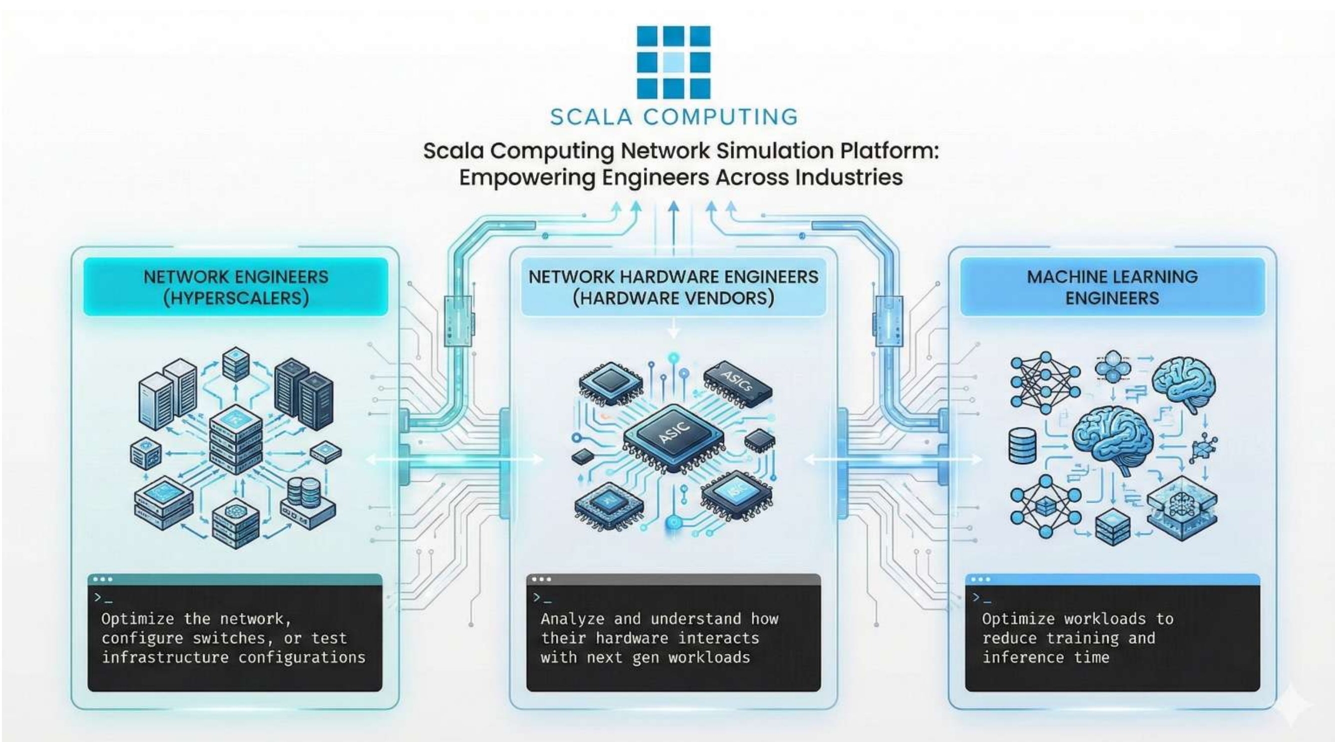
Modern AI data centers operate at a scale and level of complexity that exceeds the limits of traditional testing and validation methods. As networks grow larger, faster, and more tightly coupled to distributed AI workloads, engineering teams face critical questions that cannot be answered through lab experiments, small-scale testbeds, or post-deployment troubleshooting alone. Simulation at scale is no longer optional. It is essential.

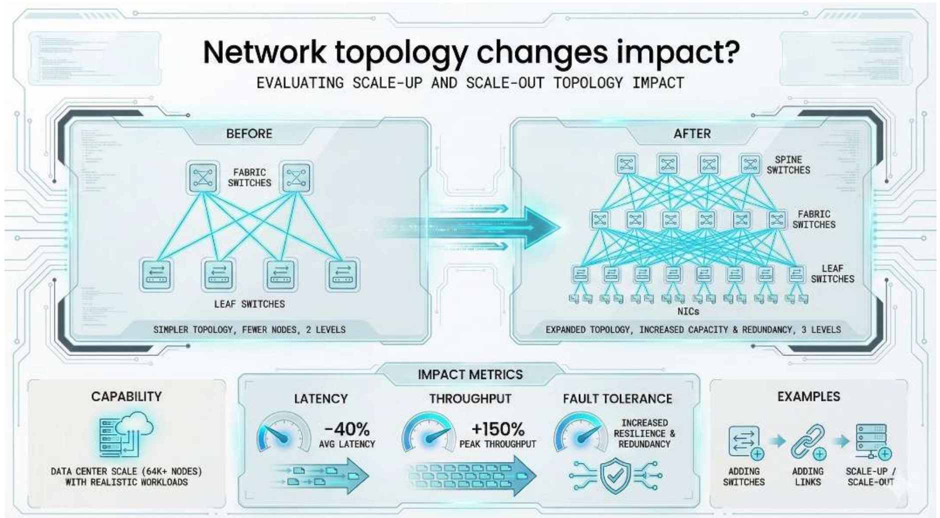
Scala Computing enables next-generation network engineering through high-fidelity simulation at datacenter scale. Our platform allows engineers to model, test, and analyze network behavior under realistic conditions, before hardware is deployed and before workloads reach production. By working closely with leading hardware vendors and data center operators, Scala Computing validates OEM hardware models and real-world workloads, allowing customers to confidently evaluate designs even when physical systems are not yet available.

