Multi-Phase Candidate Selection for Failure-Aware and Energy-Efficient SDN Networks

Kevin Saavedra
DIMES, University of Calabria
Via P. Bucci, 87036 Rende (CS), Italy
saavedrosky70@gmail.com

Mauro Tropea
DIMES, University of Calabria
Via P. Bucci, 87036 Rende (CS), Italy
m.tropea@dimes.unical.it

Mattia Giovanni Spina
DIMES, University of Calabria
Via P. Bucci, 87036 Rende (CS), Italy
mattiagiovanni.spina@dimes.unical.it

Floriano De Rango
DIMES, University of Calabria
Via P. Bucci, 87036 Rende (CS), Italy
derango@dimes.unical.it

***Abstract*—Candidate node selection is a fundamental mechanism for enabling fast traffic recovery after link failures in hybrid IP/Software-Defined Networking environments. However, existing approaches generally assume homogeneous hardware deployments and select candidate nodes without considering the additional traffic load induced by failure-driven rerouting. This paper proposes a multi-phase candidate selection framework that jointly addresses network resilience, hardware capacity planning, and energy efficiency. In the first phase, the framework evaluates the traffic load generated under every single-link failure during the busy hour and determines, for each node, its worst-case operating condition. This information is used to assign the minimum hardware tier capable of supporting the resulting load, replacing conventional fixed-capacity margins with a topology- and traffic-aware dimensioning process. In the second phase, a greedy candidate set providing complete failure coverage is constructed and subsequently extended through a structured neighbourhood exploration. The final candidate set is selected by minimising a weighted objective that combines candidate-node power consumption and recovery propagation delay. A configurable parameter allows network operators to balance these objectives without modifying the hardware allocation determined in the first phase. Evaluations on three SNDLib topologies with real traffic matrices demonstrate that the proposed framework guarantees complete single-link failure coverage while reducing total network power consumption by up to 27.8% compared with a homogeneous carrier-grade deployment.**

***Index Terms*—SDN, candidate placement, single-link failure, green networking, hardware tier selection, energy efficiency**

I. Introduction

Software-Defined Networking (SDN) decouples the control and forwarding planes, enabling centralised and programmable traffic management [1]. This flexibility has fostered innovative applications across a wide range of domains, including ad hoc networks, such as Flying Ad Hoc Networks (FANETs) [2], [3], and Artificial Intelligence (AI). As SDN deployments mature, however, operators increasingly face the challenge of selecting the most appropriate hardware for each network node. Although traffic loads can vary significantly across nodes, existing candidate-selection algorithms typically assume hardware-homogeneous infrastructures, in which every node is equipped with identical switching hardware [4]–[7]. This assumption reflects the traditional practice of over-provisioning Wide Area Networks (WANs) [8] to accommodate peak-hour demand and provide sufficient failure headroom. While this strategy improves reliability, it also requires high-end routing equipment to be deployed throughout the network [1], including at nodes whose actual traffic load could be handled by less powerful and more energy-efficient devices. A key reliability requirement is the rapid recovery from *single-link failures* (SLFs). In the conventional SDN approach, a subset of switches is designated as *candidate nodes*. When a link failure occurs, the switch adjacent to the failed link tunnels the disrupted traffic to a pre-selected candidate node, which then forwards it towards the affected destinations [7]. The selection of these nodes, commonly referred to as the Candidate Selection problem, has received considerable attention [4]–[7]. However, existing algorithms assume hardware-homogeneous networks and therefore overlook the differences in processing capacity and energy consumption among switch types. To the best of our knowledge, no previous work jointly addresses hardware-tier assignment and candidate-node placement with the objective of minimising energy consumption while guaranteeing full SLF coverage. A straightforward approach to handling SLFs is to provision a fixed capacity margin above the busy-hour traffic load. However, this margin cannot be assumed to provide sufficient protection. Failure-induced rerouting may concentrate traffic along alternative paths, causing intermediate nodes to experience loads that significantly exceed the provisioned margin [9]. Such overload conditions can only be identified by explicitly computing the traffic distribution under each failure scenario.

This paper makes two contributions:

1) **Failure-scenario capacity planning.** For each node, we compute the maximum traffic load it would carry under every possible single link failure during the busy hour

979-8-3195-0167-7/26/$31.00 ©2026 IEEE

and use this value to select the minimum viable hardware tier. This replaces an unverified fixed margin with a data-driven, topology-aware guarantee.

2) **Energy-efficient candidate placement.** Given the hardware assignments from Phase 1, we select candidate nodes by minimising a weighted combination of candidate-set energy and average recovery propagation delay over a structured candidate neighbourhood. A parameter $\alpha \in [0, 1]$ allows operators to continuously trade energy minimisation against delay minimisation.

We evaluate on three widely used benchmark network topologies from SNDLib [10] – Abilene, Germany50, and Nobel-Germany – with real 5-minute traffic matrices [10], [11]. As hardware tier representatives, we use the ZodiacFX (a low-power OpenFlow research switch with published specifications [12]) and the NEC PF5240 (a carrier-grade OpenFlow switch [13]), which provide concrete, publicly documented power and capacity values representative of the low-power and carrier-grade ends of the commercial OpenFlow switch range. The methodology itself is independent of the specific hardware parameters chosen. The remainder of this paper is organized as follows: Section II provides a brief overview of the related work. Section III and Section IV delve into the details of the proposal, with Section V describing the candidate selection approach. Section VI and Section VII showcase the experimental campaign and the discussion of the obtained results, respectively. Finally, Section VIII concludes the paper.

