NGHIÊN CỨU HIỆU NĂNG KIẾN TRÚC CHUYỂN MẠCH GÓI QUANG CHO MẠNG TRUNG TÂM DỮ LIỆU

STUDY OF THE PERFORMANCE OF AN OPTICAL PACKET SWITCH ARCHITECTURE FOR DATA CENTER NETWORKS

Cao Hồng Sơn

TÓM TẮT

Bài báo phân tích đánh giá hiệu năng kiến trúc chuyển mạch gói quang cho các mạng trung tâm dữ liệu (DCN). Các kết quả tính toán và mô phỏng cho thấy hiệu năng của kiến trúc chuyển mạch Spanke là tốt hơn so với kiến trúc chuyển mạch khác. Trên cơ sở đó trong bài báo đã đề xuất một mô hình kiến trúc chuyển mạch gói toàn quang mới cho mạng trung tâm dữ liệu. Trong kiến trúc chuyển mạch gói toàn quang này đã sử dụng kiến trúc chuyển mạch spanke cải tiến kết hợp kỹ thuật ghép phân chia theo bước sóng (WDM) và sử dụng khối xử lý mào đầu gói quang (OHP). Với kiến trúc WDM OPS kết hợp OHP đã giúp làm giảm đáng kể trễ hệ thống và tăng được thông lượng chuyển mạch trong các mạng trung tâm dữ liệu lớn.

Từ khóa: *Mạng trung tâm dữ liệu, chuyển mạch gói quang (OPS), kiến trúc chuyển mạch gói quang.*

ABSTRACT

The paper analyzes and evaluates the performance of optical packet switching architecture for data center networks (DCNs). The calculation and simulation results show that Spanke switching architecture's performance is better than other switching architectures. Based on this, the paper proposes a new all-optical packet switching architecture for the data center network. In this all-optical packet switching architecture, an improved spanke switching architecture is used, which combines wavelength division multiplexing (WDM) and the Optical Header Processor (OHP). With WDM OPS architecture combined with OHP, it has significantly reduced system latency and increased switching throughput in large data center networks.

Keywords: *Datacenter network, optical packet switching (OPS), optical packet switch architecture.*

Khoa Viễn thông 1, Học viện Công nghệ Bưu chính Viễn thông
Email: sonch@ptit.edu.vn
Ngày nhận bài: 05/11/2019
Ngày nhận bài sửa sau phản biện: 13/12/2019
Ngày chấp nhận đăng: 20/02/2020

1. GIỚI THIỆU

Xu hướng gần đây cho thấy các ứng dụng mạng chuyển từ các trung tâm dữ liệu đám mây riêng sang các trung tâm dữ liệu đám mây công cộng. Như được chỉ ra trong chỉ số đám mây toàn cầu của Cisco năm 2016, số lượng IP đám mây toàn cầu hàng năm sẽ đạt 14,1ZB vào cuối năm 2020, tăng từ 8,6ZB vào năm 2018 [1, 2]. Trong các mạng trung tâm dữ liệu hiện nay, các thiết bị chuyển mạch điện tử chưa đáp ứng được các nhu cầu về băng thông nhiều hơn, mức tiêu thụ điện năng thấp và độ trễ nhỏ [3]. Các thiết bị chuyển mạch kênh quang (OCS) có thể tăng băng thông đáng kể nhưng không phù hợp với lưu lượng thay đổi. Các DCN được đề xuất mới nhất sử dụng công nghệ chuyển mạch gói quang (OPS) có khả năng đạt được thông lượng cao, mềm dẻo phù hợp với các dịch vụ có tính bùng nổ. Đặc biệt, giải quyết được hiện tượng thắt nút cổ chai, gia tăng được thông lượng chuyển mạch. Tuy nhiên trong các công trình này đã đưa ra các mô hình kiến trúc chuyển mạch gói quang sử dụng các bộ xử lý mào đầu điện tử, do đó làm tăng thời gian xử lý gói tại các nút chuyển mạch.