With these simulations, teams can understand how modern workloads, including large-scale AI training and inference, interact with networking protocols, hardware characteristics, network topologies, and even physical data center layouts. Engineers can explore the impact of topology changes, hardware failures, congestion events, and peak traffic scenarios at full operational scale, without risk to production environments.

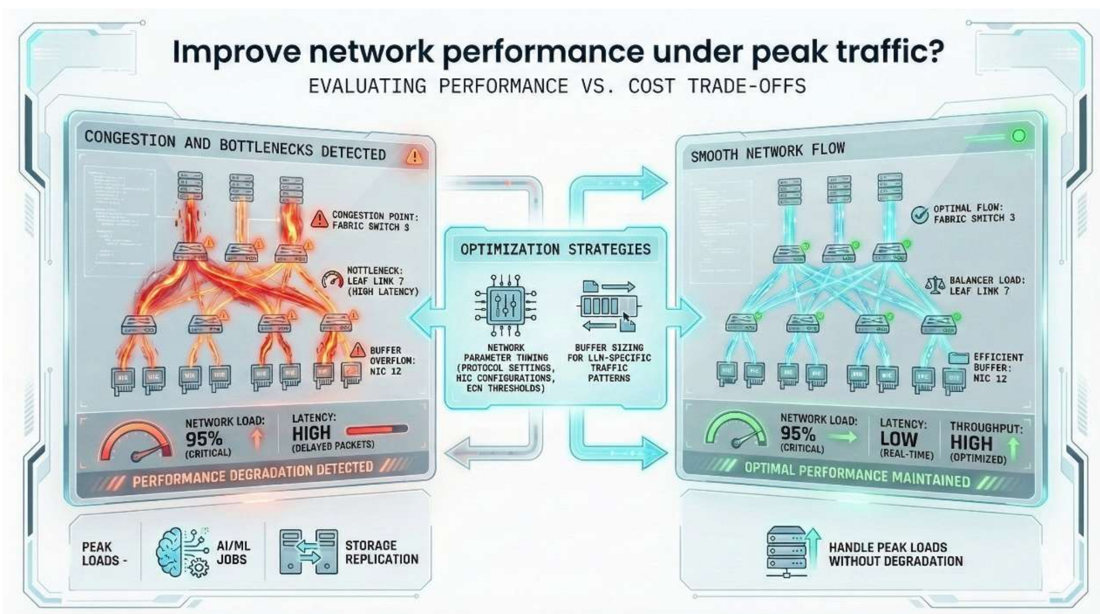


10) What is the impact of network topology changes I am considering?

- Network engineers can use Scala Computing simulation to evaluate how network modifications, such as adding switches, routers, or links, affect latency, throughput, and fault tolerance. **Only Scala Computing can do this at data center scale (64k+ nodes) and with realistic workloads**
- The Scala Computing simulation analytics engine provides insights so that network infrastructure engineers can easily understand how hardware upgrades or expanding the network will impact overall performance, before investing in real-world changes
- Network engineers can evaluate the impact of scale-up and scale-out topologies on their network performance



