

RL-Based Adaptive Scheduling for Full-Duplex Multi-Hop FSOC under Blocking Events in Smart Semiconductor Manufacturing

Siwoong Park*
Optical ICT Convergence Research
Section
Honam Research Division
Electronics and Telecommunications
Research Institute (ETRI)
Gwangju, South Korea
swp@etri.re.kr

Ji-Soo Shin
¹⁾Optical ICT Convergence Research
Section, Honam Research Division,
Electronics and Telecommunications
Research Institute (ETRI)
²⁾College of AI Convergence, Chonnam
National University (CNU)
Gwangju, South Korea
wkdltn@etri.re.kr

Chan-Il Yeo
Optical ICT Convergence Research
Section
Honam Research Division
Electronics and Telecommunications
Research Institute (ETRI)
Gwangju, South Korea
ciyeo@etri.re.kr

Abstract—Semiconductor manufacturing increasingly relies on high-capacity and interference-free communication backbones to support automation and data-intensive processes. In semiconductor fabs, free-space optical communication (FSOC) can deliver high-throughput and interference-free connectivity. However, frequent line-of-sight interruptions from moving overhead hoist transport (OHT) vehicles and sudden burst traffic significantly threaten link stability. This paper proposes a reinforcement learning (RL)-based adaptive scheduling framework for full-duplex multi-hop FSOC networks that explicitly models and mitigates blocking events. A discrete-event simulation environment reflecting realistic fab parameters, including OHT mobility, FSOC terminal density, burst arrivals, and reconfiguration delays, was developed to evaluate performance. Results show that the RL-based approach significantly improves average throughput while reducing drop rates compared to static and heuristic baselines. These findings confirm the effectiveness of learning-driven scheduling in ensuring both capacity and reliability, providing a practical pathway toward resilient FSOC backbones in next-generation semiconductor fabs.

Keywords—Adaptive scheduling, Blocking events, Free-space optical communication (FSOC), Multi-hop networking, Overhead hoist transport (OHT), Reinforcement learning (RL), Smart semiconductor manufacturing

I. INTRODUCTION

Modern semiconductor fabs demand ultra-precise inspection systems, continuous process optimization, and large-scale data transfers across production lines, pushing communication infrastructures to their limits. As next-generation semiconductor fabs adopt highly automated and intelligent systems such as overhead hoist transport (OHT) platforms and data-intensive inspection equipment, the limitations of traditional wired and radio frequency (RF)-based infrastructures are becoming increasingly evident. Conventional RF and wired infrastructures face critical drawbacks such as electromagnetic interference (EMI) vulnerability, inflexible installation requirements, escalating costs, and limited scalability for future manufacturing demands. Designing a next-generation communication backbone that simultaneously ensures high throughput, low



Fig. 1. Concept of RL-based adaptive scheduling in a full-duplex multi-hop FSOC network under OHT blocking events in smart semiconductor manufacturing.

latency, EMI immunity, and flexible reconfiguration is therefore a critical challenge [1].

Free-space optical communication (FSOC) has gained attention as a viable alternative, offering multi-Gbps data rates, resilience against EMI, unlicensed operation, and flexible topology reconfiguration suitable for dynamic fab environments [2], [3], [4]. However, operating FSOC links in dynamic semiconductor fabs is far from straightforward. Moving OHT systems frequently obstruct line-of-sight (LOS) paths, creating blocking events that degrade link stability. Moreover, inspection and control traffic exhibits burst characteristics, resulting in sudden high bandwidth demands. Conventional static or rule-based scheduling approaches struggle to cope with such uncertainties and dynamics.

To address these limitations, this paper proposes a reinforcement learning (RL)-based adaptive scheduling framework tailored to multi-hop FSOC networks in semiconductor manufacturing environments [5], [6]. The proposed approach explicitly incorporates OHT-induced blocking events into the decision-making process of the RL agent, enabling real-time relay path selection and link reconfiguration. As a result, the framework ensures stable and high-capacity transmission even under frequent blocking conditions.

The main contribution of this paper are summarized as follows:

- FSO network modeling with explicit consideration of OHT-induced blocking events
- Design of an RL-based adaptive scheduling algorithm for multi-hop FSO
- Simulation-based evaluation demonstrating improved throughput and reduced drop rate

II. PROPOSED MODEL

A. System Overview

The proposed model is built upon a multi-hop full-duplex FSO network that explicitly considers blocking events caused by OHT vehicles in next-generation semiconductor manufacturing environments. Fig. 1 illustrates the conceptual architecture of the proposed system. FSO terminals are deployed above the production line and form a flexible multi-hop relay topology. OHT vehicles often obstruct LOS links, leading to abrupt disruptions in ongoing transmissions. To mitigate these interruptions, the RL-based scheduler predicts mobility trends and adjusts relay routes in real time while adapting to burst traffic conditions. Unlike conventional static scheduling, where blocked bursts are entirely lost, or heuristic scheduling, which relies on simple rule-based rerouting, the proposed approach enables intelligent prediction and mitigation of blocking events. This ensures stable, high-throughput, and low-latency connectivity, providing a robust and interference-free communication backbone for smart semiconductor manufacturing.