Trong bài báo này, tác giả đề xuất một mô hình kiến trúc chuyển mạch gói toàn quang sử dụng kỹ thuật xử lý mào đầu gói quang (OHP) có khả năng làm giảm thời gian xử lý mào đầu hiệu quả hơn so với kỹ thuật xử lý mào đầu điện tử, do đó làm giảm đáng kể thời gian trễ truyền các gói dịch vụ đồng thời làm tăng thông lượng chuyển mạch. Thông qua mô hình giải tích và mô phỏng đã tiến hành khảo sát và đưa ra các kết quả liên quan tới các tham số hiệu năng thời gian trễ hệ thống trung bình và thông lượng chuyển mạch.

Bài báo được cấu trúc như sau: Phần một là phần giới thiệu; Tiếp theo kiến trúc chuyển mạch gói quang, phân tích hiệu năng chuyển mạch gói quang, các kết quả tính toán số và mô phỏng hiệu năng chuyển mạch gói quang được khảo sát và thảo luận cũng như kiến trúc chuyển mạch gói toàn quang đề xuất được trình bày trong phần thứ hai; Cuối cùng, phần thứ ba là phần kết luận.

2. KIẾN TRÚC CHUYỂN MẠCH GÓI TOÀN QUANG

2.1. Kiến trúc chuyển mạch gói quang

Các kiến trúc OPS có khả năng kết nối hàng nghìn cổng vào - ra trong khi có thể điều khiển theo thang thời gian nano giây. Trong các DCN thường sử dụng sơ đồ hệ thống như trên hình 1 [4].

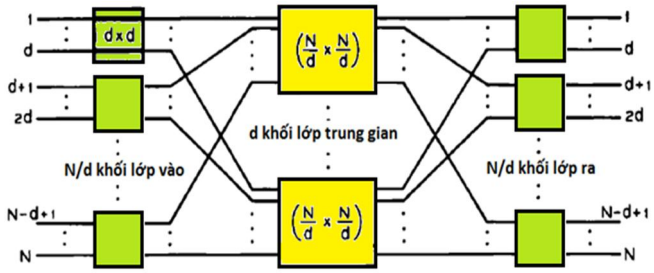
Trong đó, thông tin trao đổi nội bộ và nội cụm vẫn được điều khiển bởi các chuyển mạch điện tử, trong khi giao tiếp giữa các cụm được xử lý bởi một OPS. OPS thực hiện kết nối các cụm với nhau bằng các cổng vào/ra tốc độ bít cao.

Chuyển mạch hoạt động hoàn toàn trong miền quang. Sau đây sẽ trình bày hai kiến trúc chuyển mạch gói quang tiêu biểu là chuyển mạch điều khiển tập trung (kiến trúc Benes) và chuyển mạch điều khiển phân tán (kiến trúc Spanke).

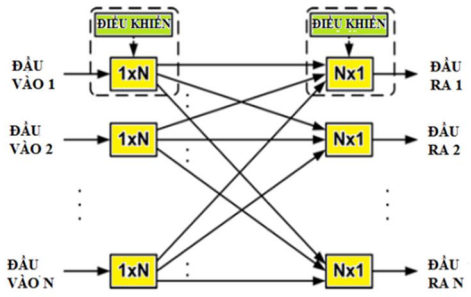
Hình 1. Sơ đồ hệ thống DCN sử dụng OPS

Kiến trúc Benes là một mạng chuyển mạch có khả năng tái cấu hình N đầu vào và N đầu ra, ký hiệu là $(N \times N)$. Cho $N = d.q$, với d và q là các ước nguyên của N. Một mạng $(N \times N)$ có thể được biểu diễn thành một lớp đầu vào và một lớp đầu ra với tổng cộng $2N/d$ mạng $(d \times d)$ và một lớp giữa gồm d mạng $(N/d \times N/d)$ như hình 2. Mạng này được gọi là một cấu trúc base-d, những mạng nhỏ hơn được gọi là các mạng con [5].

Đối với kiến trúc của Benes, khi N tăng sẽ làm tăng số lượng các khối (2×2) trung gian ở giữa qua đó làm tăng số bit điều khiển cho mỗi gói tin, làm tăng thời gian điều khiển cấu hình ma trận chuyển mạch. Thời gian điều khiển của hệ thống này phụ thuộc vào $N.\log_2 N$ khi sử dụng giải thuật điều khiển looping.