II. Related Work

In this section, literature works concerning network programmability [14] are briefly described. The authors in [15] proposed SCOPE, a cost-efficient and QoS-aware scheme for hybrid SDN deployment. The approach combines priority-based SDN switch selection with the incremental placement of controllers according to the available budget. Its objectives include increasing programmable traffic and reducing QoS violations. Nevertheless, it does not address hardware-tier assignment, failure-scenario capacity planning, or candidate-node selection. The work proposed in [16] investigates the energy-efficient migration of legacy networks towards SDN through a multistage upgrade process. Their method jointly determines which switches and controllers should be upgraded at each stage while considering deployment costs and traffic requirements. Energy savings are mainly achieved by rerouting traffic and disabling unused network resources. However, the method does not dimension switch hardware according to failure-induced loads or optimise candidate placement for SLF recovery. Wang et al. [17] proposed an energy-efficient SDN management method based on dynamic topology switching and reliability constraints. The approach reduces energy consumption by adapting the set of active network resources while maintaining backup connectivity against single-link failures. It therefore considers reliability within energy-aware network operation. However, it does not jointly determine heterogeneous hardware tiers and candidate-node placement based on the traffic generated under each failure scenario. Failures can be also caused by malicious entities that can exploit the SDN programmability to target the controller with the aim of turning it down through Denial of Service (DDoS) attacks , trying, therefore, to disrupt the whole network [18], [19]. The proposal described in [20] jointly addressed switch upgrades and controller deployment in hybrid SDN infrastructures. The optimisation maximises the amount of traffic managed through SDN while minimising switch-to-controller propagation delay under a limited deployment budget. Exact and heuristic solutions were proposed for networks of different sizes. However, the approach does not consider heterogeneous switch capacities, energy consumption, or failure-induced traffic. Chu et al. [4] introduced the first Greedy Candidate Selection Algorithm (CSA) using a binary candidate table with greedy column-sum maximisation. Li et al. [5] proposed a search-tree variant with shorter repair paths. Yang and Yeung [6] reduced candidate count via multi-tunnel repair. Vuppalapati and Venkatesh [7] established the Minimum Candidate Selection (MCS) algorithm, which yields the fewest candidates among the methods cited herein and provides the candidate table construction used in our Phase 2; the same authors extended to multi-link failures in [21]. Adjacent work addresses energy and fault tolerance in hybrid networks. Caria and Jukan [22] compute failure-induced traffic shifts to dimension link bandwidth in hybrid SDN/OSPF networks, but do not address switch hardware selection or candidate placement. Huin et al. [23] enable energy-proportional routing via SDN programmability, targeting normal-operation routing without failure-scenario capacity planning. Galán-Jiménez et al. [24] jointly optimise energy and load balancing under a fixed hardware model, without addressing failure-scenario traffic or candidate placement. The authors of [25] investigated the reliability implications of migrating legacy networks towards SDN-based infrastructures. Their analysis considers nodal reachability, network connectivity, and edge-disjoint routing under link-failure conditions. The study shows that SDN-based networks can achieve operational reliability comparable to that of traditional architectures. However, it does not address failure-induced traffic loads, heterogeneous hardware dimensioning, or candidate-node placement for fast recovery. To the best of our knowledge, no existing study jointly determines the hardware tier allocation by considering worst-case traffic loads under failure scenarios and optimises the placement of candidate nodes for energy-efficient SLF recovery. Most existing solutions address either resource dimensioning or recovery placement independently, without capturing the mutual impact of resilience requirements and energy consumption.

Main Contributions

Unlike existing candidate-selection approaches, which assume homogeneous switch hardware, and energy-aware solutions, which do not consider failure-driven candidate placement, this work jointly addresses resilience and energy efficiency. Specifically, it: (i) introduces a failure-scenario capacity-planning method that computes the worst-case load of each node across all single-link failures and assigns the

minimum viable hardware tier replacing conventional fixed-capacity margins; (ii) proposes an MCS-based candidate-placement strategy that preserves complete SLF coverage while jointly considering candidate-set power and recovery propagation delay; and (iii) provides a tunable parameter α to balance these objectives without modifying the hardware deployment determined during capacity planning. The evaluation demonstrates that the proposed approach can reduce total network power by up to 27.8% while maintaining 100% SLF coverage.

III. System Model

A. Network Graph

We model the network as an undirected graph $G = (V, E)$ where each node $n \in V$ is defined as an SDN switch and each edge $(u, v) \in E$ is defined as a bidirectional link. Propagation delay on link (u, v) is defined as $w_{uv} = d_{uv}/200\,\text{ms}$, where d_{uv} is the geographic distance in km (Haversine formula) and 200 km/ms approximates the speed of light in fibre. We restrict evaluation to the 2-edge-connected core of each topology, eliminating stub nodes and any bridge edges, retaining only the subgraph in which every node has at least two link-disjoint paths to every other node—a necessary condition for single link failure recovery.

B. Hardware Tiers

We consider two hardware tiers. Each node is assigned to exactly one tier; the assignment is a design variable. For evaluation, we use two commercially available OpenFlow switches with published power and capacity specifications as tier representatives (Table I). We define P_{base} as the baseline idle chassis power of the device, P_{port} as the incremental per-port power consumption [26], and μ^{port} as the nominal line-rate port capacity in packets per second (pps) under the traffic model defined below.

- **Low-power tier (ZodiacFX):** a low-power OpenFlow research switch used here as a representative of the low-power device class with fully published power and capacity specifications [12]. Its ASIC (Microchip KSZ8795CLX) operates at wire rate on four 100 Mbps ports. The methodology applies to any device with published specifications.
- **Carrier-grade tier (NEC PF5240):** targets high-traffic nodes. Supports 48×1 GbE ports with a 176 Gbps non-blocking fabric [13].

Port rates are computed using an Ethernet-level IMIX traffic model [27] with three frame sizes representative of backbone traffic: 64 B (minimum Ethernet frame, IEEE 802.3), 594 B (minimum IPv4 datagram reassembly size of 576 B [28] plus 18 B Ethernet overhead), and 1518 B (maximum Ethernet frame, IEEE 802.3), in a 7:4:1 packet ratio. The resulting average frame size and demand-to-load conversion factor are:

$$\bar{b} = \frac{7\times64 + 4\times594 + 1\times1518}{12} = 361.83\,\text{B}, \quad (1)$$

$$f_{\text{IMIX}} = \frac{10^6}{8\,\bar{b}} \approx 345.46\ \text{pps}\,\text{Mbps}^{-1}. \quad (2)$$

TABLE I
Hardware Tier Representatives

Tier	Device	P_{base}	P_{port}	μ^{port}
Low-power	ZodiacFX	15.0 W	0.15 W	34,546 pps
Carrier-grade	NEC PF5240	118.33 W	0.52 W	345,463 pps

μ^{port}: line-rate capacity at $\bar{b} = 361.83$ B average frame size (1)

The total power of node n is:

$$P(n) = P_{\text{base}}(\text{hw}(n)) + \deg(n) \cdot P_{\text{port}}(\text{hw}(n)), \quad (3)$$

where $\text{hw}(n) \in \{\text{low-power, carrier-grade}\}$ is the hardware tier assigned to node n, and $\deg(n)$ is its number of active ports. We neglect the control-plane power component P_{control}, which is small relative to P_{base} in both devices [26]. The effective per-node capacity is port-limited: $\mu(n) = \deg(n) \cdot \mu^{\text{port}}$.