B. Simulation Parameters

Table I summarizes the simulation parameters that characterize the operating environment and evaluation setup. Each experiment is performed over a total duration T_{sim} , where the system consists of N_{FSOC} FSO terminals forming a multi-hop relay backbone with up to N_{hop} stages, and N_{OHT} overhead hoist transport vehicles moving at an average speed V_{OHT} . At every scheduling interval, the agent observes the positions and velocities of OHTs, LOS conditions between terminals, and the state of traffic queues. Traffic is modeled as bursts of size B_{burst} generated at an average frequency F_{burst} . Network control actions, including link establishment, relay assignment, and bandwidth allocation, are constrained by the maximum link throughput R_{max} and reconfiguration latency T_{link} . Service interruptions exceeding the drop threshold θ_{drop} are counted as drop events. All parameters have been carefully selected to reflect real-world semiconductor manufacturing environments, capturing OHT mobility patterns, FSO terminal density, bursty inspection/control traffic, and realistic relay reconfiguration delays. Consequently, the defined model provides a reliable abstraction for deployment-oriented evaluations of next-generation fab networks.

C. Algorithm Framework

Algorithm 1 provides a step-by-step description of the proposed RL-based adaptive scheduling scheme. At each scheduling epoch, the agent interacts with the FSO environment through the following sequence:

- The agent collects a comprehensive state vector encompassing OHT locations and speeds, current

TABLE I. SIMULATION PARAMETERS

Parameter	Symbol	Value	Description
Simulation Time	T_{sim}	30 min	Total experiment duration
OHT Nodes	N_{OHT}	20	Number of OHT
OHT Speed	V_{OHT}	100 m/min	Average speed of OHT
FSOC Terminals	N_{FSOC}	16	Number of FSO terminals
Relay Hops	N_{hop}	3	Number of multi-hop stages
Link Setup Time	T_{link}	30 ms	Reconfiguration/repair latency
Traffic Burst Size	B_{burst}	10 GB	Control/inspection burst size
Burst Frequency	F_{burst}	1-60 /min	Average inter-burst interval
Max. Link Capacity	R_{max}	10 Gbps	Single FSO link throughput
Drop Threshold	θ_{drop}	50 ms	Minimum outage to count as drop event

relay topology, LOS availability, and instantaneous traffic load.

- **State Observation:** The system state S_t captures the instantaneous dynamics of the environment, including OHT positions and velocities (V_{OHT}), the FSO relay topology up to N_{hop} stages, the LOS blocking matrix, and the current buffer occupancy q .
- **Action Selection:** Based on the scheduling policy π , the agent determines an action A_t . For the RL policy, actions are chosen via an ϵ -greedy exploration of the Q-function $Q(S_t, a)$. The heuristic policy selects relay paths through a predefined rerouting rule, while the static policy maintains fixed link assignments.
- **Action Execution:** The chosen action is enforced on the FSO network. Reconfiguration events incur a setup delay T_{link} . Whenever an OHT-induced blockage lasts longer than the predefined threshold θ_{drop} , the incident is registered as a link outage.
- **Reward Computation:** For RL, the agent evaluates a composite reward function that balances throughput, latency, drop events, and reconfiguration costs:

$$R_t = \alpha T_{net}(t) - \beta D_{avg}(t) - \gamma_{drop} L_{drop}(t) - \delta C_{config}(t),$$

where $T_{net}(t)$ is the instantaneous throughput constrained by R_{max} , $D_{avg}(t)$ denotes the average latency, $L_{drop}(t)$ indicates drop occurrences, and $C_{config}(t)$ represents the cost of reconfiguration. The weights α , β , γ_{drop} , and δ reflect the trade-offs among competing performance objectives.

- **Experience Storage and Q-Network Update:** Each transition S_t, A_t, R_t, S_{t+1} is appended to the replay buffer $\mathcal{B}$. The Q-network parameters are then updated according to the classical Q-learning rule:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \eta \left[R_t + \gamma \max_{a'} Q(S_{t+1}, a' - Q(S_t, A_t)) \right],$$

where η is the learning rate and γ is the discount factor that determines the weight of future rewards.

