

# Multi-Objective Optimization of Latency, Energy Consumption, and Reliability in Optical Switching Networks

Sanaya Doctor

Department of Computer Science, George Mason University, Fairfax, VA, USA.  
contactsanaya@gmu.edu

Raimo Kangas

Department of Computer Science, Colorado State University, Fort Collins, CO, USA.  
raimokangas76@colostate.edu

## Abstract

Optical switching networks have become a foundational infrastructure for data centers, high-performance computing, telecommunications, and increasingly for distributed artificial intelligence systems. Their design must satisfy competing objectives: minimizing latency, reducing energy consumption, and preserving reliability under dynamic and uncertain traffic conditions. This paper examines multi-objective optimization of latency, energy consumption, and reliability in optical switching networks from a system-level perspective. It argues that these objectives are not independent technical variables but are structurally coupled through switching granularity, control-plane design, hardware redundancy, fiber topology, cooling requirements, and operational governance. The analysis compares circuit switching, wavelength switching, burst switching, packet switching, and hybrid electrical-optical architectures, showing how each paradigm distributes latency and energy burdens differently across the network stack. It then reviews multi-objective optimization approaches, including Pareto-based evolutionary methods, simulation-driven design, and software-defined orchestration, while emphasizing that optimization outcomes are shaped by institutional priorities and deployment constraints. The paper further discusses sustainability, robustness, fairness, and policy implications, noting that energy-efficient designs can either improve or worsen reliability depending on maintenance regimes, traffic fairness, and failure recovery strategies. Case illustrations from intra-data center interconnect, cloud infrastructure, and resilient backbone networks clarify the trade-offs. The conclusion proposes that future optical switching systems should be evaluated as socio-technical infrastructures rather than purely technical artifacts, with optimization frameworks that incorporate lifecycle energy, service equity, failure transparency, and adaptive governance.

## Keywords

optical switching networks; multi-objective optimization; latency; energy consumption; reliability; sustainability; resilience; software-defined optical networks.

## 1. Introduction

Optical switching networks have moved from a specialized transport technology to a central element of contemporary digital infrastructure. The rapid growth of cloud computing, machine learning workloads, real-time analytics, and distributed storage has increased demand for interconnect fabrics that can transfer enormous volumes of data with predictable

delay and high availability [1]. Electrical packet switching remains flexible and mature, but it faces fundamental limits in bandwidth density, energy per bit, and wiring complexity as rack-scale and cluster-scale systems expand [2]. Optical switching offers high bandwidth, low crosstalk, and potentially lower transmission loss, yet its benefits depend on architectural choices that determine how connections are established, how contention is resolved, and how failures are isolated [3].

The design problem is inherently multi-objective. Latency, energy consumption, and reliability interact in ways that resist simple trade-off curves. A network that reduces latency by using aggressive optical bypass may consume additional energy in transceivers, amplifiers, and control-plane signaling. A network that lowers energy by consolidating traffic onto fewer high-capacity links may become more vulnerable to correlated failures. A network that increases reliability through redundant paths and protection switching may consume more hardware, fiber, and operational attention. These interactions mean that optimization cannot be reduced to minimizing one metric while treating the others as constraints. Instead, design must identify acceptable regions within a multidimensional performance space and then align them with service-level expectations, regulatory obligations, and sustainability commitments [4], [5].

This paper addresses that problem by examining multi-objective optimization of latency, energy consumption, and reliability in optical switching networks. The discussion is system-level rather than device-level. It focuses on how switching paradigms, control architectures, deployment models, and governance arrangements shape the feasible trade-off space. It also considers how optimization methods from engineering and computer science can be adapted to socio-technical realities, including maintenance practices, energy markets, and fairness concerns. The central claim is that optical switching design should be understood as a problem of infrastructure governance as much as a problem of hardware and algorithms.

## **2. Architectural Foundations and Switching Paradigms**

Optical switching networks can be organized according to the granularity at which switching occurs. Circuit switching establishes end-to-end lightpaths and reserves capacity for the duration of a connection. It offers predictable latency and low per-packet processing overhead, but it can waste bandwidth when traffic is bursty. Wavelength switching operates at the level of wavelength channels, allowing multiple connections to share fiber while retaining optical transparency. Burst switching aggregates data into larger units and can improve utilization for bursty traffic, though it introduces contention and buffering challenges. Packet switching provides fine-grained multiplexing and statistical sharing, but it requires optical buffering, header processing, and synchronization that remain difficult to implement efficiently in the all-optical domain. Hybrid electrical-optical architectures combine electronic packet switching for control and small flows with optical circuit or wavelength switching for large elephant flows [3], [6].