C. Failure Model

When the link (u, v) fails, node u can no longer forward traffic through v. We define the *affected destination set* $\mathcal{AD}(u, v)$ as all destinations d for which v lies on at least one shortest-delay path from u to d in the intact network. By the shortest-path optimality condition [29, Lemma 24.10], this holds if and only if:

$$d \in \mathcal{AD}(u, v) \iff |w_{uv} + \delta(v, d) - \delta(u, d)| < \varepsilon, \quad (4)$$

where $\delta(\cdot, \cdot)$ is the delay-weighted shortest path length on G and $\varepsilon = 10^{-6}$ ms handles floating-point precision. Destination d is marked affected if v lies on *any* shortest-delay path from u to d—a deliberately conservative criterion: even if deterministic Dijkstra routing does not actually traverse v due to tie-breaking, the candidate table still covers that failure, guaranteeing full coverage under any single shortest-path routing policy. A node $h \neq u$ is a *valid candidate node* for failure (u, v) if, in the post-failure network $G \setminus \{(u, v)\}$: (1) u can reach h, and (2) h can reach every $d \in \mathcal{AD}(u, v)$. Upon failure, u tunnels all traffic for affected destinations to h [7], which forwards each flow via shortest-delay paths. The average recovery propagation delay is:

$$\Delta(u, v, h) = \delta_{(u,v)}(u, h) + \frac{1}{|\mathcal{AD}(u, v)|} \sum_{d \in \mathcal{AD}(u,v)} \delta_{(u,v)}(h, d), \quad (5)$$

where $\delta_{(u,v)}(\cdot, \cdot)$ denotes shortest-path delay in the post-failure graph $G \setminus \{(u, v)\}$.

IV. Failure-Scenario Capacity Planning

A. The Insufficiency of a Fixed Margin

The standard 80 % busy-hour utilisation target [9] reserves a 20 % headroom for demand variability under normal operation—insufficient for failure rerouting. When a link fails, its traffic reroutes through alternative paths, loading intermediate nodes with their own traffic *plus* rerouted flows; the spike depends on topology and link load, not on a fixed margin. We define the *dimensioning-critical node* as the node reaching the

TABLE II
BUSY-HOUR LOAD VS. WORST SINGLE-LINK FAILURE LOAD AT THE DIMENSIONING-CRITICAL NODE PER TOPOLOGY.

Topology	Busy-hour	Worst failure	Surge
Abilene	405,740 pps	690,923 pps	**+70 %**
Germany50	959,164 pps	1,061,428 pps	+11 %
Nobel-Germany	1,002,417 pps	1,176,291 pps	+17 %

highest absolute worst-case load across all single-link failures. As Table II shows, on Abilene (WASHng) the surge reaches 70 %: a node provisioned at 80 % utilisation would be driven to 136 %, well beyond its provisioned capacity.

B. Traffic Model

Let $M = \{d_{ij}\}_{i,j \in V}$ be a demand matrix (Mbps), routed via delay-weighted shortest-path forwarding (ties broken deterministically). Let G' be any connected spanning subgraph of G. The aggregate packet rate λ_n (pps) at node n under M on G' is:

$$\lambda_n(G', M) \;=\; f_{\mathrm{IMIX}} \sum_{\substack{(i,j):\, d_{ij}>0 \\ n \,\in\, p^*(G',i,j)}} d_{ij}, \tag{6}$$

where $p^*(G', i, j)$ denotes the node sequence of the shortest-delay path from i to j in G', and $f_{\mathrm{IMIX}} = 10^6/(8 \times 361.83\,\mathrm{B}) \approx 345.46\,\mathrm{pps\,Mbps^{-1}}$ converts megabit-per-second demand values to packets per second at the IMIX-derived average frame size of 361.83 B (Eq. (1)) [27]. Because evaluation is restricted to the 2-edge-connected core (Section III-A), removing any single edge leaves G' connected and λ_n well-defined. Let $\{M_t\}_{t=1}^{T}$ denote the T available 5-minute snapshots.

C. Computing the Worst-Case Load

We identify the busy-hour demand matrix M_{BH} as the snapshot with the highest aggregate network load:

$$M_{\mathrm{BH}} = \arg \max_{t \in \{1,\ldots,T\}} \sum_{n \in V} \lambda_n(G,\, M_t), \tag{7}$$

This gives a physically realizable matrix in which all node loads are temporally co-occurrent. The failure-scenario worst-case load λ_n^{fs} is the maximum load node n sustains across all single link failures:

$$\lambda_n^{\mathrm{fs}} = \max_{(u,v) \in E} \lambda_n\big(G \setminus \{(u,v)\},\, M_{\mathrm{BH}}\big). \tag{8}$$

This value is computed directly from measured traffic and the network graph with no assumptions about load distribution. By using 5-minute averages, λ_n^{fs} *understates* the true worst case: within-interval peaks are smoothed out, so surge figures in Table II are conservative lower bounds.

D. Hardware Tier Assignment

Node n is assigned the low-power tier if its failure-scenario worst-case load remains within 80 % of the low-power tier's port capacity; otherwise it receives the carrier-grade tier:

$$\mathrm{hw}(n) = \begin{cases} \text{Low-power} & \lambda_n^{\mathrm{fs}} < 0.80 \cdot \mu^{\mathrm{LP}}(n), \\ \text{Carrier-grade} & \text{otherwise,} \end{cases} \tag{9}$$

where $\mu^{\mathrm{LP}}(n) = \min\big(\mu_{\max}^{\mathrm{LP}},\ \deg(n) \cdot \mu^{\mathrm{port,LP}}\big)$ is the effective capacity of the low-power tier at node n, capped by the device's physical port limit $\mu_{\max}^{\mathrm{LP}}$. This guarantees that no node exceeds 80 % utilisation under any single link failure during the busy hour, for the observed 5-minute traffic averages. Because hardware is assigned once from the traffic data—before any candidate placement decision—the total network power

$$E_{\mathrm{total}} = \sum_{n \in V} P(n) \tag{10}$$

is fully determined at this stage via (3), independently of the subsequent candidate node selection.