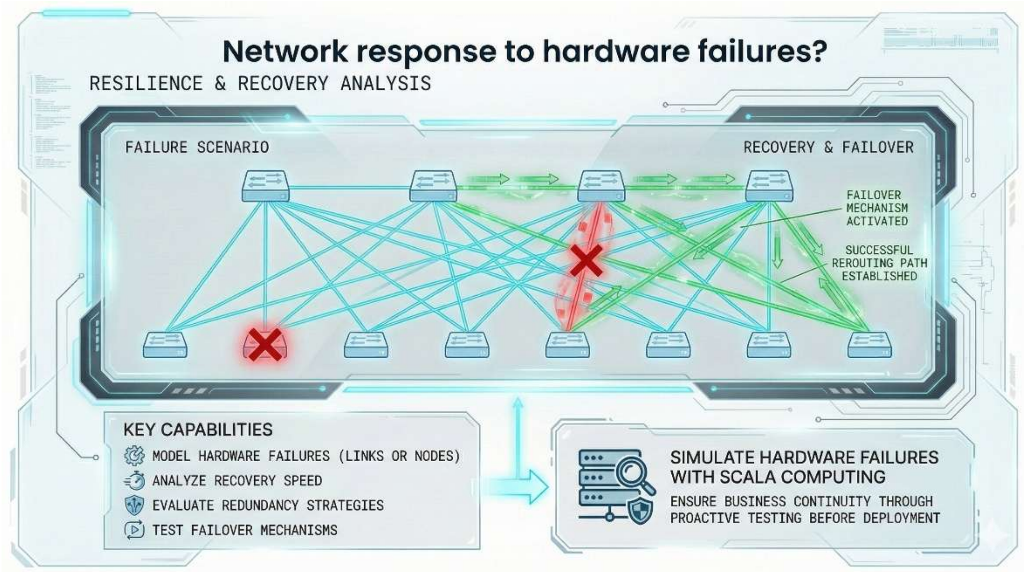
9) How can I improve my network performance under peak traffic conditions?



- Simulations help evaluate trade-offs between performance and cost, guiding decisions on over-provisioning versus under-utilization
- By modeling real-world workloads, such as AI/ML jobs or storage replication, engineers can identify the network’s capacity limits and bottle necks, determine if the network can handle peak expected loads without degradation and plan for scaling
- Network parameter tuning (protocol settings, NIC configurations, ECN thresholds) significantly impacts congestion control effectiveness
- Buffer sizing and management strategies can be adjusted to improve performance of LLM-specific traffic patterns

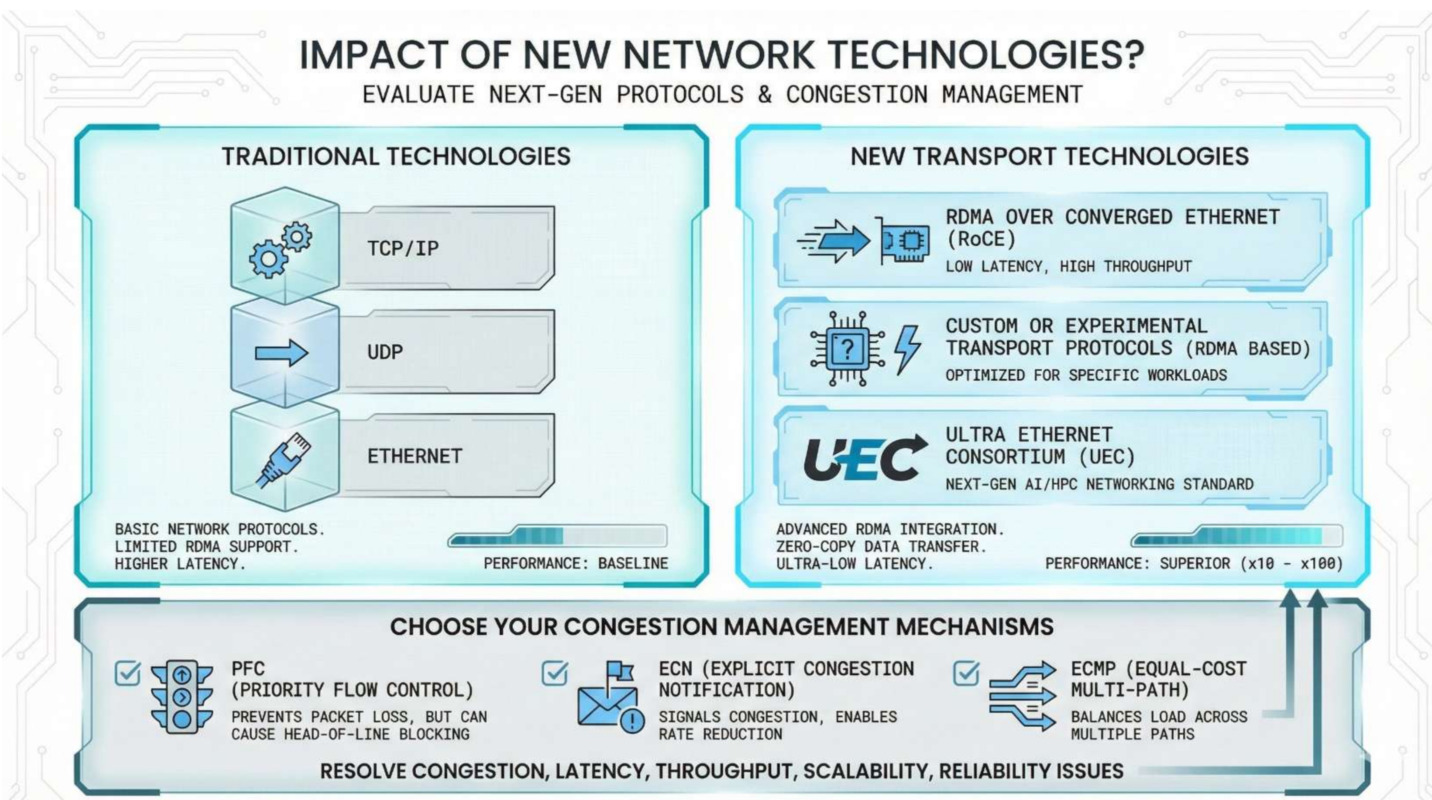
8) How does the network respond to hardware failures?