Hình 2. Kiến trúc chuyển mạch Benes (N X N) base-d



Hình 3. Kiến trúc chuyển mạch Spanke (N x N)

Kiến trúc Spanke là mạng chuyển mạch không nghẽn hoàn toàn với N cổng đầu vào/ đầu ra như hình 3. Trong đó, các thành phần chính là các chuyển mạch không gian $(1 \times N)$ và $(N \times 1)$ [5]. Kiến trúc này có cấu trúc mô-đun: số lượng chuyển mạch $(1 \times N)$ và $(N \times 1)$ tăng tuyến tính với số lượng cổng và mỗi chuyển mạch này có thể được điều khiển tự động. Do đó, độ phức tạp điều khiển và thời gian cấu hình của toàn bộ chuyển mạch đa cổng có thể giảm đáng kể. Trong trường hợp này, tổng thời gian để xác định ma trận chuyển đổi là thời gian để tạo ra một chuyển mạch $(1 \times N)$ và một chuyển mạch $(N \times 1)$ duy nhất. Trong thực tế, các mào đầu gói tại mỗi đầu vào N có thể được xử lý cục bộ bởi các bộ điều khiển độc lập.

2.2. Phân tích hiệu năng

2.2.1. Thông lượng, Throughput

Thông lượng là tỉ lệ băng thông hiệu dụng trên tổng băng thông. Thông lượng của hệ thống bằng 1 tại mỗi khe thời gian có ít nhất một gói được gửi tới đầu ra của hệ thống. Tuy nhiên thông lượng sẽ bị giới hạn bởi thời gian cấu hình t_{Switch} do không có gói nào được gửi trong khoảng thời gian này. Công thức tổng quát (1) thể hiện xác suất có ít nhất một gói được gửi tới đầu ra tính toán dựa trên thời gian trễ do cấu hình ma trận chuyển mạch [5].

$$\text{Throughput} = \frac{RTT}{RTT+t_{Switch}} \sum_{k=1}^{N} C_N^k \left(\frac{load_R}{N}\right)^k \left(1-\frac{load_R}{N}\right)^{N-k} \quad (1)$$

trong đó, RTT là thời gian khứ hồi, t_{Switch} tương ứng là thời gian cấu hình chuyển mạch, $load_R$ là tải thực, N là số cổng vào/ cổng ra.

2.2.2. Trễ hệ thống trung bình, latencyUB

Việc tính toán trễ sẽ phức tạp hơn do mô hình phân tích không tính toán đến dung lượng bộ đệm, trong khi đây là một yếu tố tác động rất nhiều đến trễ hệ thống. Tuy nhiên, dựa trên một mô hình đơn giản để có thể tính toán tổng quát và so sánh trễ khi sử dụng các kiến trúc chuyển mạch khác nhau cho cho OPS. Sử dụng xác suất truyền lại tối đa $(P_{RetrMax})$ và đặt $load = 1$ để tính toán đường bao trên của trễ hệ thống trung bình (latencyUB). Vậy, latencyUB hiệu dụng có thể được tính toán như sau [5]:

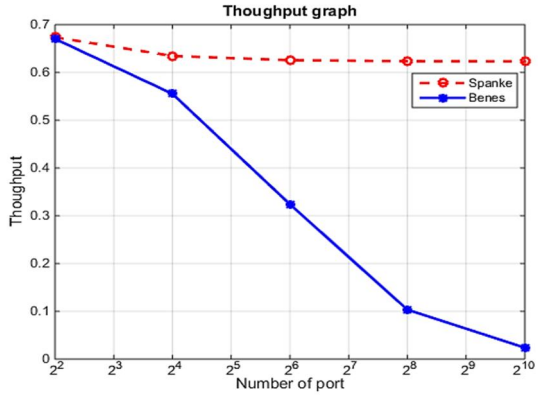
$$\text{Latency UB} = \frac{RTT+ t_{switch}}{1-P_{RetrMax}} \quad (2)$$