V. CANDIDATE PLACEMENT

A. Candidate Table

For each directed link failure $\ell = (u, v)$, we derive $\mathcal{AD}(\ell)$ via (4) and check, for each node $h \neq u$, both connectivity conditions in $G \setminus \{(u, v)\}$: (1) u can reach h, and (2) h can reach every $d \in \mathcal{AD}(\ell)$. Since G is restricted to its 2-edge-connected core (Section III-A), $\mathcal{H}(\ell)$ is non-empty for every failure with $\mathcal{AD}(\ell) \neq \emptyset$. Following [7], the candidate table has $T(\ell, h) = 1$ iff h satisfies both conditions for failure ℓ.

B. Greedy Base Set and Neighbourhood

Greedy set-cover on T finds a minimal $\mathcal{B} \subseteq V$ such that every non-trivial failure has a valid candidate in $\mathcal{B}$ [7]. We extend $\mathcal{B}$ with up to two additional nodes:

$$\mathcal{C} = \{\, \mathcal{B} \cup X : X \subseteq V \setminus \mathcal{B},\ |X| \leq 2 \,\}. \tag{11}$$

where $|X| = 0$ recovers $\mathcal{B}$, $|X| = 1$ adds one node, and $|X| = 2$ adds a pair; every set in $\mathcal{C}$ covers all failures by construction. The limit $|X| \leq 2$ balances solution quality against search cost: on Germany50 ($|\mathcal{C}| = 1{,}177$) the full sweep completes in under 3 min; $|X| \leq 3$ would expand $|\mathcal{C}|$ by more than an order of magnitude ($\approx 15\times$) for marginal gain, since the baseline $\mathcal{B}$ already covers every failure.

C. Objective and Candidate Node Selection

Let $\mathcal{F} = \{\ell \in E : \mathcal{AD}(\ell) \neq \emptyset\}$ be the set of non-trivial failures, equivalently the rows of T with at least one valid candidate. For each candidate set $\mathcal{S} \in \mathcal{C}$, the average recovery propagation delay is:

$$D(\mathcal{S}) = \frac{1}{|\mathcal{F}|} \sum_{\ell \in \mathcal{F}} \min_{h \in \mathcal{S} \cap \mathcal{H}(\ell)} \Delta(\ell, h), \tag{12}$$

Algorithm 1 Failure-Scenario Sizing and Candidate Placement

Input: $G = (V, E)$, traffic snapshots $\{M_t\}$, $\alpha \in [0, 1]$
Output: Hardware map $\mathrm{hw}(\cdot)$, candidate set $\mathcal{S}^*$

// Phase 1: Failure-Scenario Capacity Planning
$M_{\mathrm{BH}} \leftarrow \arg\max_t \sum_n \lambda_n(G, M_t)$
$\lambda_n^{\mathrm{fs}} \leftarrow 0$ **for all** $n \in V$
for each undirected link $(u, v) \in E$ **do**
  // one removal per physical link
  Remove (u, v); route M_{BH} on $G \setminus \{(u, v)\}$
  $\lambda_n^{\mathrm{fs}} \leftarrow \max(\lambda_n^{\mathrm{fs}}, \lambda_n)$ **for all** $n \in V$
  Restore (u, v)
end for
Assign $\mathrm{hw}(n)$ via (9) **for all** $n \in V$ *// E_{total} fixed here*
// Phase 2: Candidate Placement
Pre-compute $\delta(\cdot, \cdot)$ on G
for each directed link $\ell = (u, v) \in E$ **do**
  $\mathcal{AD}(\ell) \leftarrow$ (4); $\mathcal{H}(\ell) \leftarrow$ valid candidates in $G \setminus \{\ell\}$
end for
$\mathcal{B} \leftarrow$ GREEDYSETCOVER$(\mathcal{H})$; $\mathcal{C} \leftarrow$ neighbourhood via (11)
$\mathcal{S}^* \leftarrow \arg\min_{\mathcal{S} \in \mathcal{C}} \left[\alpha \frac{E(\mathcal{S})}{E_{\max}} + (1-\alpha) \frac{D(\mathcal{S})}{D_{\max}}\right]$ (13)
return $\mathrm{hw}(\cdot)$, $\mathcal{S}^*$

where $\mathcal{H}(\ell) = \{h : T(\ell, h) = 1\}$ is the valid candidate set for failure ℓ from the candidate table T, and $\Delta(\ell, h) \triangleq \Delta(u, v, h)$ (5) for $\ell = (u, v)$ is the recovery propagation delay. The selected set minimises:

$$\mathcal{S}^*(\alpha) = \arg\min_{\mathcal{S} \in \mathcal{C}} \left[\alpha \cdot \frac{E(\mathcal{S})}{E_{\max}} + (1 - \alpha) \cdot \frac{D(\mathcal{S})}{D_{\max}}\right], \quad (13)$$

where $E(\mathcal{S}) = \sum_{n \in \mathcal{S}} P(n)$ (3) is the candidate-set power, $E_{\max} = \sum_{n \in V} [P_{\mathrm{base}}^{\mathrm{CG}} + \deg(n) \cdot P_{\mathrm{port}}^{\mathrm{CG}}]$ is the all-carrier-grade upper bound (Table I), and $D_{\max}$ is a fixed normalisation parameter, set to 100 ms across all experiments; this value exceeds the maximum observed recovery delay in our evaluation (20 ms on Abilene). Both denominators are constant across candidate sets, so score rankings are stable across evaluations.