- Network engineers can Simulate with Scala Computing to model hardware failures (links or nodes) and analyze how quickly the network recovers
- Create a set of simulations to evaluate and compare various redundancy strategies and failover mechanisms with a variety of workloads

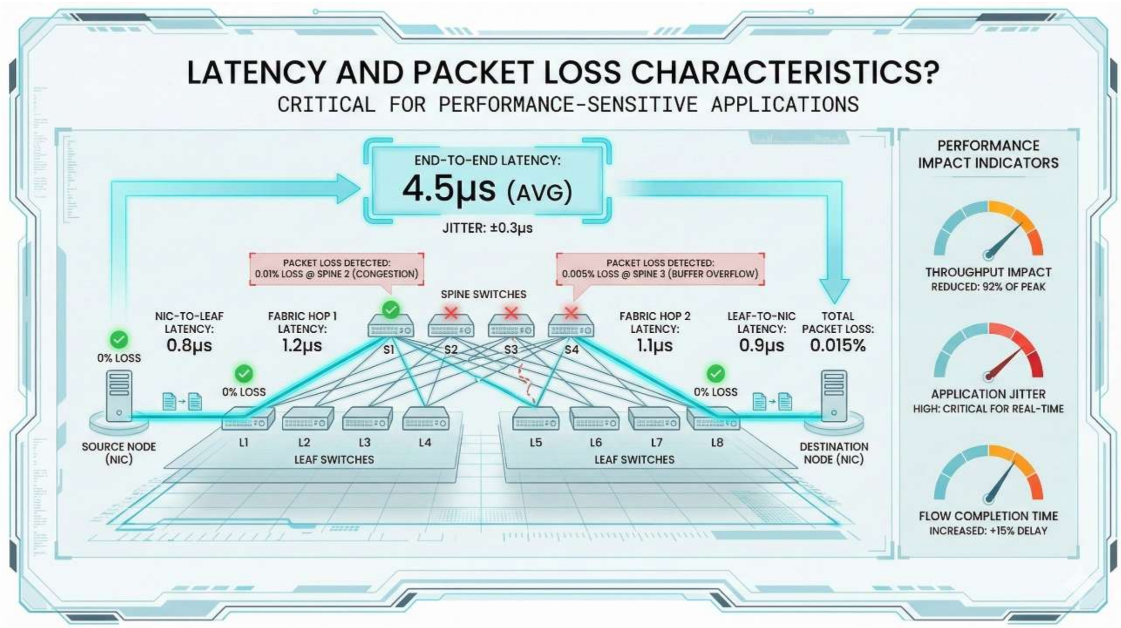


7) What is the impact of new network technologies?

- Network Engineers can simulate RDMA over Converged Ethernet (RoCE) and compare it with the latest cutting-edge transports from the Ultra Ethernet Consortium (UEC) or test the impact of different custom or experimental transport protocols (RDMA based) on performance
- Engineers can use congestion management and notification mechanisms, such as PFC, ECN and/or ECMP to understand and resolve issues related to congestion, latency, throughput, scalability and reliability

