

# Mango GPUBoost™ RDMA

Achieve Unparalleled Performance and Efficiency with Our Cutting-edge RDMA Solution



## EXECUTIVE SUMMARY

RoCEv2 (RDMA over Converged Ethernet v2) has gained attention for its unique performance advantages over traditional network protocols like TCP - maintaining infrastructure compatibility and offering exceptional scalability through the IP layer. As a result, many hyperscalers are opting to deploy RoCEv2-based RNICs in their data centers, which often host tens of thousands of GPUs.

However, existing RoCEv2-based RNICs often face a "connection scalability" issue as they struggle to maintain performance with large number of concurrent QPs (even with few hundreds of QPs). In large-scale environments where each RNIC handles thousands of QPs, this becomes a huge performance bottleneck. Moreover, these RNICs are "hard-wired," meaning they lack the flexibility to customize protocols to meet specific user requirements. This limitation leads to missed opportunities for optimizing performance for specific workloads or applications.

**Mango GPUBoost™ RDMA** addresses these challenges with an FPGA-based RoCEv2 RNIC. This solution provides greater connection scalability and customizability, enabling the management of thousands of QPs with tailored workflow for individual users.

## HIGHLIGHTS

### HIGH PERFORMANCE RDMA SOLUTION

- Line-rate performance (up to 2x100GbE) on each network port
- Higher connection scalability (maximum # of concurrent without perf. degradation)

### INTEROPERABILITY AND COMPATIBILITY WITH STANDARDS

- Compliant with major kernel distributions
- Interoperable with commercial RNICs and network switches

### TAILORABLE FOR ANY USER ENVIRONMENT

- Customized RoCEv2 to accelerate user applications with configurable headers
- Easy-to-use with the provided development kits

## SPECIFICATIONS

### USE CASES

- Large-scale AI/ML training system
- Disaggregated storage (NVMeoF)
- High Performance Computing (HPC)

### CAPABILITIES

- RoCEv2 (IBTA 1.4)
- 2048 Queue Pairs
- Hardware-based reliable transport
- Base Memory Management Extensions
- RoCE Congestion Management
- Priority Flow Control (PFC)
- GPU – RNIC peer-to-peer communication (compatible with AMD/NVIDIA GPUs)
- Support customized header

### INTERFACE CONFIGURATION

- PCIe Interface
  - > PCIe 3.0 x16 or PCIe 4.0 x8
- Ethernet Interface
  - > 2x 100GbE QSFP28
  - > Direct-attach copper or optical transceiver
- Portable Design
  - > AMD/Xilinx Alveo™ FPGAs (LP, FHHL)
  - > Intel/Altera Agilex™ FPGAs

### PERFORMANCE SUMMARY

NVMe over RoCEv2 \*

4KB (IOPS)	Random Read	5.4M
	Random Write	5.4M
	Random RW	8.4M
128KB (GB/s)	Sequential Read	23.0
	Sequential Write	22.8

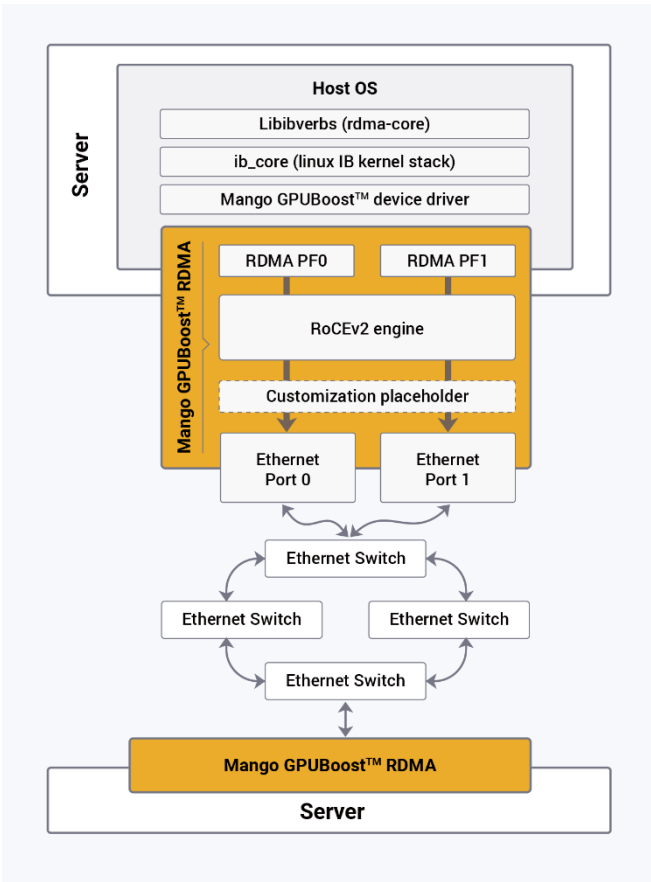
\* Tested with FIO, numjobs=32, iodepth=32. Used nullblk for target storage device

Collective communications †

rccl-test (GB/s)	All-reduce	91
	All-gather	90

† Tested with rccl-test on 4 GPU nodes, each with 8 MI300X GPUs and 8 RNICs. Bandwidth is aggregated across 8 RNICs. Data size is fixed at 16GB.