Each paradigm distributes latency and energy differently. Circuit and wavelength switching reduce electronic processing and queuing delays, but they require signaling and configuration time that can dominate for short transfers. Packet switching reduces configuration latency but increases energy use in electronic switching, buffering, and header processing. Hybrid architectures attempt to capture the benefits of both by routing delay-sensitive or small flows through electronics and high-volume flows through optics. The result is a segmented system in which latency and energy outcomes depend on traffic classification, flow-size distribution, and the responsiveness of the control plane [7]. Reliability also varies. A failed lightpath in a

circuit-switched network affects a bounded set of connections, while a failed control-plane component in a hybrid network may disrupt many flows simultaneously.

The control plane is therefore as important as the data plane. Software-defined optical networking separates control logic from switching hardware, enabling centralized or hierarchical orchestration of lightpaths, wavelengths, and protection policies [12]. This separation can improve energy efficiency by allowing the network to reconfigure resources according to demand, but it also creates new dependencies. If the controller becomes unavailable or compromised, the network may lose the ability to adapt, even if the optical hardware remains operational. Studies of OpenFlow-based and GMPLS-controlled optical networks show that control-plane unification can simplify multi-layer operations, but it also raises questions about failure domains, vendor interoperability, and operational skill requirements [13], [14]. Architectural choice is thus not merely a matter of performance; it determines who can govern the network, how quickly failures can be diagnosed, and how easily sustainability goals can be enforced.

### **3. The Coupled Nature of Latency, Energy, and Reliability**

Latency in optical switching networks includes propagation delay, transmission delay, switching delay, queuing delay, and control-plane delay. Propagation delay is fixed by distance and fiber properties, though hollow core fiber and advanced materials may reduce it [16]. Transmission delay depends on link capacity and serialization. Switching delay depends on the switching technology, reconfiguration time, and contention resolution. Queuing delay depends on traffic load and buffering strategy. Control-plane delay depends on signaling, topology discovery, and policy enforcement. Reducing one component can increase another. For example, optical bypass reduces queuing and electronic processing, but if reconfiguration is slow, short flows may experience greater delay than they would in an electronic packet switch [1], [7].

Energy consumption is similarly distributed. Optical transmission can be more energy-efficient than electrical transmission over long distances, but transceivers, amplifiers, cooling systems, and control processors consume significant power. Data center interconnects face additional constraints because energy used for cooling and power conversion compounds the energy used by switching and transmission [4], [5]. Aggressive optimization of energy per bit may lead to higher aggregate energy if it requires frequent reconfiguration, over-provisioning, or complex control signaling. Conversely, static lightpaths may reduce control energy but waste capacity and force other parts of the system to compensate. Energy optimization therefore requires lifecycle accounting rather than instantaneous measurement alone [19], [20].

Reliability is coupled to both latency and energy. Protection switching can reduce outage duration but requires redundant capacity that may remain idle and consume energy. Restoration can use spare capacity more efficiently but may increase recovery latency and create transient congestion. Hardware reliability improves with redundancy, but redundancy increases cost, cooling load, and operational complexity. In large-scale systems, reliability is not simply the absence of component failure; it is the ability to maintain acceptable service under partial failure, attack, misconfiguration, and maintenance [17], [18]. A network optimized only for average latency and average energy may perform poorly during correlated failures or traffic shifts. Multi-objective optimization must therefore consider tail latency, worst-case energy spikes, and recovery-time distributions, not only means.

### **4. Multi-Objective Optimization Frameworks**

Multi-objective optimization provides a formal language for navigating these coupled objectives. In engineering design, Pareto-based methods identify solutions for which no objective can be improved without worsening another. Evolutionary algorithms such as NSGA-II and related approaches are widely used because they can explore non-convex and discontinuous trade-off spaces without requiring gradient information [8], [9]. They are particularly useful when the network model includes discrete choices, such as whether to deploy a particular switch, which wavelengths to assign, or how much redundancy to include. Survey work on multi-objective methods emphasizes that the choice of algorithm is less important than the clarity of objectives, constraints, and decision variables [10], [11]. If reliability is defined only as link availability, for example, the optimizer may ignore correlated failures, software faults, or human error.