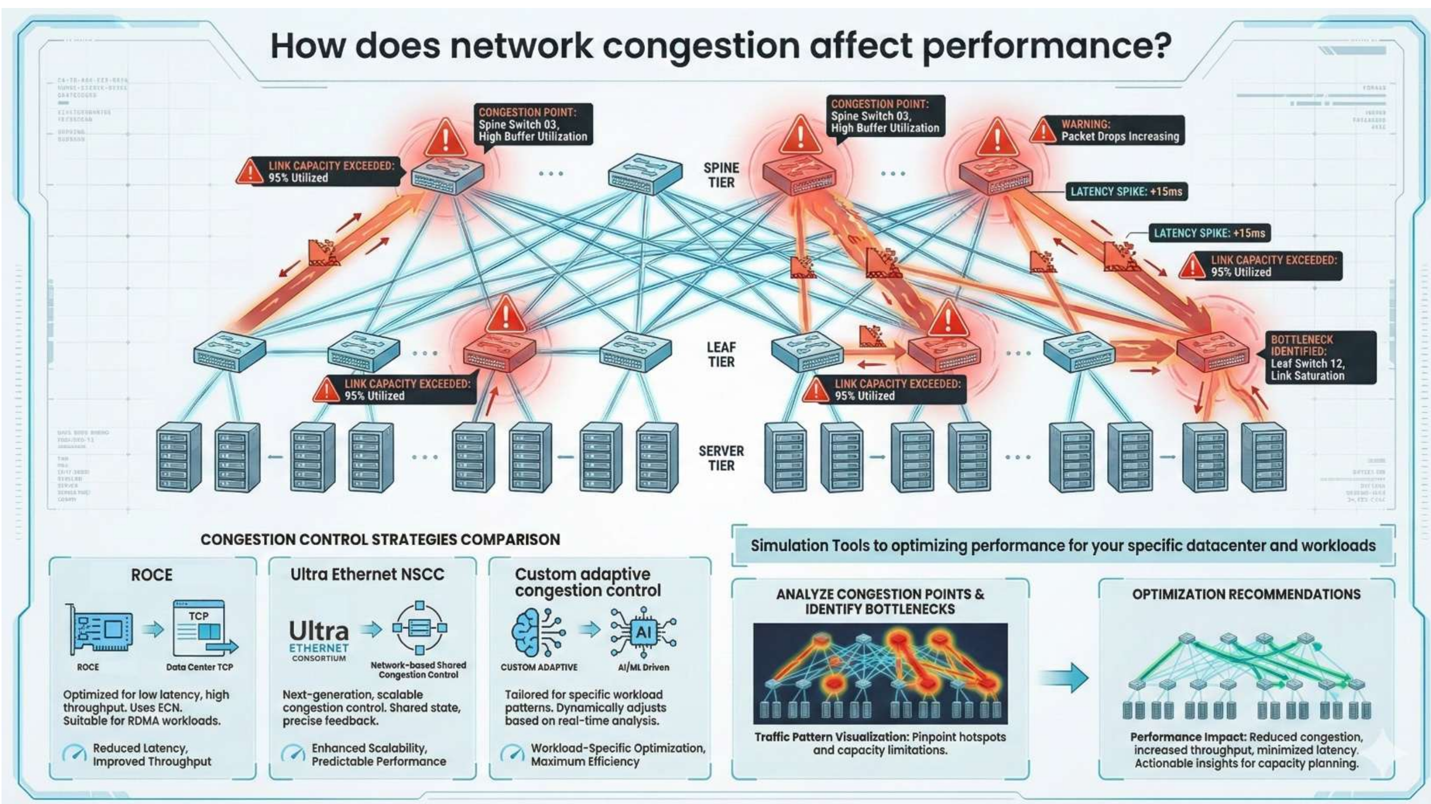


6) What are the latency and packet loss characteristics?



- Scala analytics provides precise measurements of end-to-end latency that can answer questions about the effects of packet loss under various scenarios. This understanding is critical for performance-sensitive applications

5) How does network congestion affect performance?



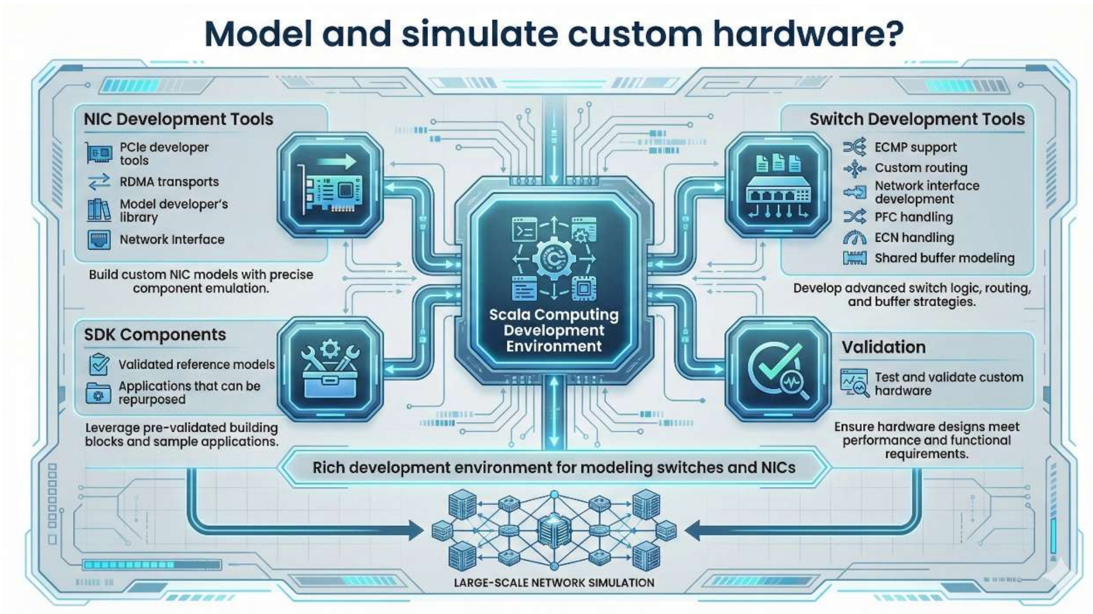
- Simulating traffic patterns helps engineers analyze congestion points and identify bottlenecks
- The Scala Compute platform provides tools to simulate and compare different congestion control strategies (e.g., ROCE DCTCP or Ultra Ethernet NSCC or even custom adaptive congestion control) to identify the best approach for optimizing performance for your specific datacenter and workloads

4) Can I model and simulate my own hardware as part of a larger network?