Algorithm 1 RL-based Adaptive Scheduling for Multi-Hop FSOC with Blocking Event

Input: T_{sim} , N_{OHT} , N_{FSOC} , V_{OHT} , N_{hop} , T_{link} , B_{burst} , F_{burst} , R_{max} , θ_{drop} , Policy $\pi \in \text{static, heuristic, RL}$, Buffer size L , RL hyperparameters, Reward Weights, Blocking model constant: $k > 0$ (Poisson-derived)

Output: Scheduling actions A_t , Updated Q-network parameters, Performance metrics: average throughput $\bar{T}$, drop rate $\bar{L}_{drop}$

```

1: Initialization:
2:   Initialize Q-network, replay buffer  $\mathcal{B} \leftarrow []$ 
3:   Set  $t \leftarrow 0$ ; delivered  $\leftarrow 0$ ; demanded  $\leftarrow 0$ ;
   drops  $\leftarrow 0$ ;  $q \leftarrow 0$  // buffer occupancy  $q$ 
4: while  $t < T_{sim}$  do
5:   Burst arrival:
6:      $t \leftarrow t + \Delta t$ ,  $\Delta t \sim \text{Exp}(\lambda)$ ,  $\lambda = F_{burst}/60$ 
7:     demanded  $\leftarrow$  demanded +  $B_{burst}$ ;
     bursts  $\leftarrow$  bursts + 1
8:   Buffer admission:
9:     if  $q + B_{burst} \leq L$  then  $q \leftarrow q + B_{burst}$ 
10:    else drops  $\leftarrow$  drops + 1
11:  end if
12:  Compute blocking probability:
13:     $P_{block} = 1 - \exp(-kF_{burst})$ 
14:    blocked  $\sim \text{Bernoulli}(P_{block})$ 
15:  Observe state:
16:     $S_t = (\text{blocked}, \text{relay depth}, \text{latency}, q)$ 
17:  Select action:
18:    if  $\pi = \text{RL}$  then
19:      Choose  $A_t$  using  $\epsilon$ -greedy over  $Q(S_t, a)$ 
20:    else if  $\pi = \text{heuristic}$  then
21:      Choose reroute path based on heuristic rule
22:    else if  $\pi = \text{static}$  then
23:      Use fixed link assignment
24:    end if
25:  Transmission reaction:
26:    if blocked = 1 then
27:      if reroute attempt fails ( $P_{fail}$ ) then
28:        drops  $\leftarrow$  drops + 1
29:      else
30:        Transmit  $\min(q, R_{max} \cdot \Delta t)$ ; update  $q$ 
31:      end if
32:    else
33:      Transmit  $\min(q, R_{max} \cdot \Delta t)$ ; update  $q$ 
34:    end if
35:  Reward calculation (if RL):
36:     $R_t = \alpha T_{net}(t) - \beta D_{avg}(t) - \gamma_{drop} L_{drop}(t) - \delta C_{config}(t)$ 
37:    where:
38:     $T_{net}(t)$ : Actual network throughput ( $\leq R_{max}$ )
39:     $L_{drop}(t)$ : Incremented if link outage  $> \theta_{drop}$ 
40:     $D_{avg}(t)$ : Average latency
41:     $C_{config}(t)$ : Reconfiguration cost ( $T_{link}$ )
42:  Q-network update (if RL):
43:    Store  $(S_t, A_t, R_t, S_{t+1})$  in buffer  $\mathcal{B}$ 
44:    Perform Q-learning update
45: end while
46: Compute results:
47:    $\bar{T} = \text{delivered}/T_{sim}$ 
48:    $\bar{L}_{drop} = \text{drops}/\text{bursts}$ 

```

By embedding the system parameters of Table I (e.g., N_{OHT} , N_{FSOC} , V_{OHT} , B_{burst} , F_{burst} , R_{max} , T_{link} , θ_{drop}) directly into the state, action, and reward design, the framework enables robust and reproducible optimization. This integration ensures that the RL agent progressively learns to mitigate blocking events and buffer overflows, thereby achieving superior throughput and reduced drop rates compared to static or heuristic scheduling baselines.