## DESIGN OVERVIEW



### RoCEv2 Acceleration

Mango GPUBoost™ RDMA integrates the RoCEv2 protocol with specialized hardware components, and it is compatible with existing software components such as rdma-core, and linux infiniband module. This ensures easy system integration of GPUBoost™ RDMA without requiring any application-level modifications. Designed to support line-rate bandwidth for each network port, Mango GPUBoost™ RDMA is well-suited for high-performance scenarios such as large-scale AI/ML workloads or HPC applications.

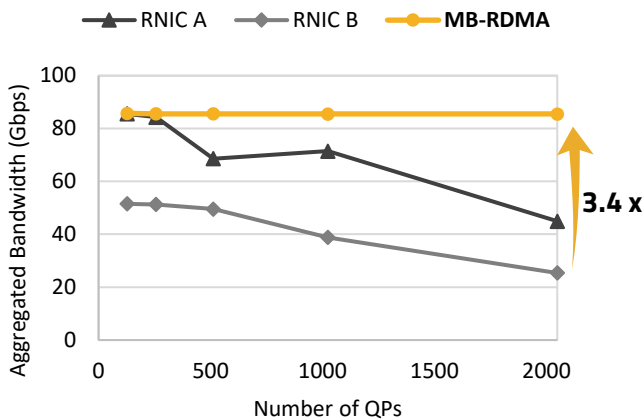
### Customized RNIC solution

Beyond its high-performance communication capabilities, RNICs provide a new opportunity to offload compute-intensive tasks or extend the RoCEv2 protocol for custom use. However, most companies offering RNIC solutions tend to have limited flexibility to meet diverse demands. Unlike existing solutions, FPGA-based Mango GPUBoost™ RDMA provides a placeholder that allows users to easily deploy customized processing and transmit necessary metadata via RoCEv2 extended headers. This high-level of available customization enables additional acceleration of applications.

## EVALUATION RESULTS

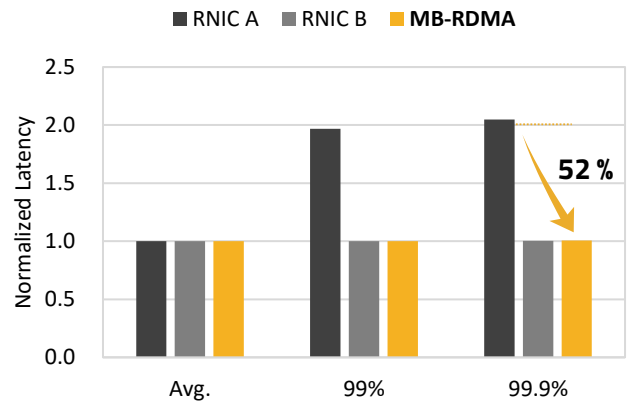
### Connection scalability: Bandwidth

Perftest, RDMA READ, 512B, # of threads = 8



### Connection scalability: Latency

Perftest, RDMA READ, 128KB, # of threads = 32, # of QPs = 32



#### DISCLAIMERS

The performance claims in this document are based on the internal cluster environment. Actual performance may vary depending on the server configuration. Software and workloads used in performance tests may have been optimized for performance only on MangoBoost products. Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. Results that are based on pre-production systems and components as well as results that have been estimated or simulated using MangoBoost reference platform for informational purposes only. Results may vary based on future changes to any systems, components, specifications, or configurations. Statements in this document that refer to future plans or expectations are forward-looking statements. These statements are based on current expectations and involve many risks and uncertainties that could cause actual results to differ materially from those expressed or implied in such statements. MangoBoost does not guarantee any specific outcome. Nothing contained herein is, or shall be relied upon as, a promise or representation or warranty as to future performance of MangoBoost or any MangoBoost product. The information contained herein shall not be deemed to expand in any way the scope or effect of any representations or warranties contained in the definitive agreement for MangoBoost products.

The information contained herein may not be reproduced in whole or in part without prior written consent of MangoBoost. The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions and typographical errors. The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. MangoBoost assumes no obligation to update or otherwise correct or revise this information and MangoBoost reserves the right to make changes to the content hereof from time to time without any notice. Nothing contained herein is intended by MangoBoost, nor should it be relied upon, as a promise or a representation as to the future.

MANGOBOOST MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION. © 2024 MangoBoost, Inc. All rights reserved.