- Scala Computing provides a rich development environment of modeling tools, and an SDK with validated reference models and applications, so that network hardware vendors can model their switches and NICs, and then test and validate them

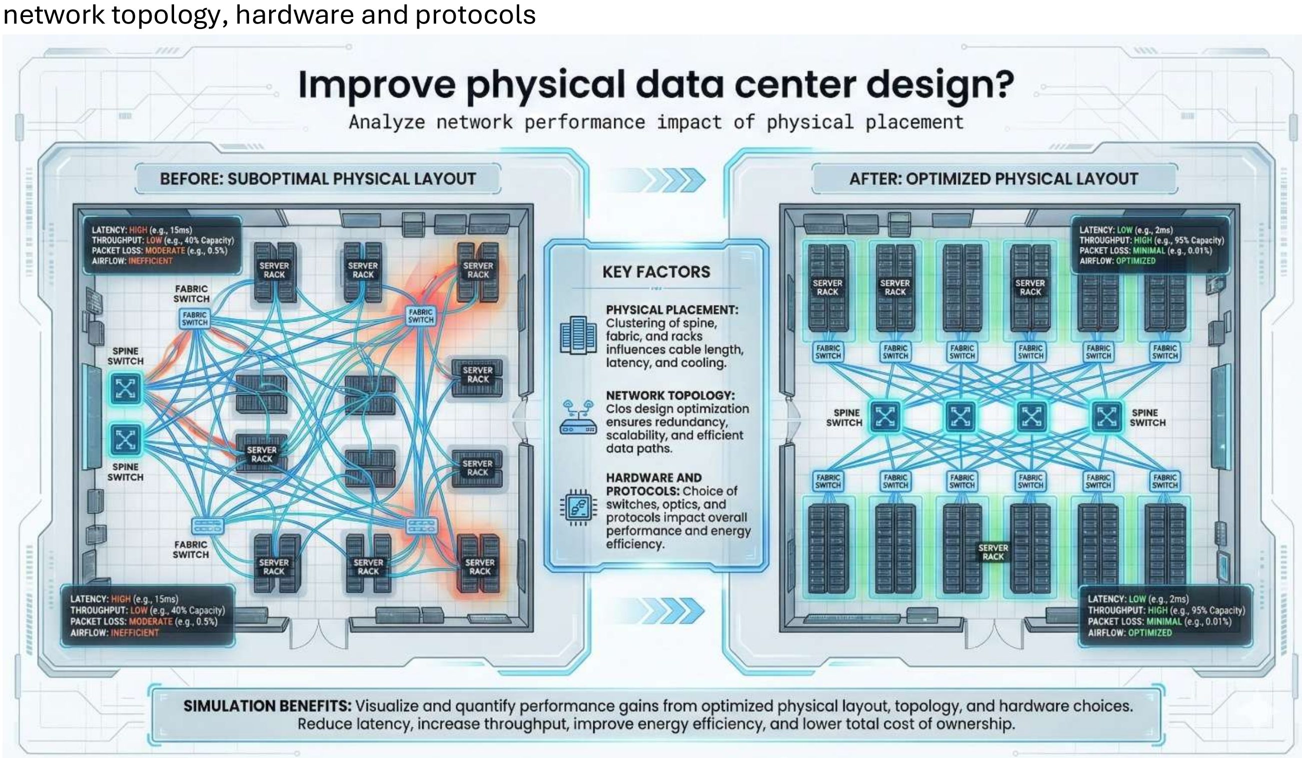
- We provide tools for developing NICs, including PCIe developer tools, RDMA transports, a model developer's library, Network Interface, etc.

- We provide tools for developing Switches, including support for ECMP, custom routing, network interface development, PFC handling, ECN handling, shared buffer modeling, etc.



