Simulating AI at Datacenter Scale

With Scala Computing



Executive Summary

This white paper explores how ns-3 simulation on the Scala Compute Platform has helped networking industry leaders innovate at the leading edge of AI/LLM/ML networking technology. By leveraging advanced network simulation, our world class partners and customers have accelerated their development, reduced costs, enhanced quality and robustness, and optimized performance.

They chose to use the Scala Compute Platform for ns-3 because it provides the power, fidelity and flexibility they needed to simulate at data center scale (64k node clusters and beyond). They also needed the industry's first network simulation consortium, cofounded by Scala Computing, because it provides users with official OEM virtual hardware models, defined and validated by industry

leading vendors, as well as the advanced AI/ML workloads provided by well-known data center operators.

Scala will continue to help our customers as they compete to deliver AI, at scale, across cloud, data center, edge, and client with their hardware and software portfolios.

1. Introduction

In the fast-paced world of network hardware manufacturing, staying ahead of the competition requires constant innovation. Network Simulation has emerged as a powerful tool for companies to understand the tradeoffs and achieve their goals, including improved performance and reliability with AI/ML workloads.

For example, customers have used our platform as a virtual data center to explore enhancements to their RDMA protocols. The integrated consortium provides the validated RDMA NIC and switch models, defined by the vendor for use along with the Scala generic models. Customers could assemble their own AI/ML workloads using collectives from Scala, or use the ones provided thru the Consortium by well-known data center operators. Engineers were able to test their RDMA protocols with various configurations, including the latest RoCE protocols. Comparing the results across various configurations showed the performance tradeoffs with drops, incast and drift.

The analytics that comes from simulation has enabled our customers to enhance their critical networking hardware and software products to be lower cost, more robust and more tolerant of network congestion, with fewer drops and minimal switch overloads.

2. Key Benefits of Simulation

2.1 Accelerated Development

Simulation allows hyperscalers and network vendors, including startups, to rapidly prototype workloads, protocols and hardware, significantly reducing their time-to-market. This acceleration in development cycles enables them to respond more

quickly to market demands and technological advancements, empowering them to satisfy the market's exploding demand for AI results.

For example, it is intuitively understood that ideal load balancing would improve flow completion time (FCT). However, knowing the exact settings to achieve a more balanced and efficient route utilization, in a real-world production environment, is much faster and more cost effective with simulation. Another example would be using simulation to understand congestion management protocols. Scala computing is focused on helping you understand your performance and reliability tradeoffs, with the newer AI workloads, at scale.

By using Scala Computing with ns-3 simulation to understand network performance, engineers can assess potential performance tradeoffs associated with load balancing methods and settings for different machine learning workloads running over various network sizes and topologies and with different reorder buffer parameters. For example, customers can use this understanding to accelerate their development of an improved method to handle out-of-order packets and recover from reorder buffer overflows and packet losses.

2.2 Cost Reduction

Simulating product performance has enabled customers to identify where their hardware can be scaled down to save costs. For example, simulation studies have focused on memory requirements for packet reorder buffer sizes or for transmit and receive queues, and the resulting gain in understanding has lead to reduced hardware specifications without sacrificing network performance.

Other customers have studied performance sensitivity to factors such as the network size, mask size and number of connections, to discover how much redundancy was required for robust operation. This is a tradeoff against the cost of additional fabric switches and rack switches within each Pod. In these scenarios, better understanding has led to reduced redundancy and further cost reduction.

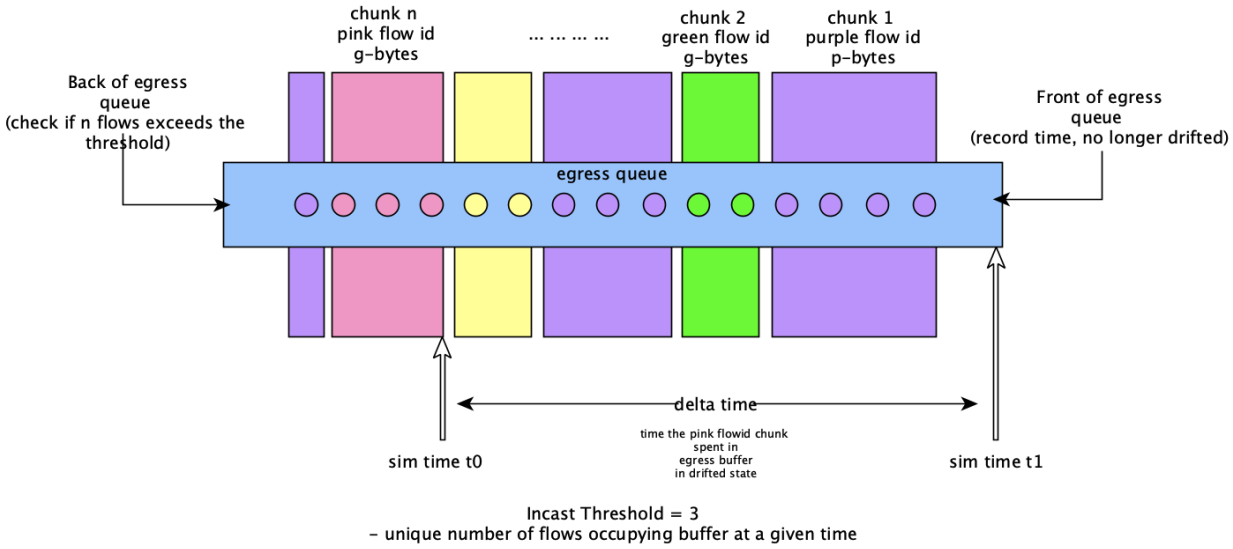
More generally, network simulation helps networking vendors and data center operators resolve potential issues before changes are implemented in real world production environments. This approach substantially reduces the many costs associated with physical prototypes and testing. Testing with simulations results

