Research Statement

Sushovan Das

susdas@ethz.ch
sushovan-das.net

Modern packet-switched datacenter networks (DCNs) face a growing mismatch with emerging workloads: distributed AI/ML training, HPC, and large-scale graph analytics generate long-lived, structured collective communication, yet DCNs still rely on a generic "packet" abstraction designed for short, bursty, uncorrelated flows. At the same time, CMOS-based packet-switch ASICs are reaching power and cost limits, making energy the dominant bottleneck for next-generation clusters.

My research vision addresses this dual challenge by operating at the intersection of device-level capabilities, network fabric design, and application-level communication structure. I pursue three complementary thrusts: (i) rethinking next-generation switch hardware through a seamless fusion of low-energy optics with electronic packet pipelines, (ii) designing high-performance and energy-efficient DCN architectures that leverage optical circuit-switching for predictable and scalable communication, and (iii) extending application-aware circuit–packet co-design principles for cross domains e.g., LEO satellite networks and quantum clouds.

Equipped with a cross-disciplinary ECE–CS background, I position myself as a researcher who bridges device-level intuition with systems and architecture design where it can have real impact. My work has appeared in major networking and systems venues, including *IEEE INFOCOM*, *IEEE/ACM Transactions on Networking*, *USENIX NSDI*, *ACM HotNets*, *IEEE/Optica JLT*, and *IEEE ICC*.

Research Overview

The fundamental abstraction of today's networks—*that everything is a packet*—has served computing well for decades. However, the workloads driving modern packet-switched DCNs are rapidly evolving. Distributed AI/ML training, Mixture-of-Experts models, large-scale graph processing, and scientific computing all exhibit structured, high-volume communication patterns fundamentally different from the short, bursty flows packet switching was designed for. These workloads are inherently *circuit-friendly*: long-lived, predictable, and dominated by collective motifs such as all-reduce, all-gather, MoE token dispatch, and pipeline boundaries.

Moreover, the post–*Moore's law of networking* era is exposing energy and scalability limits of CMOS packet-switch ASICs. Networking hardware accounts for a growing share of datacenter power, and next-generation AI clusters threaten to push energy budgets into unsustainable territory. Optical circuit switching (OCS) offers a compelling alternative—data-rate agnostic, low-power, and no buffer—but naïve deployments lack multicast support, struggle with traffic skewness, and incur severe tail latency during circuit reconfiguration.

My research addresses these challenges through a cross-layer methodology connecting optical device capabilities with network fabric design and application-level communication structure. My hybrid ECE (undergraduate and master's) and CS (PhD and postdoc) background uniquely positions me to unify the strengths of both communities. During my PhD, I redesigned optical DCN architectures across three fronts: *Shufflecast* added native physical-layer multicast, *OSSV* introduced optical-edge techniques to mitigate skewness, and *Phoenix* applied space–time opportunistic traffic correction to overcome OCS downtime. Together, these works showed how integrating optical primitives with workload structure can close long-standing gaps in all-optical DCNs.

Building on these insights, my postdoctoral research expands the same principles beyond optical cores. I am developing architectures for multi-ASIC packet switches that use circuit-switched indirection to overcome bandwidth bottlenecks, and exploring how structured ML collectives can inform device–fabric co-design more broadly. This progression—from optical DCN architectures (PhD) to hybrid optical–electrical switch substrates (postdoc)—naturally motivates the broader research vision outlined below.

My research program is organized around three interconnected thrusts: (i) a *hardware vision* for rethinking switch substrates through optical–electronic fusion and lane-level elasticity; (ii) an *architecture vision* that leverages optical circuit-switching principles to design predictable and energy-efficient DCN architectures for AI/HPC workloads; and (iii) a *cross-domain vision* that extends circuit–packet co-design to emerging platforms such as LEO satellite networks and hybrid quantum clouds. Together, these thrusts outline a unified agenda centered on structured communication, constrained reconfiguration, and cross-layer design.

Thrust 1: Hardware Vision

Future AI/HPC clusters demand switching substrates that are neither purely packet-switched nor purely optical. Thrust 1 focuses on the device and hardware-level foundations needed to build multi-modal switch fabrics that exploit optical predictability and electronic flexibility. The goal is to establish optical–electronic co-design as a first-class principle for next-generation cluster switches.