2.3. Kết quả đánh giá hiệu năng và thảo luận

Trong phần này sẽ so sánh các tham số hiệu năng của chuyển mạch Benes và chuyển mạch Spanke. Khi khảo sát lấy khoảng cách host-switch là 40m. RTT sẽ là 400ns. Chu kỳ gói là 40ns. Thời gian cấu hình chuyển mạch trung bình là 6ns cho kiến trúc chuyển mạch Spanke. Ở mỗi cổng vào, xác suất có dữ liệu đến (load hay tải) theo phân phối nhị thức sẽ được đặt cố định. Tất nhiên sẽ có xác suất xảy ra tranh chấp và coi như các đầu ra là như nhau.

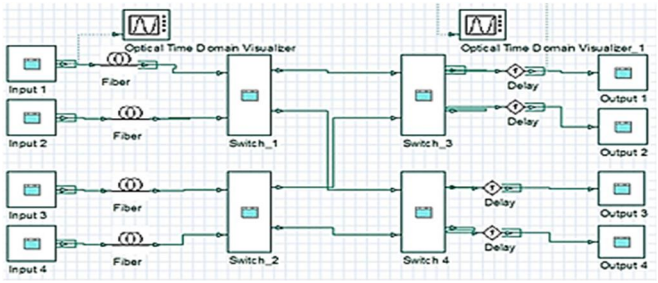
Đầu tiên sẽ phân tích các kết quả tính toán số cho thông lượng trung bình hệ thống dựa trên biểu thức (1) [5]. Kết quả đưa ra trên hình 4 biểu thị thông lượng trung bình hệ thống thay đổi theo số cổng đầu vào trong hai kiến trúc chuyển mạch: Spanke, Benes với $load_R = 1$. Có thể thấy thông lượng trong trường hợp sử dụng kiến trúc Benes phụ thuộc rất nhiều vào số lượng cổng và giảm rất nhanh khi kích thước ma trận chuyển mạch tăng. Còn thông lượng khi sử dụng kiến trúc Spanke hầu như giữ nguyên không đổi. Thông lượng hệ thống hoàn toàn độc lập so với số cổng và tăng tuyến tính theo tải cho tới khi giá trị tải đạt mốc 0,7.

Tuy có thể thấy một sự giảm nhẹ đối với những ma trận chuyển mạch lớn, nhưng hiện tượng này có thể giải thích bằng việc nghiễm nhiên xác suất tranh chấp sẽ tăng lên khi số lượng cổng tăng, dẫn đến thông lượng giảm chứ không phải do ảnh hưởng từ kiến trúc.

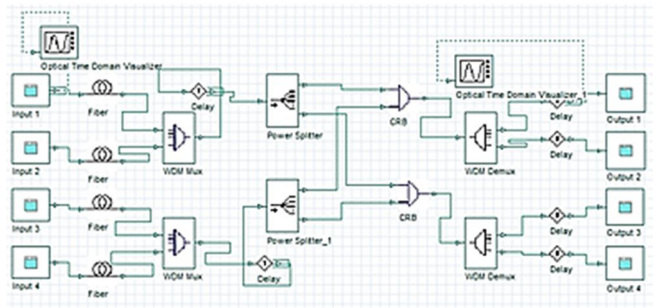


Hình 4. Thông lượng trung bình hệ thống thay đổi theo số cổng đầu vào trong hai kiến trúc: Spanke, Benes

Tiếp theo, để so sánh trễ trong mô hình kiến trúc chuyển mạch Benes và Spanke, thực hiện tiến hành thiết lập một mô hình mô phỏng với ba kịch bản tương ứng với các kiến trúc chuyển mạch có số cổng đầu vào/đầu ra lần lượt là $N = 4, 8$ và 16 cổng trong gói phần mềm OptiSystem. Trên hình 5a và 5b tương ứng là mô hình thiết lập cho kiến trúc chuyển mạch Benes và Spanke 4 đầu vào ra.



