A Quantum Data Center Network Architecture

Yufeng Xin
RENCI, University of North Carolina at Chapel Hill
Chapel Hill, NC, USA

Liang Zhang
ESnet, Lawrence Berkeley National Laboratory
Berkeley, CA, USA

2024 IEEE International Conference on Quantum Computing and Engineering (QCE) | 979-8-3315-4137-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/QCE60285.2024.10387

***Abstract*—Quantum entanglement generation and distribution is key to the vision for a Quantum Internet. Existing research has focused on entanglement routing schemes and QR performance in a general network setting with strong physical layer assumptions. In reality, the required quantum memory and quantum repeater devices along with the control protocols for general quantum networks remain grand challenges in the foreseeable future. Main performance concerns at the networking layer include entanglement distribution rate and fidelity. We thereafter propose to focus more research efforts on the quantum data center network due to its structured topology and potential to support emerging distributed quantum computing applications. This paper presents a quantum data center (QDC) network architecture with novel switch architecture and protocols that feature locally buffered entanglement qubits and centralized on-demand entanglement generation and distribution. Queueing models for the switches and large-scale network simulation with the NetSquid quantum simulator demonstrate the high network throughput, scalability, and fidelity performance for the proposed QDC network architecture and protocol.**

I. Introduction

The vision for a Quantum Internet relies on the transportation and processing of quantum state information among quantum devices connected through high-quality quantum and classical channels. The main performance concerns at the networking layer are successful entanglement distribution and end-to-end fidelity. However, existing studies often rely on high-level abstractions of physical components. This project proposes a quantum data center (QDC) network architecture to bridge the gap between network design and physical realization. QDCs, with their structured topology and deterministic routing, can support scalable quantum applications.

II. Quantum Data Center Network

Linking multiple quantum processors faces challenges in local processing and network communication, measured by entanglement generation rate, quality, and qubit utilization. Quantum networking typically uses abstract channels between nodes for heralded entanglement generation and ebit distribution. Quantum switches or repeaters perform entanglement swapping to extend reach, but performance is hindered by low swapping success ($\sim$ 50%) and exponential fidelity degradation over distance and time, complicating multi-hop entanglement paths. Practical issues include the cost of heralding stations, protocol complexity, and inflexible entanglement generation. The abstraction of quantum links often requires discarding unused ebits, wasting valuable resources. These factors limit the capacity and performance of current quantum network architectures, making them unsuitable for meaningful quantum applications.

Our QDC design benefits from smaller scale, less stringent physical constraints, structured classical data center designs, and a centralized network control framework. Our first innovation moves the ebit generation function from quantum links to quantum switches, enhancing network performance with continuous ebit generation and storage. Centralized nodes for entanglement distribution and heralding improve performance and cost efficiency [1]. Recent prototypes of quantum repeater nodes support the feasibility of this design [2]. Our second innovation separates the ebit swapping function from generation, eliminating storage needs at hosts and resulting in a scalable network architecture with improved throughput.

The QDC architecture uses a spine-leaf fat-tree topology, supporting high capacity and speed. It has three node layers: spine switches, leaf switches, and hosts. Leaf switches facilitate intra-cluster entanglement, while spine switches manage cross-cluster entanglement. Quantum leaf switches generate and store entanglement locally, providing on-demand distribution and routing. Spine switches handle swapping for cross-cluster demands, while hosts focus on quantum computing tasks, dynamically requesting ebits as needed.

III. Quantum Spine Switch and Quantum Leaf Switch

A. Quantum Leaf Switch

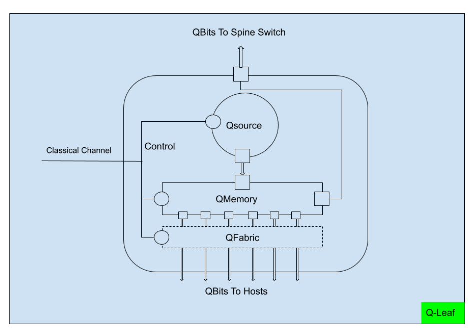

Fig. 1: Quantum Leaf Switch Architecture

The quantum leaf switch features a centralized ebit source, a qubit memory, and a qubit fabric with ports to the hosts. This new architecture greatly simplifies the operational structure of the switch and boost the switch performance in terms of throughput and buffer size while offering more control.

Our proposed leaf switch can be abstracted as a capacity $M/M/C/K$ queue with reneging. When considering the dephasing of the qubits in the buffer, an ebit in the buffer will expire after its fidelity drops below a certain threshold and will leave the queue, *i.e.*reneging. The buffer size is K, the number of ebits that can be stored in the memory. C represents the number of servers in the queue. Contrary to the common network queue model, in this quantum switch queue model, the customer arrival process is the ebit generation process while the connection requests assume the role of servers in the queue.