Remark on α. E_{total} (10) is fixed by Phase 1; $E(\mathcal{S})$ is the aggregate power of the candidate set. Although E_{total} does not change with candidate selection, minimising $E(\mathcal{S})$ at $\alpha = 1$ preferentially assigns the candidate role to lower-power, lower-load nodes—which by Phase 1 construction have spare capacity headroom during failure rerouting. At $\alpha = 0$ the algorithm minimises recovery delay; intermediate α trades delay against candidate-set power at no cost to total network energy.

The full procedure is given in Algorithm 1. On Germany50 (50 nodes, 88 links) the complete procedure runs in under 3 min on commodity hardware, which is acceptable for offline capacity planning.

VI. Performance Evaluation

A. Setup

We evaluate on three benchmark network topologies from SNDLib [10]:

- **Abilene**: 11 nodes, 14 links. Zhang et al. [11] traffic dataset: 51 snapshots at 5-min granularity (2004-03-01, 00:00–04:10).
- **Germany50**: 50 nodes, 88 links. DFN-2005 dataset [10]: 52 snapshots at 5-min granularity (2005-02-15, 00:00–04:15).
- **Nobel-Germany**: 17 nodes, 26 links. DFN-2005 dataset [10]: 52 snapshots at 5-min granularity (2005-02-15, 00:00–04:15).

These topologies are used as representative network graph benchmarks. The hardware parameters from Table I provide concrete, published values for the two tiers; the methodology is applicable to any pair of hardware specifications. Algorithm 1 is evaluated for selected $\alpha \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$; representative values are reported below.

B. Load Surge Under Link Failure

Table II quantifies the surge at each topology's dimensioning-critical node. On Abilene (WASHng), the worst failure raises load from 405,740 to 690,923 pps—a 70 % surge that drives a node provisioned at 80 % utilisation to a capacity ceiling of 507,175 pps, resulting in 136 % utilisation, exceeding the node's provisioned capacity. On Germany50 and Nobel-Germany the surges (11 % and 17 %) fall within 20 %, but this can only be established by running the Phase 1 computation; a fixed margin gives no such guarantee a priori.

C. Energy Savings and Hardware Composition

Phase 1 assigns 16 of 50 Germany50 nodes to the low-power tier, dropping total power from 6,008.0 W to **4,337.7 W** (**27.8 % saving**). On Nobel-Germany, 1 of 17 nodes qualifies, yielding 5.1 % savings (1,934.6 W vs. 2,038.7 W). On Abilene, every failure-scenario load exceeds the ZodiacFX threshold; all 11 nodes are carrier-grade and no saving is recorded—the correct outcome. Fig. 1 compares total power against the uniform carrier-grade baseline. Fig. 2 shows tier distribution across candidate and passive roles: at $\alpha = 0.4$ on Germany50, candidate nodes account for only 166 W (3.8 % of total), with 13 low-power passive nodes contributing the bulk of the savings. Candidate designation does not change hardware or power draw—E_{total} is fully determined by Phase 1.

D. Recovery Delay, α, and Candidate Placement

Table III and Fig. 3 show performance across α. Total power is identical at all α within each topology, confirming Phase 2 operates energy-neutrally within the Phase 1 hardware budget. On Germany50, delay spans 2.57 ms ($\alpha = 0$, four candidates) to 3.57 ms ($\alpha = 1$, two candidates)—a 39 % spread at zero energy cost. Increasing α reduces candidate count ($4 \to 3 \to 2$) while shifting composition toward lower-power

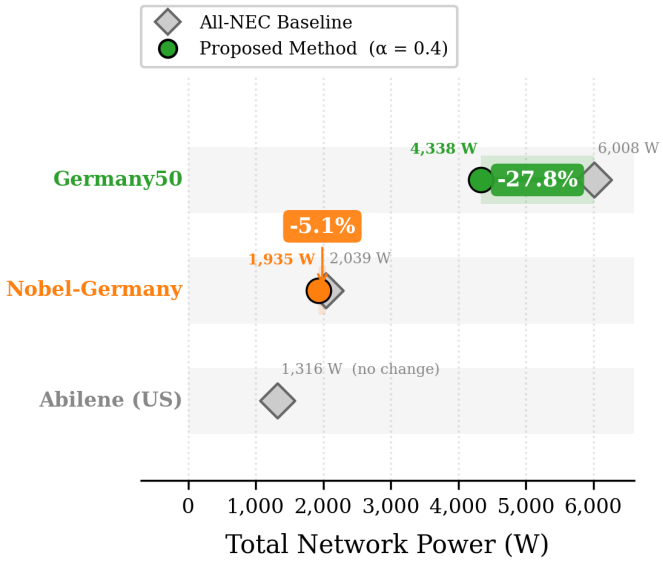


Fig. 1. Total network power vs. uniform carrier-grade baseline at $\alpha = 0.4$. Savings derive entirely from Phase 1 hardware assignment; candidate placement does not affect total power.

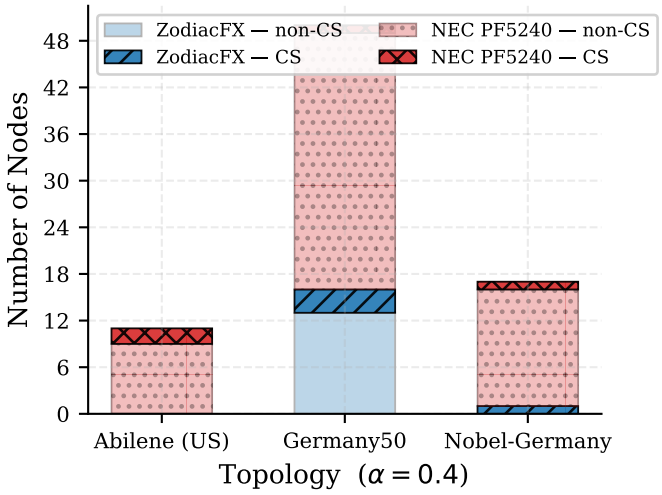


Fig. 2. Hardware composition at $\alpha = 0.4$: low-power (ZodiacFX) and carrier-grade (NEC) nodes by candidate node and passive role.

nodes. Nobel-Germany spans 2.53–3.37 ms; Abilene spans 17–20 ms (all carrier-grade, longer geographic distances). Fig. 4 shows geographic placement at $\alpha = 0.4$: Germany50 candidates (Aachen, Bremerhaven, Hamburg, Regensburg) are distributed across the graph for short average repair paths. Fig. 5 decomposes power by role, confirming passive nodes dominate total consumption.