III. PERFORMANCE EVALUATION

A. Simulation Setup

To assess the effectiveness of the proposed RL-based adaptive scheduling framework, a discrete-event simulation environment was developed based on the parameters summarized in Table I. The total simulation time was set to 30 minutes (T_{sim}), during which 20 OHT vehicles (N_{OHT}) moved along the production line at an average speed of 100 m/min (V_{OHT}). Each OHT acted as a dynamic blocking obstacle for FSOC links, introducing stochastic LOS blocking events that directly affect communication reliability. A total of 16 FSOC terminals (N_{FSOC}) were deployed across the fab floor, enabling up to three relay stages (N_{hop}) to establish multi-hop connectivity when direct links were unavailable. The maximum throughput of a single FSOC link was constrained to 10 Gbps (R_{max}), while link setup and reconfiguration required an average latency of 30 ms (T_{link}). Traffic was modeled as burst arrivals of 10 GB (B_{burst}) with a frequency ranging from 1 to 60 events per minute (F_{burst}). Any service interruption exceeding 50 ms (θ_{drop}) was counted as a drop event. To ensure statistical robustness, multiple simulation runs were conducted with independent random seeds. Performance was evaluated in terms of average throughput and drop rate under three scheduling policies: static, heuristic, and RL-based.

B. Throughput Performance

The throughput performance shown in Fig. 2 was obtained through a discrete-event simulation framework. For each burst frequency ranging from 1 to 60 events per minute, multiple

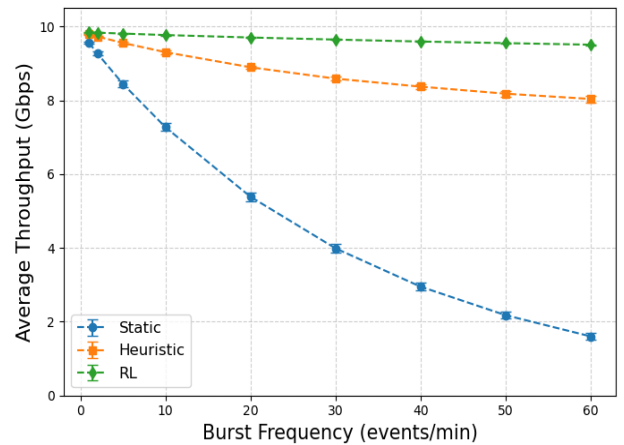


Fig. 2. Average throughput performance under varying burst frequencies for static, heuristic, and RL-based scheduling policies.

independent simulation runs were conducted to capture statistical variability. Three scheduling policies were implemented: a static baseline, a heuristic rerouting scheme, and the proposed RL-based adaptive scheduling. In the static scheme, bursts were either successfully transmitted or dropped depending on blocking events without any adaptive mechanism. The heuristic approach attempted rerouting with a fixed probability of success, reflecting limited adaptability. In contrast, the RL-based agent employed a Q-learning framework, selecting actions based on an ϵ -greedy exploration strategy, and continuously updating its Q-network through temporal-difference learning. Throughput in each run was computed as the ratio of successfully delivered burst traffic to the total offered load, normalized with respect to the maximum FSOC link capacity, which was fixed at R_{max} of 10 Gbps. The final performance curves illustrate the average throughput across 30 simulation runs, and standard deviation error bars are included to demonstrate statistical confidence.

C. Drop Rate Analysis

Fig. 3 illustrates the drop rate performance of the three scheduling policies under varying burst frequencies. The static scheme exhibits a rapid increase in drop ratio, exceeding 80% when the burst frequency reaches 60 events per minute. This sharp degradation arises from its inability to adapt to frequent blocking events, leading to buffer overflows and persistent link outages. The heuristic scheme demonstrates improved resilience by reducing the drop rate significantly compared to the static baseline, yet the ratio still climbs steadily with higher traffic intensity, reflecting its limited adaptability under dynamic fab conditions. In contrast, the RL-based policy maintains a consistently low drop rate, remaining close to zero even under the most demanding traffic conditions. This outcome highlights the ability of reinforcement learning to anticipate and mitigate blocking-induced disruptions by leveraging its learned rerouting strategies. Overall, these results confirm that the RL-based scheduling framework not only sustains high throughput (Fig. 2) but also ensures reliability by effectively suppressing drop events across a wide range of operating conditions.

D. Discussion

The performance trends observed in Figs. 2 and 3 provide several key insights. First, the static scheduling scheme suffers a severe throughput collapse as burst frequency increases, confirming its inability to cope with frequent blocking events in dynamic fab environments. The heuristic approach mitigates this degradation to some extent by attempting rerouting, but its reliance on a fixed success probability results in limited performance gains under highly variable traffic and mobility conditions. In contrast, the RL-based policy consistently sustains near-optimal throughput while maintaining lower drop rates, demonstrating its ability to learn and exploit environment-specific blocking patterns. These findings highlight the critical role of adaptive, learning-driven strategies in ensuring reliable FSOC backbone performance for next-generation semiconductor manufacturing. At the same time, the computational overhead and convergence time of RL remain important factors to be addressed in real-time deployment, suggesting directions for lightweight or hardware-assisted implementations. It should be noted that average throughput and drop rate, although displaying inverse trends in the figures, represent complementary aspects of network performance rather than interchangeable measures. Throughput characterizes the effective data delivery capacity

