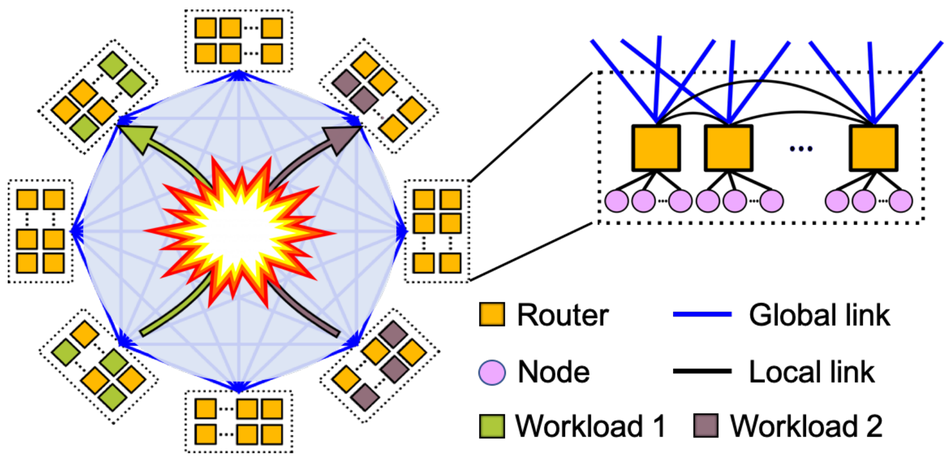
Preventing Workload Interference with Intelligent Routing and Flexible Job Placement Strategy on Dragonfly system

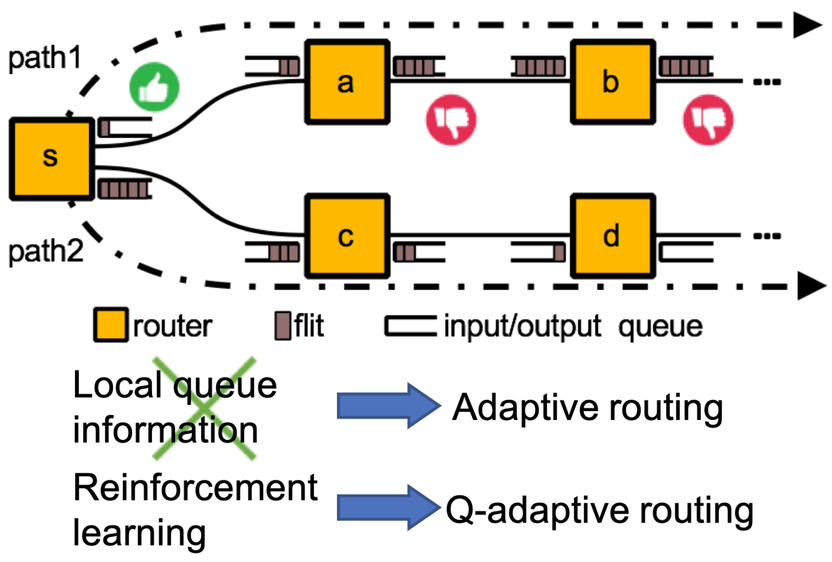
XIN WANG, YAO KANG, ZHILING LAN

Motivation

- Dragonfly networks suffer from inter-job interference due to shared global links.