E. Comparison with MCS Baseline

Table IV compares the MCS baseline (Vuppalapati 2022 [7]: uniform NEC hardware, greedy base set $\mathcal{B}$) against our method at $\alpha = 0$ (delay-optimal) and $\alpha = 1$ (energy-optimal). MCS baseline delay is recomputed using the same propagation-delay metric (5) applied to $\mathcal{B}$, ensuring a metric-consistent

TABLE III
RESULTS AT SELECTED α VALUES. † = ZODIACFX (LOW-POWER); OTHERS = NEC PF5240.

Topology	α	Candidates	Power (W)	Delay (ms)
Abilene	0.0	4 (CG): ATLAng, CHINng, HSTNng, SNVAng	1,316.2	17.45
	1.0	2 (CG): ATLAng, CHINng	1,316.2	20.38
Germany50	0.0	4 (1LP+3CG): Bremerhaven†, Aachen, Braunschweig, Wuerzburg	4,337.7	2.57
	0.4	4 (3LP+1CG): Bremerhaven†, Hamburg†, Regensburg†, Aachen	4,337.7	2.92
	0.6	3 (2LP+1CG): Bremerhaven†, Regensburg†, Aachen	4,337.7	3.11
	1.0	2 (1LP+1CG): Bremerhaven†, Aachen	4,337.7	3.57
Nobel-Germany	0.0	4 (1LP+3CG): Norden†, Hannover, Frankfurt, Nuernberg	1,934.6	2.53
	1.0	2 (1LP+1CG): Norden†, Hannover	1,934.6	3.37

TABLE IV
PROPOSED METHOD VS. MCS BASELINE [7] AT $\alpha = 0$ (DELAY-OPTIMAL) AND $\alpha = 1$ (ENERGY-OPTIMAL). SAVING: POWER REDUCTION VS. UNIFORM-NEC BASELINE; **BOLD** = IMPROVEMENT OVER MCS.

Topology	Method	$\lvert\mathcal{S}\rvert$	Power (W)	Saving (%)	Delay (ms)
Abilene	MCS baseline	2	1,316.2	0.0	20.38
	Ours α=0	4	1,316.2	0.0	**17.45**
	Ours α=1	2	1,316.2	0.0	20.38
Germany50	MCS baseline	2	6,008.0	0.0	3.57
	Ours α=0	4	**4,337.7**	**27.8**	**2.57**
	Ours α=1	2	**4,337.7**	**27.8**	3.57
Nobel-Ger.	MCS baseline	2	2,038.7	0.0	3.37
	Ours α=0	4	**1,934.6**	**5.1**	**2.53**
	Ours α=1	2	**1,934.6**	**5.1**	3.37

comparison. On Abilene every node requires carrier-grade hardware, so Phase 1 produces no energy saving relative to uniform NEC; our method still reduces average recovery delay by 14 % at $\alpha = 0$ by extending the candidate set to four nodes. On Germany50, Phase 1 hardware assignment alone reduces total network power by 27.8 % relative to the uniform-NEC baseline; at $\alpha = 0$ the additional candidates further cut recovery delay by 28 %. At $\alpha = 1$ our two-candidate set matches MCS delay exactly while consuming 27.8 % less power. Nobel-Germany shows the same pattern at a smaller energy margin (5.1 %). Our method strictly dominates MCS on at least one objective – with no degradation on the other – across all three topologies.

VII. RESULTS DISCUSSION

Why a fixed margin cannot replace failure-scenario computation. Table II illustrates two distinct situations that a fixed margin conflates. On Germany50 and Nobel-Germany, the surge (11 % and 17 %) happens to fall within a 20 % margin – but this can only be verified by computing the

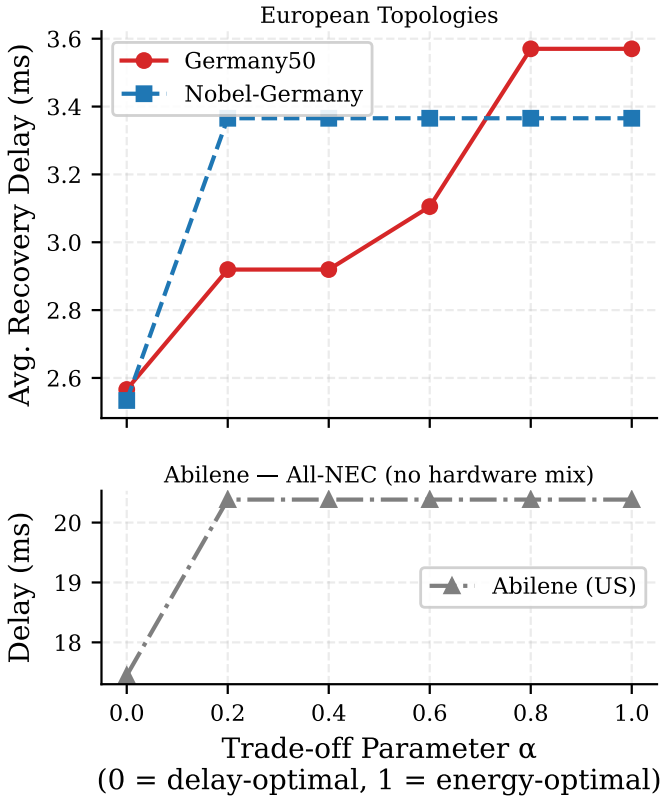


Fig. 3. Average recovery propagation delay vs. α. Total power is constant at all α values. Top: Germany50 and Nobel-Germany (2–4 ms). Bottom: Abilene (17–20 ms, all carrier-grade).

actual failure-scenario traffic. Without the computation, an operator has no way to know whether the margin is sufficient. On Abilene, the surge (70 %) far exceeds the 20 % margin. A fixed-margin approach would provision the dimensioning-critical node to $405{,}740/0.80 = 507{,}175$ pps, leaving it at 136 % utilisation during the worst failure – a condition that Phase 1 explicitly avoids by computing the actual worst-case load. The computation costs one shortest-path routing per link in the topology – tractable for WAN-scale networks of up to hundreds of nodes.

