agent on a separate host system via the SmartNICs or DPUs connected to both the hosts. The NICs can also be connected via switching infrastructure as well.

The CyberFlood DPU and SmartNIC testing solution enables organizations to comprehensively evaluate the real-world impact of NICs and tune their configurations to support critical AI workloads, as they ensure high-performance, low-latency, and scalable GPU-GPU communications. Proactive testing and optimization can unlock the full potential of their AI infrastructure investment and drive more efficient model training and inference.

Technical Specifications

Product Feature	Description
RDMA Emulation	• Stateful Reliable connections • Synthetic Read & Write with variable payload size • Custom Job definition • Configurable queue pairs • N:1, 1:N, and N:M workload distribution • Dynamic Load Control – Bandwidth, Simulated User, and Simulated Users Per Second
Port-speed and Q-pairs	Port speed and Q-pair scale depend on the NIC capabilities
Statistics	• Bandwidth • Transactions – Attempted, Successful, Unsuccessful, Aborted • Jobs – Attempted, Successful, Unsuccessful • Job Completion Latency – Avg, Min, Max • RDMA Operations – Attempted, Successful, Unsuccessful, Latency • Custom Stats – NIC-specific counters and statistics
Agent Deployment Requirements	Minimum requirement: 1 x vCPUs per test agent and sufficient memory for the intended Q-pair scale

Ordering Information

Part Number	Description
CF-RDMA-AGENT-02-SUB	CyberFlood RDMA Emulation supports up to 2 test agents, 1YR SUB
CF-RDMA-AGENTL-04-SUB	CyberFlood RDMA Emulation supports up to 4 test agents, 1YR SUB
CF-RDMA-AGENT-08-SUB	CyberFlood RDMA Emulation supports up to 8 test agents, 1YR SUB
CF-RDMA-AGENT-16-SUB	CyberFlood RDMA Emulation supports up to 16 test agents, 1YR SUB

Requires CyberFlood Advanced Mixed Traffic license. Perpetual license options are also available.

Contact Us: +1 844 GO VIAVI | (+1 844 468 4284). To reach the VIAVI office nearest you, visit viavisolutions.com/contact

© 2025 VIAVI Solutions Inc. Product specifications and descriptions in this document are subject to change without notice. Patented as described at viavisolutions.com/patents

cyberflood-dpusmartnic-ds-hse-nse-ae
30194745 900 1225