in faster development iterations that can get a better product to market sooner and with lower production and operational costs.

2.3 Enhanced Quality and Robustness

Simulation enables our customers to perform extensive virtual testing to address potential failure modes early in the development process, leading to more robust and reliable network hardware and protocols with increased performance, especially for AI/ML workloads.

Customers have explored network delays by monitoring the simulated egress queues of the rack switches (RSWs) with our incast detector. These simulations tracked the number of unique flows occupying the queues and the duration they spent within each.



Our clearly charted results provide an understanding of how the delays are caused by incast and drift. Users can study the impact of acknowledgement (ACK) settings on idle time and reduced average bandwidth over the interval. The simulation results can make it possible to fully understand what was causing the queue build up in the network and the resulting Head of Line (HOL) blocking. This issue is particularly problematic in high-bandwidth, low-latency environments like data centers. To mitigate incast, customers can simulate advanced congestion control mechanisms that dynamically reroute traffic and reduce bottlenecks and clearly see the results.

Another sweep study focused on the duration between the receipt of the final data packet and the receipt of the final acknowledgment for that packet. The results track every message and can be shown for any given NIC instance. These waits were studied for a variety of algorithms, such as an all-to-all collective, to understand how the algorithms were reducing the average effective bandwidth of the report interval. Users can study the influence due to the transmit or receive pipelines of the specific NIC as compared to the sensitivity to the drift phenomenon of the specific network switch used. Understanding these performance sensitivities and tradeoffs can lead to innovative solutions to the problem that enhance overall network performance.

Users have investigated some unexplained anomalies that were seen in their production environment in order to understand performance improvements of a variety of NIC-driven load balancing techniques. Our simulated environment provides the ability to run a matrix of tests to simulate with PFCs turned on or off, with these load balancing techniques on or off, with various selective ACK settings, and with the network fully loaded and beyond. Understanding the performance in these corner cases can lead to significantly reduced packet drops.

3. Future Prospects

In partnership with industry leaders, Scala Computing is continuing to enhance their capability and scalability. This includes scale-up/scale-out, Chakra coordinators for LLMs and other AI or ML workloads running on large networks of more than 64k GPU clusters. We are implementing new transport protocols such as the Ultra Ethernet Consortium (UEC), RoCE with enhancements, and have plans to add others, such as Falcon. For global network topologies, we are adding support for superspines so customers can simulate traffic between distant data centers.

We are always focused on continually improving ease of use by further streamlining the topology and workload configuration process and developing automated ranking algorithms for maximum compute resource efficiency and reduced solve time.

We are increasing the value of the the simulations by developing specialized network Drilldown charts to provide more detailed views of the traffic at every

layer, from the data center down to the individual NICs. We are continuously enhancing the performance of our Data Analysis suite and implementing additional features, such as uplink heatmaps for ECMP and improved visualization of buffer stats and port traffic fairness.

4. Conclusion

Simulation technology from Scala Computing provides our industry leading customers and partners with a powerful set of tools to securely leverage the computational power of the cloud and simplify advanced network modeling techniques. Scala Computing offers a cost-effective, efficient, and flexible alternative to traditional physical testing or simulation on internal software/hardware. It not only reduces direct costs related to materials, labor, and equipment but also contributes to faster development cycles, enhanced innovation, and improved overall AI networking performance. By embracing the network simulation capabilities on the Scala Compute Platform, our customers continue to drive innovation, reduce costs, and deliver high-quality products to the market.