Hardware sizing and candidate placement are decoupled. E_{total} (10) depends only on the hardware assignments produced by Phase 1. Phase 2 selects which already-deployed nodes act as candidates, changing the routing policy but not the hardware. An operator can therefore reconfigure candidate placement – adjusting α, adding candidates for redundancy, or updating the candidate table after a topology change – without any hardware modification.

Limitations. The methodology selects M_{BH} as the single snapshot of maximum aggregate load and computes failure-scenario loads only for that matrix. This assumes the worst-case failure load at every node coincides with the busy-hour period; nodes whose surge peaks during off-peak hours may be under-provisioned. Extending Phase 1 to take the element-wise maximum across all snapshots is straightforward and is left for future work. The candidate neighbourhood (11) considers the greedy base set extended by up to two nodes; larger extensions are not evaluated. Recovery delay is modelled as propagation delay only; queuing at elevated post-failure loads is deferred to future work.

VIII. Conclusion

We showed that the standard 20 % capacity margin used in traffic engineering is insufficient for failure rerouting in SDN networks: actual load surges computed from real traffic matrices reach 70 % on benchmark topologies—more than three times the margin. We proposed a failure-scenario capacity planning methodology that determines, for each node, the worst-case load across all single link failures during the busy hour and uses it to assign the minimum viable hardware tier. On the Germany50 benchmark, this reduces total network power by 27.8 % relative to a uniform carrier-grade deployment while guaranteeing full SLF coverage and sub-4 ms average recovery delay on Germany50 and Nobel-Germany (Abilene requires all carrier-grade nodes, yielding 17–20 ms over longer geographic distances). The parameter α provides an energy-neutral knob to tune the balance between delay performance and candidate-set simplicity. Both contributions—hardware sizing and candidate node placement—follow directly from measured traffic and topology data, requiring no assumptions about the distribution of failure-induced load beyond what the network itself determines.

Acknowledgment

This work was partially supported by Project PON SUNS - Secure Ubiquitous Networks and Services financed by MISE, (Azione 1.1.2) – CUP B29H26000360007. Period:2026-2028.

References

[1] S. Jain, A. Kumar, S. Mandal, J. Ong, L. Poutievski, A. Singh, S. Venkata, J. Wanderer, J. Zhou, M. Zhu, J. Zolla, U. Hölzle, S. Stuart, and A. Vahdat, "B4: Experience with a globally-deployed software defined WAN," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 3–14, 2013.

[2] M. Tropea, C. E. Quispe Aldana, E. P. de Freitas, and F. De Rango, "SFEM3: A performance comparative analysis for SDN-based FANET Emulation using Mininet-WiFi and ns-3," *Ad Hoc Networks*, vol. 177, p. 103859, Oct. 2025.

[3] M. G. Spina, M. Tropea, and F. De Rango, "SURA-LB: Software-defined IDS with UAV Resource Aware Load-Balancing in FANET disaster scenarios," *Comput. Commun.*, vol. 223, pp. 101–114, Jul. 2024.

[4] C.-Y. Chu, K. Xi, M. Luo, and H. J. Chao, "Congestion-aware single link failure recovery in hybrid SDN networks," in *Proc. IEEE INFOCOM*, 2015, pp. 1086–1094.

[5] N. Li, Y. Shi, Z. Zhang, J. F. Martinez, and X. Yuan, "Search-tree based SDN candidate selection in hybrid IP/SDN network," in *Proc. IEEE ICNP*, 2020, pp. 1–6.

[6] Z. Yang and K. L. Yeung, "SDN candidate selection in hybrid IP/SDN networks for single link failure protection," *IEEE/ACM Trans. Netw.*, vol. 28, no. 1, pp. 312–321, 2020.

[7] N. Vuppalapati and T. G. Venkatesh, "Minimum candidate selection algorithm for hybrid IP/SDN networks with single link failures," *IEEE Networking Letters*, vol. 4, no. 3, pp. 152–156, 2022.

[8] F. De Rango and M. Tropea, "DTN Architecture With Resource-Aware Rate Adaptation for Multiple Bundle Transmission in InterPlanetary Networks," *IEEE Access*, vol. 10, pp. 47 219–47 234, Apr. 2022.

[9] Cisco Systems, "Best practices in core network capacity planning," Cisco Systems, White Paper C11-728551, 2013.