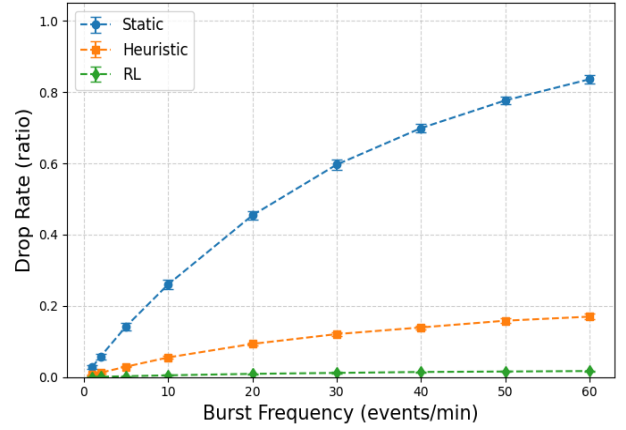


Fig. 3. Drop rate performance under varying burst frequencies for static, heuristic, and RL-based scheduling policies.

of the FSOC backbone, incorporating the impact of successful transmissions, latency, reconfiguration delays, and bandwidth utilization. By contrast, drop rate quantifies the fraction of burst traffic lost due to blocking or buffer overflow, serving as a direct indicator of link reliability. While an increase in drop rate inevitably reduces throughput, the two metrics emphasize different dimensions of performance. Evaluating them jointly provides a more comprehensive assessment, demonstrating that the proposed RL-based scheduling framework can simultaneously sustain high capacity and ensure robust reliability under dynamic semiconductor fab conditions.

IV. CONCLUSION

This study introduces a blocking-aware RL scheduling strategy for full-duplex multi-hop FSOC systems, enhancing throughput and reliability in highly dynamic semiconductor manufacturing settings. Unlike conventional scheduling schemes, the proposed framework explicitly models and mitigates blocking events caused by OHT movements, enabling more resilient and reliable link maintenance in highly dynamic fab conditions. Through discrete-event simulations, the framework is shown to significantly improve average throughput while simultaneously reducing drop rates compared to static and heuristic baselines. These results confirm the effectiveness of reinforcement learning in adapting to blocking events and traffic variability, thereby ensuring both high capacity and robust reliability in FSOC backbones. Future work will focus on reducing the training overhead and convergence time of RL algorithms, as well as developing lightweight and hardware-assisted implementations to ensure practical real-time deployment in next-generation fabs.

ACKNOWLEDGMENT

This research was supported by Ministry of Trade, Industry and Energy (MOTIE) and the Korea Institute for Advanced of Technology (KIAT), under the corporate demand-driven challenge and Innovative R&D Program for Next-Generation Researchers (Grant No.: RS-2025-15373143)

REFERENCES

- [1] H. Jie, Z. Zhao, Y. Zeng, Y. Chang, F. Fan, C. Wang, and K. Y. See, "A review of intentional electromagnetic interference in power

- electronics: Conducted and radiated susceptibility,” *IET Power Electron.*, vol. 17, no. 12, pp. 1487-1506, 2024.
- [2] M. Safari, and M. Uysal, “Relay-assisted free-space optical communication,” *IEEE Trans. Wirel. Commun.*, vol. 7, no. 12, pp. 5441-5449, 2008.
- [3] M. A. Khalighi and M. Uysal, “Survey on free space optical communication: A communication theory perspective,” *IEEE Commun. Surv. Tutor.*, vol. 16, no. 4, pp. 2231-2258, 2014.
- [4] S. Park, C. I. Yeo, Y. S. Heo, J. H. Ryu, H. S. Kang, S. C. Kim, and J. H. Jang, “Common path-based mobile free-space optical terminal with adaptive beamforming function for Gbps out-of-band full-duplex connectivity to UAVs,” *Opt. Commun.*, vol. 494, no. 127041, 2021.
- [5] Z. Gao, M. Eisen and A. Ribeiro, “Resource allocation via model-free deep learning in free space optical communications,” *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 920-934, 2022.
- [6] J. Liu, H. Luo, H. Tao, J. Liu, and J. Zhou, “JLOS: A cooperative UAV-based optical wireless communication with multi-agent reinforcement learning,” *IEEE Trans. Netw. Serv. Manag.*, vol. 22, no. 2, pp. 1345-1356, 2025.