Multi-ASIC packet switches with optical indirection. As ASICs hit limits in I/O density, power, and thermal budgets, vendors increasingly assemble high-radix switches from multiple chiplets. The resulting inter-ASIC bandwidth bottleneck creates persistent hotspots. As an ongoing work, I design an architecture that inserts a lightweight circuit-switched indirection layer before ASIC ingress. By dynamically remapping ports based on observed traffic structure, this layer localizes heavy flows and dramatically reduces inter-ASIC load. Early prototypes show monolithic-switch performance while cutting interconnect demand by more than half.

Switch 2.0: a unified hardware substrate. Beyond multi-ASIC designs, I envision Switch 2.0—a unified switching substrate that tightly integrates packet-switch ASICs, optical circuit fabrics, passive optical primitives (e.g., splitters, WDM modules), and emerging elastic multi-lane transceivers. Modern transceivers expose tens of independently controllable optical lanes, yet today's packet switches treat these lanes as a single indivisible channel. Switch 2.0 takes the opposite view: lane-level elasticity is a fundamental resource. By activating packet forwarding, circuit paths, optical multicast, or lane-parallel striping on demand, the fabric can match its mode and granularity to the algebraic structure of ML collectives (rings, recursive doubling, MoE bipartite exchanges). Co-packaged optics further strengthens this direction by lowering electrical loss and enabling denser, energy-efficient lane-level optical I/O, making per-lane agility practical at scale. This creates a new class of hardware–software interfaces where the network adapts its bandwidth, power state, optical mode, and forwarding semantics at iteration boundaries. Realizing Switch 2.0 requires (i) lane-level control policies, (ii) hybrid buffering strategies for mixed packet–circuit paths, (iii) stable power-state transitions under tightly synchronized collectives, and (iv) new compiler/runtime APIs that expose collective structure to the fabric. My long-term goal is to build the first switch architecture that is at once optical, electronic, elastic, and collective-aware.

Cross-layer hardware–software co-design. A central direction is mapping collective structures—rings, recursive doubling, MoE bipartite exchange, and pipeline DAGs—onto packet routes, circuit schedules, and per-lane activation patterns. I will explore combinatorial optimization for joint packet–circuit scheduling, queueing models for mixed bufferless paths, and control-theoretic analyses of safe reconfiguration. These efforts will be supported by a multi-layer simulator coupling packet dynamics, OCS reconfiguration, transceiver elasticity, and energy models.

Thrust 2: Architecture Vision

While Thrust 1 establishes the switching substrate, Thrust 2 focuses on the architectural principles that make optical circuit switching practical, predictable, and high-performance at datacenter scale. My work shows that combining predictable OCS cores with flexible optical edges provides a powerful foundation for structured, high-performance communication. Looking ahead, I aim to extend these principles towards elastic, collective-aware optical architectures that adapt resources to application demand.

Optical multicast at the core. Modern workloads rely heavily on multicast-like patterns: all-reduce/broadcast in ML training, replication in storage, and tree-structured propagation in HPC. Yet OCS fabrics only support point-to-point circuits. During my PhD, I designed Shufflecast, that introduces an optical multicast substrate built using low-cost passive splitters arranged in a ToR-level topology. By designing splitter connectivity and precomputing static multicast routes, Shufflecast enables simultaneous line-rate multicasts without dynamic tree construction or packet-core involvement. Analysis and testbed evaluation demonstrate substantial energy and cost savings over IP multicast while preserving throughput and reliability.

Optical edge reconfiguration. OCS cores face two bottlenecks: traffic skewness and reconfiguration-induced downtime. During my PhD, I designed OSSV that introduces spatial flexibility by dynamically reshaping rack membership via small OCS devices at the edge, aligning workloads with the core's periodic schedule. My ongoing work Phoenix generalizes this to space–time coordination, using opportunistic traffic correction to align reconfigurations with low-impact phases of collective progress. Both systems embody a unified principle: the optical edge can recover flexibility that the OCS core inherently lacks, and space–time manipulation is a powerful lever for predictable optical DCNs.

Elastic routing and collective-aware optical architectures. Emerging multi-lane optical transceivers enable *elastic optical networking* within datacenters. Each lane is independently controllable, allowing circuits to be allocated bandwidth proportional to flow demand: some flows may require only a few lanes, while others may need an entire bundle. This creates opportunities for elastic circuit routing, improved circuit multiplexing beyond the time-division schemes of systems like Sirius or Opera, and physical isolation for different traffic classes (e.g., collectives vs. RPCs). My goal is to design algorithms, control policies, and topologies that integrate this elasticity with packet switching, enabling DCNs that adapt their optical resources to application semantics while retaining strong energy proportionality.