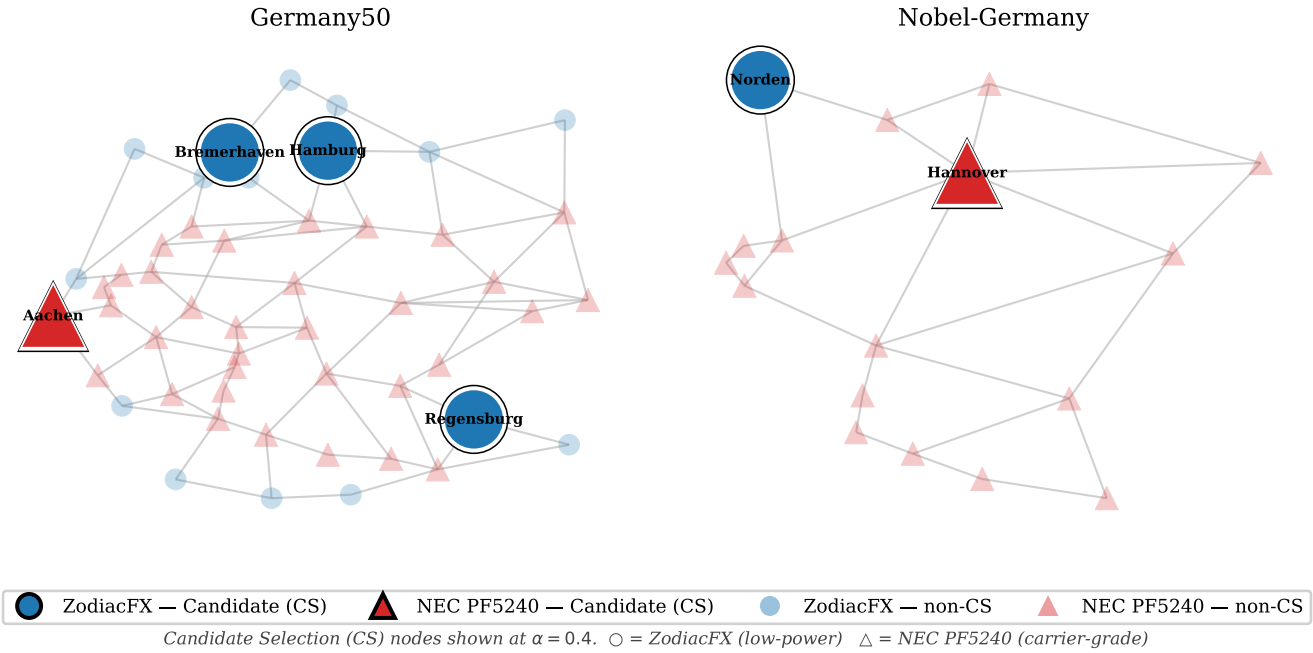


Fig. 4. Candidate node placement at $\alpha = 0.4$. Red: carrier-grade; blue: low-power tier. Bold border: designated candidate.

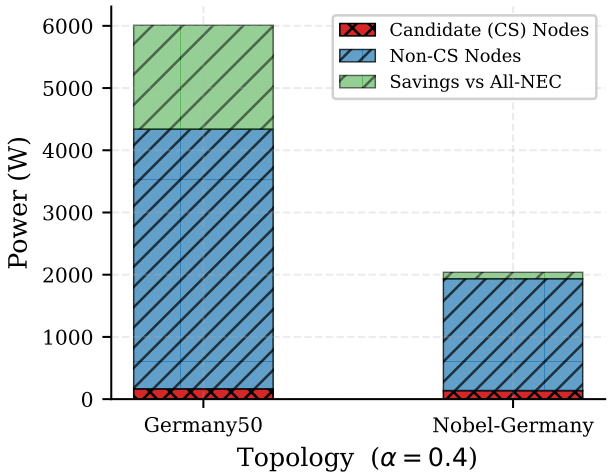


Fig. 5. Power decomposition at $\alpha = 0.4$: candidate node power, passive power, and savings vs. uniform carrier-grade. Passive nodes dominate total consumption.

[10] S. Orlowski *et al.*, "SNDlib 1.0 — survivable network design library," *Networks*, vol. 55, no. 3, pp. 276–286, 2010.

[11] Y. Zhang *et al.*, "Fast accurate computation of large-scale IP traffic matrices from link loads," in *Proc. ACM SIGMETRICS*, 2003, pp. 206–217.

[12] Northbound Networks, "ZodiacFX openflow switch datasheet," 2016, technical specification.

[13] NEC Corporation, "UNIVERGE PF5240 openflow switch specifications," 2014, technical specification document.

[14] M. G. Spina, F. De Rango, E. Scalzo, F. Guerriero, and A. Iera, "Distributing Intelligence in 6G Programmable Data Planes for Effective In-Network Intrusion Prevention," *IEEE Network*, vol. 39, no. 3, pp. 319–325, Feb. 2025.

[15] I. Maity, S. Misra, and C. Mandal, "SCOPE: Cost-efficient QoS-aware switch and controller placement in hybrid SDN," *IEEE Systems Journal*, vol. 16, no. 3, pp. 4873–4880, 2022.

[16] L. Hiryanto, S. Soh, K.-W. Chin, D.-S. Pham, and M. M. Lazarescu, "Green multistage upgrade of switches and controllers for bundled-links SDNs," *IEEE Access*, vol. 10, pp. 84 038–84 055, 2022.

[17] Y. Wang *et al.*, "Energy-efficient method based on dynamic topology switching and reliability in SDNs," *IEEE Transactions on Sustainable Computing*, vol. 7, no. 2, pp. 427–440, 2022.

[18] M. G. Spina, M. Tropea, and F. De Rango, "A Dynamic Entropy-Based Approach for DDoS Protection in Software-Defined Networks," in *2026 IEEE 23rd Consumer Communications & Networking Conference (CCNC)*. IEEE, 2026, pp. 09–12.

[19] F. Salatino, M. G. Spina, M. Tropea, and F. De Rango, "Detecting DDoS Attacks Through AI driven SDN Intrusion Detection System," in *2024 IEEE 21st Consumer Communications & Networking Conference (CCNC)*. IEEE, 2024, pp. 06–09.

[20] Z. Guo *et al.*, "Joint switch upgrade and controller deployment in hybrid software-defined networks," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 5, pp. 1012–1028, 2019.

[21] N. Vuppalapati and T. G. Venkatesh, "Candidate selection algorithms for hybrid IP/SDN networks with multi-link failures," *IEEE Trans. Netw. Serv. Manag.*, 2024.

[22] M. Caria and A. Jukan, "Link capacity planning for fault tolerant operation in hybrid SDN/OSPF networks," in *Proc. IEEE ICNP*, 2016.

[23] N. Huin *et al.*, "Bringing energy routing closer to reality with SDN hybrid networks," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 4, pp. 1128–1139, 2018.

[24] J. Galán-Jiménez, M. Polverini, and F. G. Lavacca, "Joint energy efficiency and load balancing optimization in hybrid IP/SDN networks," *Ann. Telecommun.*, vol. 78, pp. 1–14, 2023.

[25] Y. A. Mtawa, A. Haque, and H. Lutfiyya, "Migrating from legacy to software defined networks: A network reliability perspective," *IEEE Transactions on Reliability*, vol. 70, no. 4, pp. 1525–1541, 2021.

[26] B. G. Assefa and Ö. Özkasap, "RESDN: A novel metric and method for energy efficient routing in software defined networks," *IEEE Trans. Netw. Serv. Manag.*, vol. 17, no. 2, pp. 736–749, 2020.

[27] A. Morton, "IMIX genome: Specification of variable packet sizes for additional testing," IETF, Tech. Rep. RFC 6985, 2013.

[28] J. Postel, "Internet Protocol," RFC 791 (Internet Standard), Sep. 1981.

[29] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. Cambridge, MA, USA: MIT Press, 2009.