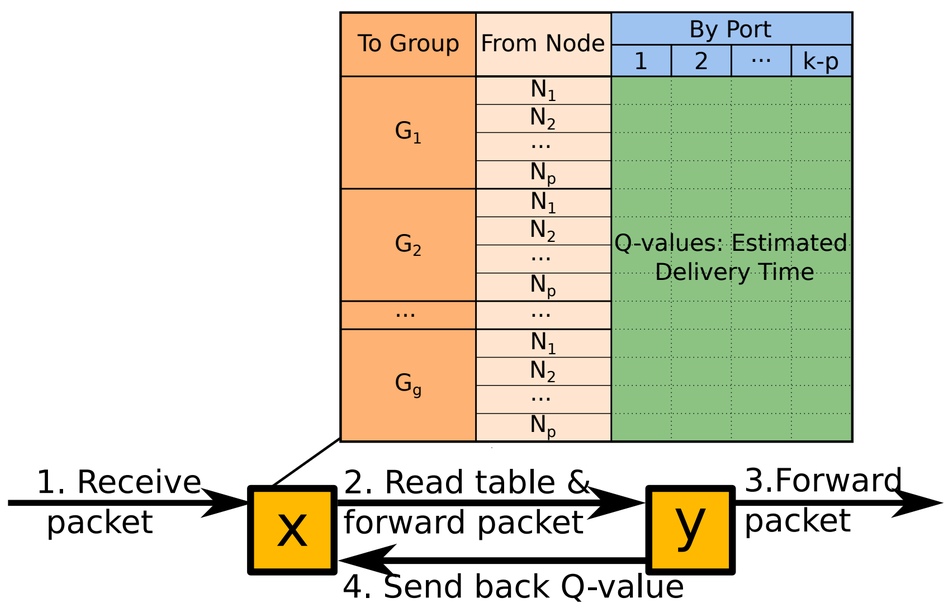
- Existing routing algorithms fail to account for system-wide job patterns and link congestion.



- Modern HPC systems offer flexible job placement, allowing users to specify workload locations for optimized performance.

Methodology

- **Q-Adaptive Routing**: a reinforcement learning based routing policy that considers global congestion.
  - Uses multi-agent reinforcement learning with two-level Q-tables.
  - Each router learns network state and selects low-delay paths.
  - Tabular Q-routing ensures low overhead and fast adaptation.
  - Fully distributed design, no extra hardware or wiring needed.
  - Updates Q-values online to reflect dynamic congestion.



- **Flexible Job Placement (FP)**: reduces inter-job network contention by placing jobs based on their communication and isolation needs.
  - Confines communication-intensive jobs to exclusive group(s) to minimize interference.
  - Isolates performance-sensitive jobs in dedicated group(s), avoiding interference from other workloads.

Experiment Setup

- System: 33-group, 1,056-node Dragonfly
- Workloads: 6 represented applications are categorized into 5 patterns seen in both scientific and machine-learning applications.

Pattern	App	Node Count	Group Count	Total Msg (MB)	Execution time (ms)	Injection Rate (GB/s)	Peak Ingress Volume
Random	UR	128	4	2853.06	13.20	216.14	3.07KB
Sweep	LU	128	4	3149.97	11.92	264.33	30.0KB
Alltoall	FFT3D	128	4	14926.35	37.35	399.67	51.68KB
Stencil	LQCD	256	8	7949.22	12.55	633.50	4.60MB
	Stencil5D	243	7.59	4536.16	9.39	483.12	14.0MB
Allreduce	CosmoFlow	128	4	572.06	12.87	44.44	2.25MB