Thrust 3: Cross-Domain Vision

The principles underlying my optical DCN work—structured communication, constrained reconfiguration, and hybrid circuit–packet control—arise naturally in several high-impact networking domains. Thrust 3 extends these ideas where physical constraints vary, yet the underlying challenges mirror optical DCNs.

LEO satellite networks. LEO constellations form time-varying networks with free-space optical ISLs that appear and disappear predictably with orbital motion. ISLs behave like slowly reconfiguring circuits under strict pointing, power, and thermal limits, making scheduling and routing fundamentally different from terrestrial systems. I plan to develop application-aware LEO–terrestrial routing that forecasts connectivity windows and opportunistically offloads latency-sensitive traffic during ultra-low-latency opportunities. Beyond routing, I will design time-varying circuit schedules that jointly optimize orbital motion, handovers, and load distribution while avoiding excessive reconfigures. I will also explore lightweight in-orbit contention management and multi-layer LEO–ground co-design, where satellites handle rapid local adaptation while ground stations provide global coordination.

Hybrid classical–quantum clouds. Quantum networks rely on entanglement distribution and teleportation, orchestrated by classical control-plane traffic. Here too, the underlying primitives resemble optical circuits: links are fragile, coherence-limited, and require carefully timed scheduling. Building on insights from optical DCNs, I will design quantum-aware routing that aligns forwarding decisions with decoherence windows, as well as congestion-control mechanisms that adapt to stochastic loss and variable fidelity. Beyond routing, I aim to explore entanglement scheduling and purification strategies that jointly optimize throughput, reliability, and resource usage, and to investigate how classical networks can act as intelligent orchestration layers stabilizing distributed quantum workloads. These directions collectively aim to build a practical, scalable foundation for hybrid classical–quantum cloud platforms.

Selected Publications

1. **Sushovan Das**, *"Towards All-Optical Circuit-Switched Datacenter Network Architectures With Low Energy and High Performance"*, *PhD Thesis*, Rice University, 2024.
2. **Sushovan Das**, Arlei Silva, T. S. Eugene Ng, "Rearchitecting Datacenter Networks: A New Paradigm with Optical Core and Optical Edge", *IEEE INFOCOM*, 2024.
3. Shalini Choudhury, **Sushovan Das**, Sanjoy Paul, Prasanthi Maddala, Ivan Seskar, Dipankar Raychaudhuri, "MEC-Intelligent Agent Support for Low-Latency Data Plane in Private NextG Core", *IEEE ICC*, 2024.
4. **Sushovan Das**, Debasish Datta, "Revisiting Packet-switched WDM Rings for Metro Networks: A Comprehensive Cross-Layer Assessment", *IEEE/Optica Journal of Lightwave Technology*, 2023.
5. **Sushovan Das**, Arlei Silva, T. S. Eugene Ng, "Near Non-blocking Performance with All-optical Circuit-switched Core", *ACM SIGCOMM Poster Program*, 2023 (Third Place).
6. **Sushovan Das**, Afsaneh Rahbar, Xinyu Crystal Wu, Zhuang Wang, Weitao Wang, Ang Chen, T. S. Eugene Ng, "Shufflecast: An Optical, Data-Rate Agnostic and Low-Power Multicast Architecture for Next-Generation Compute Clusters", *IEEE/ACM Transactions on Networking*, 2022.
7. Weitao Wang, Dingming Wu, **Sushovan Das**, Afsaneh Rahbar, Ang Chen, T. S. Eugene Ng, "RDC: Energy-Efficient Data Center Network Congestion Relief with Topological Reconfigurability at the Edge", *USENIX NSDI*, 2022.
8. **Sushovan Das**, Weitao Wang, T. S. Eugene Ng, "Towards All-optical Circuit-switched Datacenter Network Cores: The Case for Mitigating Traffic Skewness at the Edge", *ACM SIGCOMM Workshop on Optical Systems (OptSys)*, 2021.
9. Weitao Wang, **Sushovan Das**, T. S. Eugene Ng, "Abstractions for Reconfigurable Hybrid Network Update and A Consistent Update Approach", *ACM SIGCOMM Workshop on Optical Systems (OptSys)*, 2021.
10. Weitao Wang, **Sushovan Das**, Xinyu Crystal Wu, Zhuang Wang, Ang Chen, T. S. Eugene Ng, "MXDAG: A Hybrid Abstraction for Emerging Applications", *ACM HotNets*, 2021.