(a)

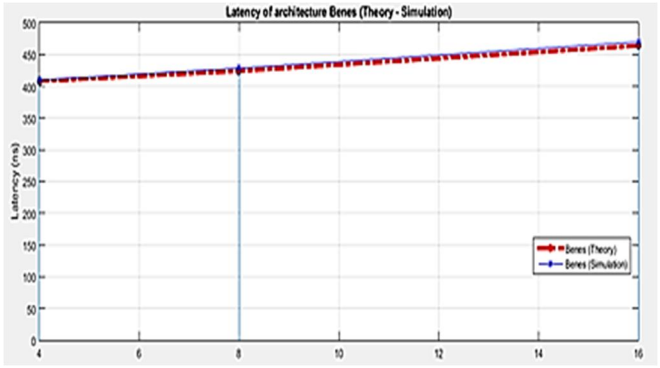


(b)

Hình 5. Mô hình thiết lập cho chuyển mạch 4 cổng vào/ cổng ra: a) Benes và b) Spanke trong phần mềm OptiSystem

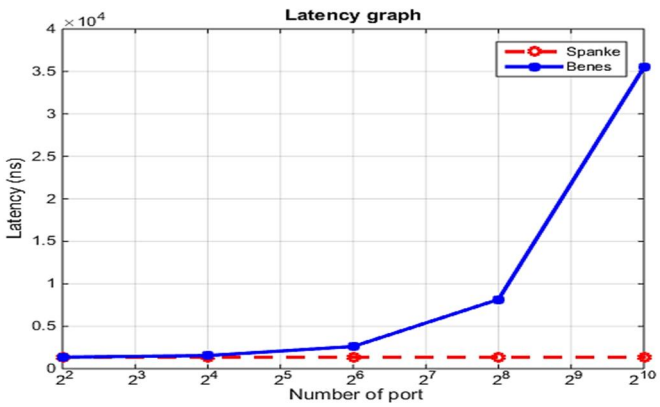
Trong mô hình mô phỏng, các gói quang được phát ở tốc độ 100 Gbit/s trên các cổng vào. Công suất phát quang trung bình của gói là 1mW.

Hình 6 và 7 là kết quả biểu thị trễ phụ thuộc vào số cổng N của kiến trúc chuyển mạch Benes và Spanke theo lý thuyết [5] và theo mô phỏng. Có thể quan sát rõ ràng trễ của hệ thống sử dụng các kiến trúc chuyển mạch tính toán theo lý thuyết khá sát với mô phỏng trên mô hình hệ thống bằng phần mềm OptiSystem.



Hình 6. Trễ của kiến trúc Benes theo lý thuyết và theo mô phỏng

Hình 7. Trễ của kiến trúc Spanke theo lý thuyết và theo mô phỏng



Hình 8. Trễ của hệ thống sử dụng kiến trúc chuyển mạch Spanke và Benes

Hình 8 đưa ra trễ của hệ thống sử dụng kiến trúc chuyển Spanke và kiến trúc chuyển mạch Benes thay đổi theo số cổng vào/ cổng ra dựa trên biểu thức (2) [5]. Trên hình 8 có thể thấy rõ rằng hiệu năng của hệ thống sử dụng kiến trúc chuyển mạch Benes giảm một cách đáng kể khi kích thước ma trận chuyển mạch tăng lên. Với việc trễ tối đa cho phép của nhiều dịch vụ trong môi trường DC là 1µs, kiến trúc chuyển mạch Benes sử dụng giải thuật loop không thể đáp ứng được yêu cầu khi số lượng cổng $N > 16$. Trong khi kiến trúc chuyển mạch Spanke luôn duy trì trễ hệ thống ở mức dưới 1µs dù lượng cổng sử dụng là bao nhiêu và tương tự như tham số thông lượng, không có sự khác biệt nhiều về trễ khi lượng cổng tăng lớn hơn 64. Cả hai thông số trên

