



Chakra Workloads: Bring AI Network Simulation to Life



A New Era in AI Network Simulation

Scala Computing invites you to explore the next generation of **AI network simulation** using **Chakra workloads** on our advanced simulation platform.

The Chakra format was developed by MLCommons to solve a critical gap between **AI workload behavior** and **infrastructure evaluation**. It is a standardized sharable representation of how large-scale AI workloads (like LLM training or distributed deep learning) behave on the network.

With this new release, Scala's high-performance simulation environment now supports both captured Chakra traces and synthetic Chakra workloads, offering an unprecedented ability to study how modern AI training and inference jobs interact with your network. This marks a major step forward in accurately modeling realistic distributed training and communication patterns at massive scale, up to and beyond 64,000 nodes, empowering engineers to analyze, predict, and optimize AI workload performance at *true data center scale*.

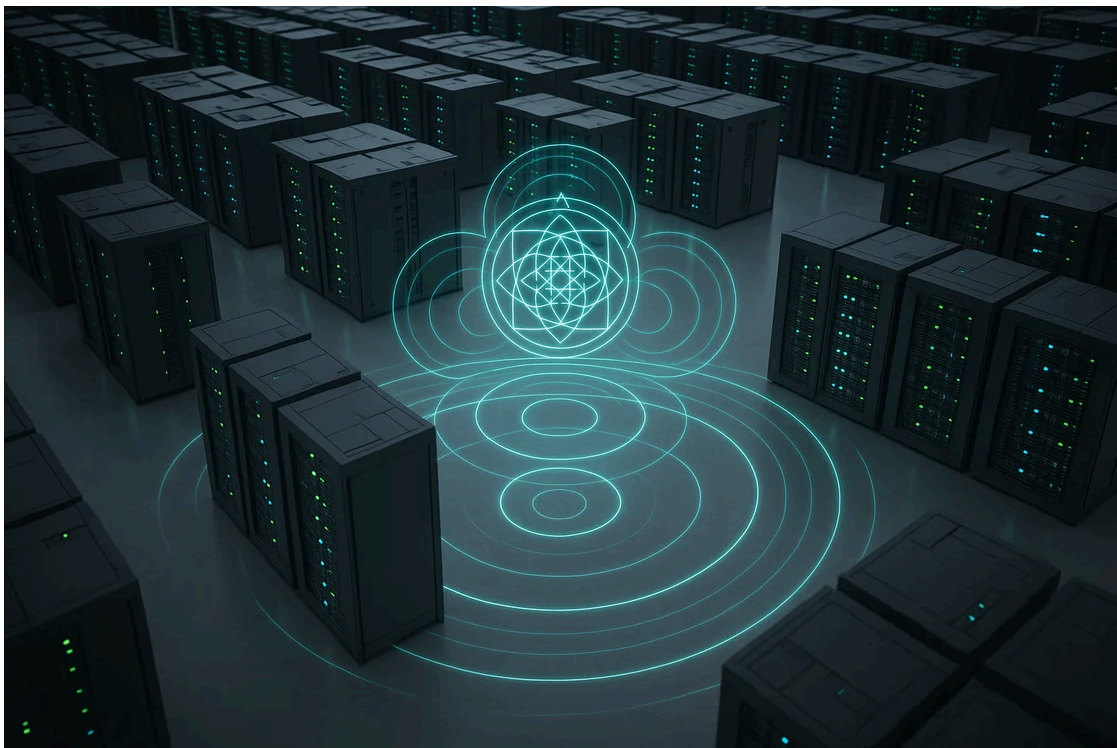
Why This Matters

As AI training continues to dominate data center traffic, understanding how these workloads behave under realistic network conditions has become mission-critical. Physical benchmarking at scale is often impractical, requiring thousands of GPUs, large-scale coordination, and enormous energy consumption.

Scala's Chakra-enabled simulation platform eliminates that barrier, allowing teams to explore real-world performance scenarios with precision and at a fraction of the cost.

With Chakra workloads, your team can now:

- **Simulate LLM and AI training workloads** with realistic communication and synchronization patterns.
- **Replicate real-world GPU behavior** captured directly from production-scale training environments.
- **Scale up to data center topologies** with vendor-validated switch, NIC, and NPU models.
- **Compare network configurations, routing algorithms, and congestion controls** under authentic AI traffic.
- **Optimize workload-to-network coupling** to improve end-to-end training performance and efficiency.



Captured vs. Synthetic Chakra? Both Have Strategic Roles

Scala's implementation supports both **captured** and **synthetic** Chakra workloads, each providing unique advantages for research and optimization:

- **Captured Chakra traces** are derived directly from real-world GPU clusters during active AI training. They deliver unmatched realism, enabling engineers to study genuine workload behaviors, traffic patterns, and congestion dynamics.
- **Synthetic Chakra workloads** are *parameterized, configurable traces* that mirror real-world communication behavior while giving researchers flexibility to explore *what-if scenarios*, such as scaling to larger clusters, testing new AI models, or evaluating future architectures.

- Synthetic traces can be tuned to represent a wide range of workloads (GPT, LLaMA, Mistral, etc.) or GPUs (A100, H100, etc.).
- They are ideal for collaborative studies, enabling users to share results *without exposing proprietary model details*.
- Synthetic Chakras can also be validated against captured traces, ensuring correlation with real-world performance.

By combining both types of Chakra workloads, Scala customers gain the best of both worlds, the realism of actual captured AI communication and the flexibility and scalability of synthetic exploration.

Powering Network Research and AI Infrastructure Design

Network engineers, architects, and researchers can use Chakra workloads to:

- **Validate next-generation network architectures** for AI and HPC workloads.
- **Optimize collective communication patterns** (All-Reduce, All-to-All, Broadcast, etc.).
- **Study bottlenecks** in flow control, ECN thresholds, or buffer management at scale.
- **Evaluate and tune protocols** for networks at data center scale.
- **Predict performance and capacity requirements** for future large-scale deployments.

By accurately simulating workload and network interaction, teams can minimize idle GPU time, reduce training completion time, and uncover

performance inefficiencies long before deploying physical infrastructure.



Get Started

Ready to see it in action?

Scala Computing is now offering demos and evaluation platforms for customers interested in exploring the benefits of **Chakra workload simulation**.

Run simulations with real or synthetic Chakras, analyze end-to-end communication efficiency, and see how your network performs under genuine AI workloads, *without the hardware costs*.

[Contact Us to request a Demo or Evaluation Platform](#)



Scala Computing's Chakra-enabled simulation platform represents a major leap forward for AI networking, combining accuracy, scalability, and flexibility to help engineers design the next generation of data center and AI infrastructure.

[For more on the value of Chakra for Network Simulations](#)

[A case study on simulating distributed LLM training](#)

Scala Computing Inc, 369 Lexington Ave, Suite 206, New York, NY 10017, USA

[Unsubscribe](#) [Manage preferences](#)