B. Quantum Spine Switch

The spine switch is designed as an *ebit* swapping station to facilitate the entanglement connection between two hosts under different leaf switches. Upon such a connection request, the two corresponding leaf switches will each send a *qubit* out of an *ebit* to a spine switch for a swapping operation while sending the other *qubit* to the hosts respectively.

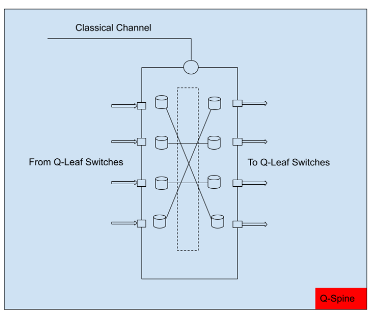

Fig. 2: Quantum Spine Switch Architecture

We model the spine-leaf network as a product assembly line system, where the swapping operation is modeled as an assemble service that needs two *ebits*, the parts, available. We can also add an assembly success probability in the model, in which a swapping failure will automatically lead to using the next pair of available *ebits*.

IV. Experiments and Evaluation

We evaluate the performance of the proposed QDC architecture using a queuing theoretical model-based simulation implemented with the Simpy event simulator. We show the performance of a QDC leaf switch with the M/M/c/K with reneging model in Fig. 3 and results on a basic QDC network consisting of 1 spine, 2 leaf, and 3 hosts per leaf switch with the assembly line queue model in Fig. 4.

V. Conclusions

The presented quantum data center network architecture provides a high-performance solution to bridge the gap between high-level network layer designs and the physical realization of quantum communication systems. Via queue models and simulation, we show that the proposed quantum leaf and spine switch design can efficiently generate and distribute entangled qubits and ensure high throughput and fidelity. The

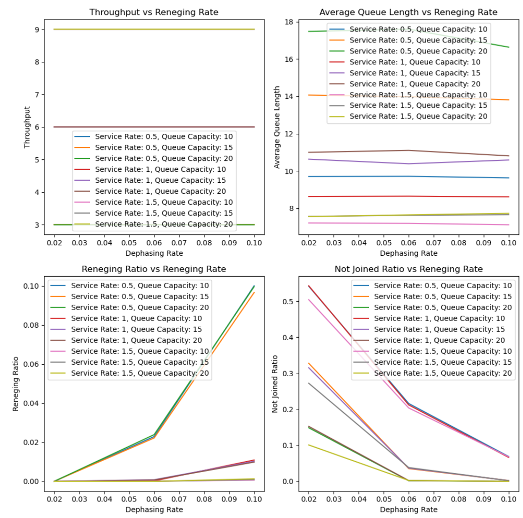

Fig. 3: Quantum Leaf Switch Performance (ebit generation rate = 10)

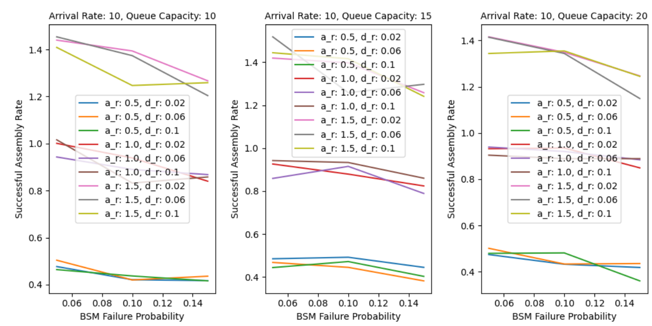

Fig. 4: Quantum Spine Switch Performance (Leaf ebit generation rate = 10)

integration of locally buffered entanglement links and on-demand entanglement distribution further optimizes performance. For the future work, we have developed a high-fidelity QDC network simulation framework using the NetSquid quantum simulator to evaluate large-scale QDC networks with realistic quantum settings. We will further refine the physical design of the quantum switches and develop advanced network control functions to support emerging quantum computing applications.

VI. Acknowledgments

We would like to thank RENCI, ESnet, and especially the QUANT-NET project for their support and contributions.

References Cited

[1] N. K. Panigrahy, T. Vasantam, D. Towsley, and L. Tassiulas, “On the capacity region of a quantum switch with entanglement purification,” in *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE, 2023, pp. 1–10.

[2] V. Krutyanskiy, M. Canteri, M. Meraner, J. Bate, V. Krcmarsky, J. Schupp, N. Sangouard, and B. P. Lanyon, “Telecom-wavelength quantum repeater node based on a trapped-ion processor,” *Physical Review Letters*, vol. 130, no. 21, p. 213601, 2023.