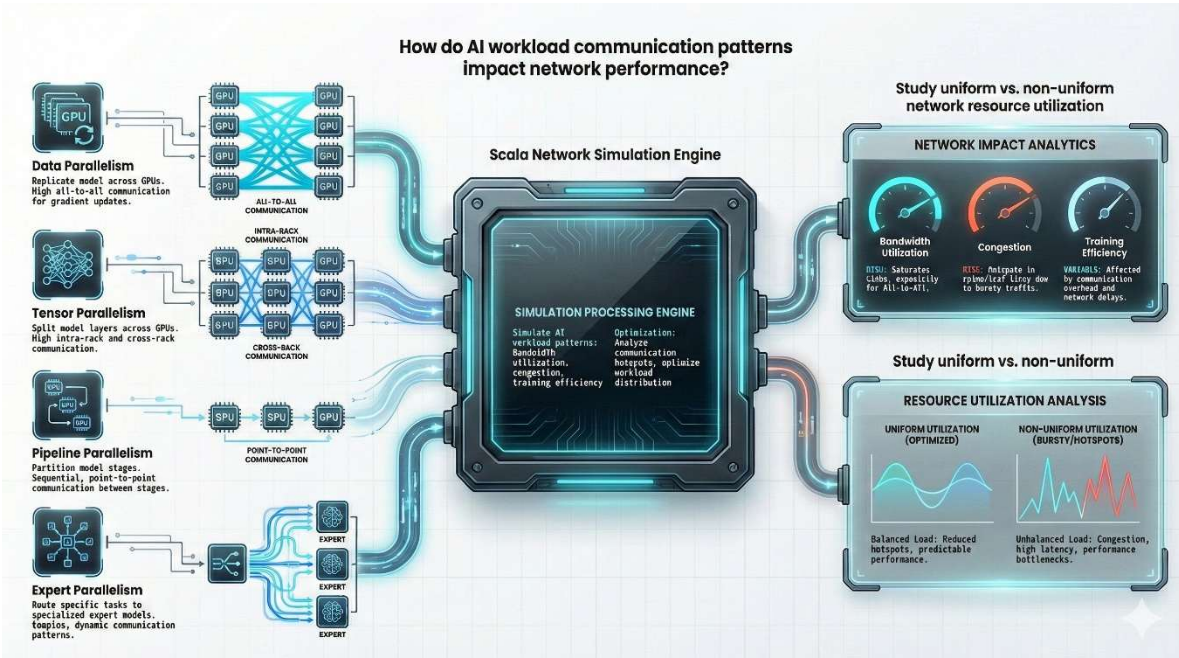
3) Can I improve the physical design of my Data Center?

- Network engineers can use Scala Computing Simulations to analyze how the network performance is affected by the physical placement (clustering) of spine switches, fabric switches, and racks, along with network topology, hardware and protocols



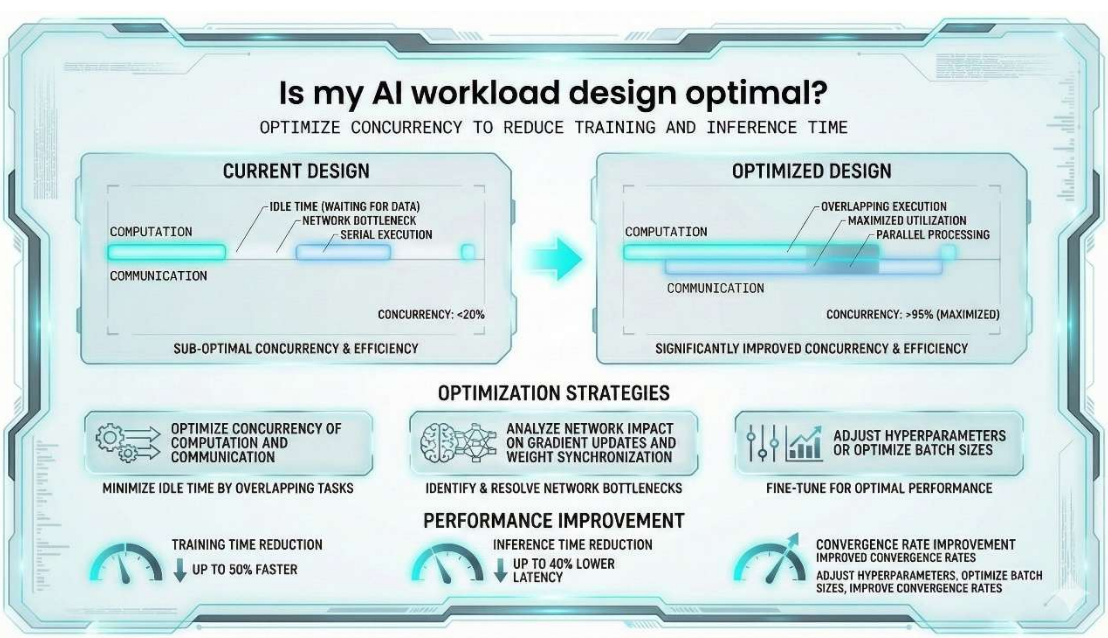
2) How do different AI workload communication patterns impact network performance?