Table 1: Summary of application configuration

Evaluation Results

- **Effectiveness of Q-adaptive routing**

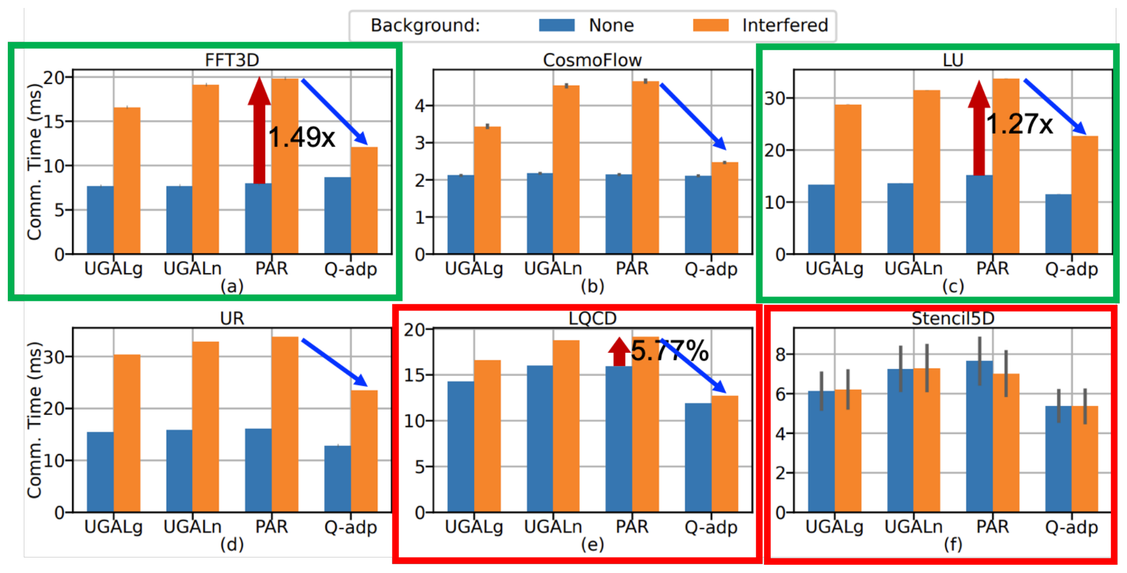


Figure 1: Communication performance of each application when run standalone and as part of a mixed workload under random placement.

- **Effectiveness of FP: Reduce system-wide interference**

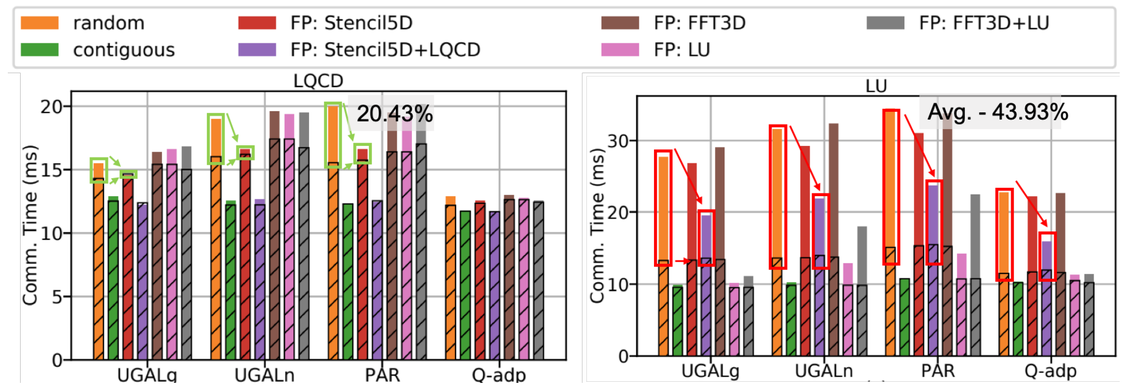


Figure 2: Communication time of LQCD and LU under different placement strategies. Colored bars show mixed workload cases; shaded bars show standalone cases. The gap reflects delay from network interference.

- **Effectiveness of FP: Protect target application**

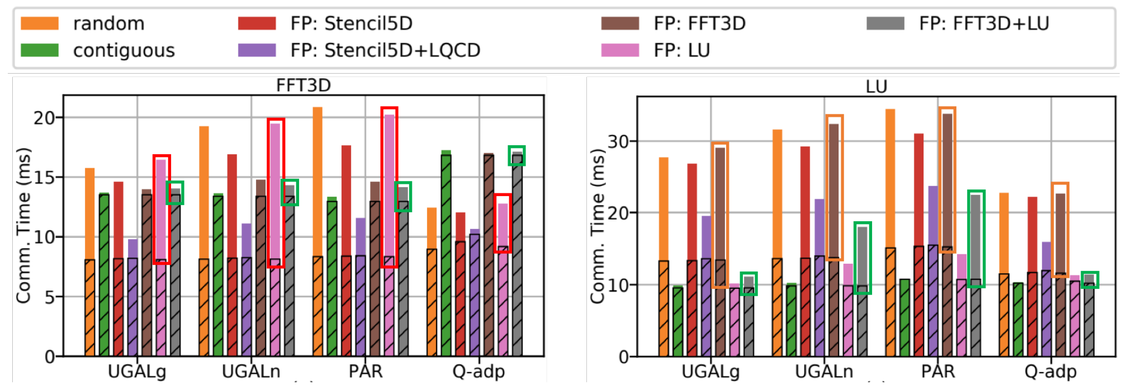


Figure 2: Communication time of FFT3D and LU under different strategies.

Summary:

- Q-Adaptive Routing outperforms traditional methods and significantly reduces communication delay.
- FP reduces interference by isolating communication-heavy jobs and assigning user jobs to dedicated groups.
- Our method is available at: https://doi.org/10.5281/zenodo.7882307

UNIVERSITY OF ILLINOIS CHICAGO