Simulation-driven optimization is common in optical network design because analytical models struggle to capture traffic dynamics, protocol interactions, and failure propagation. Simulation allows designers to evaluate latency, energy, and reliability under different demand patterns, but it also introduces assumptions about traffic generation, failure rates, and operational behavior. These assumptions are not neutral. A model that treats energy as a function of average utilization will favor consolidation, while a model that includes cooling dynamics and peak demand may favor distribution. A model that treats failures as independent will underestimate the value of diversity and geographic separation [15], [18]. Multi-objective optimization therefore cannot be separated from model governance. The selection of scenarios, weights, and constraints reflects institutional priorities and should be documented as part of the design process.

Software-defined orchestration can support multi-objective optimization by collecting telemetry, applying policies, and reconfiguring optical resources in response to changing conditions [12], [14]. However, centralized optimization can create scalability and trust problems. If a single controller optimizes latency for high-priority tenants, it may shift energy costs or reliability risks onto others. If optimization is distributed across domains, coordination becomes difficult and local decisions may produce global inefficiency. Hybrid approaches that combine centralized policy with local adaptation are promising, but they require clear interfaces, conflict-resolution rules, and auditability. In this sense, multi-objective optimization is not only a computational problem; it is a governance architecture for allocating scarce network resources among competing claims [10], [20].

## **5. Governance, Deployment, and Sustainability**

Deployment choices shape the achievable trade-offs. Greenfield data centers can design fiber topology, cooling, and power distribution around optical switching from the outset. Brownfield networks must retrofit optical equipment into existing racks, power budgets, and operational procedures. The retrofit path often favors hybrid architectures because they preserve existing electrical switching while introducing optical bypass for high-volume flows [2], [3]. This incremental approach can reduce disruption, but it may lock in inefficiencies and complicate end-to-end optimization. Governance must therefore consider not only the best technical design but also the transition path, including procurement cycles, workforce training, and vendor interoperability [13].

Sustainability adds another layer. Energy consumption in optical networks includes direct electricity use and embodied energy in manufacturing, transportation, and disposal. A network that reduces operational energy but requires frequent hardware replacement may have a worse lifecycle footprint. Similarly, a design that improves reliability through

redundancy may increase material use and electronic waste. Policy frameworks that focus only on operational energy can therefore miss important impacts. Multi-objective optimization should incorporate lifecycle assessment, carbon intensity of electricity sources, and circular economy principles [5], [19]. This is especially important as data centers face scrutiny over water use, land use, and grid impacts.

Governance also affects reliability. Service-level agreements, regulatory requirements, and public safety obligations determine how much redundancy is required and how quickly failures must be repaired. In some contexts, reliability is treated as a market differentiator; in others, it is a baseline public good. The same optical switching architecture may be optimal under a profit-maximizing cloud provider and suboptimal under a public research network with different priorities. Multi-objective optimization must therefore be embedded in a policy process that articulates whose latency, whose energy, and whose reliability matter [6], [17]. Without such articulation, optimization can become a technical exercise that legitimizes existing power arrangements rather than improving infrastructure for broader social benefit.

## **6. Robustness, Fairness, and Policy Implications**

Robustness is the capacity to maintain service across a range of disturbances, including fiber cuts, equipment failures, cyberattacks, and demand shocks. Optical networks are vulnerable to physical threats because fiber routes can be geographically correlated. A single backhoe strike or natural disaster can sever multiple logical paths if they share conduits. Robust design requires diversity, but diversity has costs in fiber, rights-of-way, and energy. Multi-objective optimization can help identify designs that balance cost and resilience, but it must be paired with risk assessment that includes low-probability, high-impact events [6], [18]. It must also consider recovery time objectives, not just availability percentages. A network that is available most of the time but takes hours to recover from a major failure may be unacceptable for emergency services or financial trading.

Fairness is less commonly discussed in optical network optimization, but it is increasingly relevant. If latency and reliability are optimized for premium customers, other users may experience degraded service. If energy-saving policies concentrate traffic on specific links, communities or institutions connected through those links may bear higher failure risk. If deployment prioritizes dense urban areas, rural and remote regions may remain underserved. These are distributional questions, not merely technical ones. Multi-objective optimization can incorporate fairness metrics, such as worst-case latency across regions or equitable access to redundancy, but doing so requires explicit value choices [10], [20]. Policymakers can support fairness by requiring transparency in network performance, mandating minimum resilience standards for critical services, and funding open testbeds that allow smaller operators to evaluate optical switching options.