- AI Engineers can simulate various AI workload patterns (e.g., data parallelism, tensor parallelism, pipeline parallelism) to determine their effects on bandwidth utilization, congestion, and overall training efficiency
- By analyzing communication hotspots, engineers can optimize workload distribution to improve training time and resource efficiency
- Machine Learning Engineers can study if their workload utilizes network resources uniformly or results in periods of over-utilization (congestion) and under-utilization



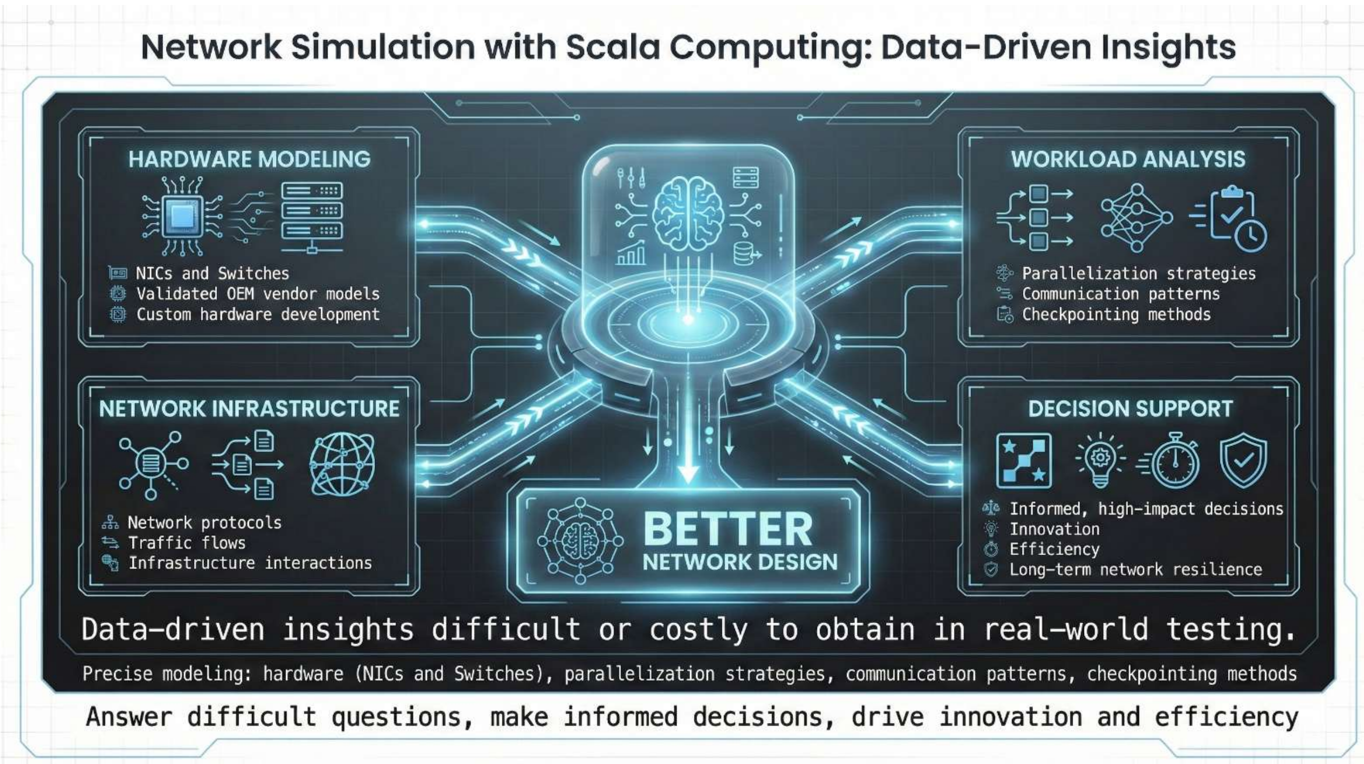
1) Is the design of my AI Workload Optimal?

- Machine Learning Engineers can simulate to understand or redesign their workload to optimize concurrency of computation and communication tasks to reduce training and inference time
- By analyzing network impact on gradient updates and weight synchronization, AI researchers can adjust hyperparameters or optimize batch sizes to improve convergence rates



Conclusion

These first 10 questions are just a start. Network simulation with Scala Computing offers a powerful and flexible platform for answering many other highly specific and detailed questions about your **network design, AI workloads, and hardware performance.** Scala can provide **data-driven insights** that are often difficult or costly to obtain in real-world testing. By enabling precise modeling of **hardware** (NICs and Switches), **parallelization strategies, communication patterns, and checkpointing methods**, engineers can **enhance efficiency and scalability** without physical deployments.



By offering an environment where **network protocols, traffic flows, and infrastructure interactions** can be rigorously tested, Scala Computing enables organizations to answer difficult questions and make **informed, high-impact decisions** that drive innovation, efficiency, and long-term network resilience.