Privacy and security also intersect with reliability. Optical switching can reduce some attack surfaces by limiting electronic processing, but it can also make traffic analysis and fault diagnosis more difficult. Encrypted traffic may be opaque to network operators, complicating anomaly detection. Control-plane centralization can create a high-value target. Policy responses should include segmentation, authentication, audit trails, and graceful degradation modes that preserve essential services during cyber incidents [12], [18]. The goal is not perfect security, which is unattainable, but resilient operation under adversarial conditions.

## **7. Case Illustrations and Cross-Domain Comparisons**

Intra-data center interconnect provides a clear illustration of coupled objectives. Machine learning training clusters require frequent synchronization and low tail latency, while energy costs from switching, cooling, and power conversion are substantial. Hybrid electrical-optical fabrics can route large flows optically and small control messages electrically, reducing both latency and energy for dominant traffic patterns [1], [3]. Recent work on wavelength-switched and clock-synchronized interconnects using hollow core fiber suggests that latency reduction can be pursued alongside synchronization and reliability, though deployment remains constrained by cost and manufacturing maturity [16]. The optimization problem is not solved once; it evolves as workloads change, hardware improves, and energy prices shift.

Cloud and content delivery networks face a different balance. They must serve geographically distributed users with variable demand, while maintaining availability across regions. Optical backbone switching can reduce transport energy by bypassing electronic routers for long-haul flows, but it requires coordination with caching, load balancing, and content placement [5], [14]. Reliability is often addressed through multi-region redundancy, which increases energy use and cost. A multi-objective framework can help operators choose between redundancy, restoration, and adaptive routing, but the choice depends on contractual obligations and user expectations [17]. Cross-domain comparison with aviation or rail infrastructure is instructive: both rely on scheduled and on-demand capacity, and both face trade-offs between efficiency and resilience. Optical networks differ in their speed and automation, but the governance challenges of safety, fairness, and public accountability are similar.

High-performance computing and research networks add another dimension. These systems often prioritize latency and deterministic performance over energy minimization, especially for scientific workflows with strict deadlines. However, energy costs and sustainability mandates are increasingly important. Multi-objective optimization can reveal configurations that preserve scientific throughput while reducing idle power and cooling demand [19], [20]. Public research networks can also serve as testbeds for fair and transparent optimization, because their governance structures include diverse stakeholders. Lessons from these environments can inform commercial deployments, particularly regarding open interfaces, reproducible benchmarks, and independent auditing.

## **8. Conclusion**

Multi-objective optimization of latency, energy consumption, and reliability in optical switching networks is a system-level challenge that cannot be resolved by isolated hardware improvements or algorithmic tuning. Latency, energy, and reliability are coupled through switching granularity, control-plane design, topology, redundancy, and operational governance. Circuit, wavelength, burst, packet, and hybrid switching each produce distinct trade-off profiles, and no single architecture dominates across all contexts. Pareto-based optimization and simulation-driven design can help identify feasible compromises, but their outputs depend on assumptions about traffic, failures, costs, and social priorities. These assumptions must be made explicit and subject to stakeholder scrutiny.

The paper has argued for a socio-technical approach that treats optical switching networks as infrastructure with distributional consequences. Sustainability requires lifecycle accounting, not just operational energy reduction. Robustness requires attention to correlated failures, recovery time, and graceful degradation. Fairness requires consideration of who benefits from low latency and high reliability and who bears the risks of energy-saving consolidation or deferred maintenance. Policy can support these goals through transparency requirements, resilience standards, open testbeds, and procurement rules that value lifecycle performance.

Future research should develop integrated models that combine optical physics, network control, energy markets, and governance processes. Such models will not produce a single optimal design, but they can help societies make more defensible choices about the infrastructures on which digital life depends.

## References

1. Kachris, C., & Tomkos, I. (2012). A survey on optical interconnects for data centers. *IEEE Communications Surveys & Tutorials*, 14(4), 1021–1036. <https://doi.org/10.1109/SURV.2011.122111.00069>
2. Benner, A. F., Ignatowski, M., Kash, J. A., Kuchta, D. M., & Ritter, M. B. (2005). Exploitation of optical interconnects in future server architectures. *IBM Journal of Research and Development*, 49(4/5), 755–775. <https://doi.org/10.1147/rd.494.0755>
3. Farrington, N., Porter, G., Radhakrishnan, S., Bazzaz, H. H., Subramanya, V., Fainman, Y., Papen, G., & Ford, J. (2010). Helios: A hybrid electrical/optical switch architecture for modular data centers. *ACM SIGCOMM Computer Communication Review*, 40(4), 339–350. <https://doi.org/10.1145/1851182.1851223>
4. Tucker, R. S. (2011). Green optical communications—Part II: Energy limitations in networks. *IEEE Journal of Selected Topics in Quantum Electronics*, 17(2), 261–274. <https://doi.org/10.1109/JSTQE.2010.2051216>
5. Baliga, J., Ayre, R. W. A., Hinton, K., & Tucker, R. S. (2011). Green cloud computing: Balancing energy in processing, storage, and transport. *Proceedings of the IEEE*, 99(1), 149–167. <https://doi.org/10.1109/JPROC.2010.2060451>
6. Grover, W. D. (2004). *Mesh-based survivable networks: Options and strategies for optical, MPLS, SONET, and ATM networking*. Prentice Hall.
7. Ramaswami, R., Sivarajan, K. N., & Sasaki, G. H. (2010). *Optical networks: A practical perspective* (3rd ed.). Morgan Kaufmann.
8. Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182–197. <https://doi.org/10.1109/4235.996017>
9. Zitzler, E., Deb, K., & Thiele, L. (2000). Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary Computation*, 8(2), 173–195. <https://doi.org/10.1162/106365600568202>
10. Marler, R. T., & Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26(6), 369–395. <https://doi.org/10.1007/s00158-003-0368-6>
11. Coello Coello, C. A., Lamont, G. B., & Van Veldhuizen, D. A. (2007). *Evolutionary algorithms for solving multi-objective problems* (2nd ed.). Springer.
12. Channegowda, M., Nejabati, R., & Simeonidou, D. (2013). Software-defined optical networks technology and infrastructure: Enabling software-defined optical network operations. *Journal of Optical Communications and Networking*, 5(10), A274–A282. <https://doi.org/10.1364/JOCN.5.00A274>
13. Liu, L., Peng, D., Nejabati, R., & Simeonidou, D. (2014). Field trial of an OpenFlow-based unified control plane for multilayer multigranularity optical switching networks.

Journal of Lightwave Technology, 32(4), 805–817.

<https://doi.org/10.1109/JLT.2013.2285150>

14. Casellas, R., Martinez, R., Munoz, R., & Vilalta, R. (2015). SDN orchestration of OpenFlow and GMPLS flexi-grid networks. *Journal of Optical Communications and Networking*, 7(1), A1–A9. <https://doi.org/10.1364/JOCN.7.0000A1>
15. Simmons, J. M. (2014). *Optical network design and planning* (2nd ed.). Springer. <https://doi.org/10.1007/978-3-319-05227-4>
16. Zhou, Zichuan, et al. "Low-latency wavelength-switched clock-synchronized intra-data center interconnects enabled by hollow core fiber." *Optics Express* 31.15 (2023): 24739-24748.
17. Vasseur, J.-P., Pickavet, M., & Demeester, P. (2004). *Network recovery: Protection and restoration of optical, SONET-SDH, IP, and MPLS*. Morgan Kaufmann.
18. Sterbenz, J. P. G., Hutchison, D., Çetinkaya, E. K., Jabbar, A., Rohrer, J. P., Schöller, M., & Smith, P. (2010). Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines. *Computer Networks*, 54(8), 1245–1265. <https://doi.org/10.1016/j.comnet.2010.03.005>
19. Barroso, L. A., Clidaras, J., & Hölzle, U. (2013). *The datacenter as a computer: An introduction to the design of warehouse-scale machines* (2nd ed.). Morgan & Claypool. <https://doi.org/10.2200/S00516ED2V01Y201306CAC024>
20. Kliazovich, D., Bouvry, P., & Khan, S. U. (2012). GreenCloud: A packet-level simulator of energy-aware cloud computing data centers. *The Journal of Supercomputing*, 62(3), 1263–1283. <https://doi.org/10.1007/s11227-010-0504-1>