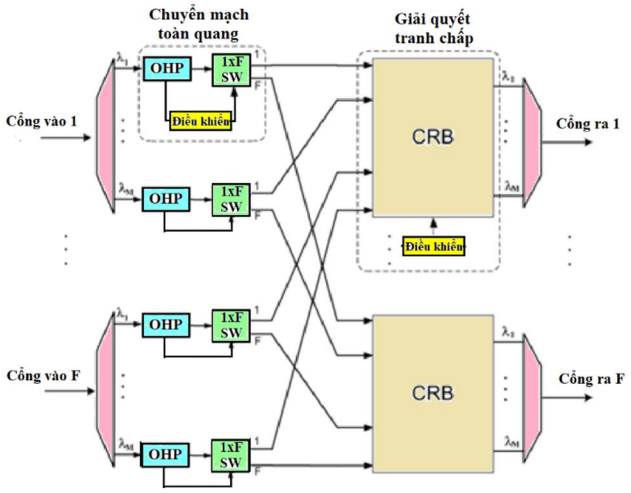
đều trở nên bão hòa khi tải tăng lớn hơn 0,7. Ngoài ra, việc tính toán trên đây không xét đến dung lượng bộ đệm, một nhân tố ảnh hưởng rất nhiều đến trễ khi có tải lớn. Thông số này tăng chậm theo số lượng cổng do tính tất yếu của việc tăng xác suất cần truyền lại ở số lượng cổng lớn. Tuy nhiên mức tăng là không đáng kể đối với các hệ thống có nhiều hơn 64 cổng.

Qua kết quả tính toán số và mô phỏng có thể thấy rõ là tuy mô hình kiến trúc chuyển mạch Benes sử dụng kỹ thuật điều khiển tập trung có giải thuật điều khiển đơn giản, nhưng khi số lượng cổng đầu vào ra tăng trên 64 cổng thì mô hình chuyển mạch phức tạp, thông lượng giảm nhanh, trễ xử lý là quá lớn và không thể chấp nhận được. Đối với mô hình kiến trúc Spanke cổ điển sử dụng kỹ thuật điều khiển phân tán đã giải quyết được các nhược điểm của kiến trúc Benes về thông lượng truyền dẫn và trễ xử lý tuy nhiên cũng rất phức tạp và chi phí đầu tư tốn kém khi số lượng cổng vào ra tăng trên 64 cổng.

Để cải thiện được các nhược điểm của kiến trúc chuyển mạch Spanke cổ điển, tác giả đề xuất một kiến trúc chuyển mạch gói toàn quang sử dụng kiến trúc chuyển mạch Spanke cải tiến kết hợp WDM sử dụng kỹ thuật xử lý mào đầu gói quang (OHP).

2.4. Kiến trúc chuyển mạch gói toàn quang đề xuất

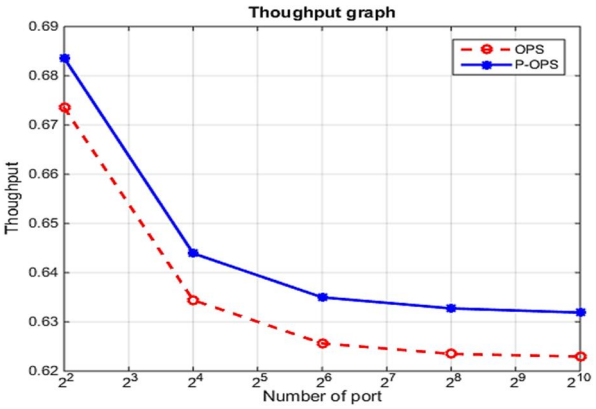
Kiến trúc chuyển mạch gói toàn quang đề xuất (P-OPS) như hình 9. Kiến trúc P-OPS này có khả năng mở rộng dễ dàng, đảm bảo được các yêu cầu kỹ thuật của trung tâm dữ liệu quy mô lớn. Trong kiến trúc có F cổng (sợi) vào và F cổng ra. Mỗi cổng vào/ cổng ra F có M bước sóng. Do đó, tổng số cổng vào/ cổng ra logic là $(N = F \times M)$. Mỗi mô-đun quang (chuyển mạch toàn quang) có một đầu vào và F đầu ra khác nhau, được chuyển đến khối giải quyết tranh chấp (CRB) ở đầu ra của kiến trúc P-OPS. Các khối CRB giải quyết tranh chấp của các gói đến từ các mô-đun quang khác nhau và xuất ra cùng một cổng ra.



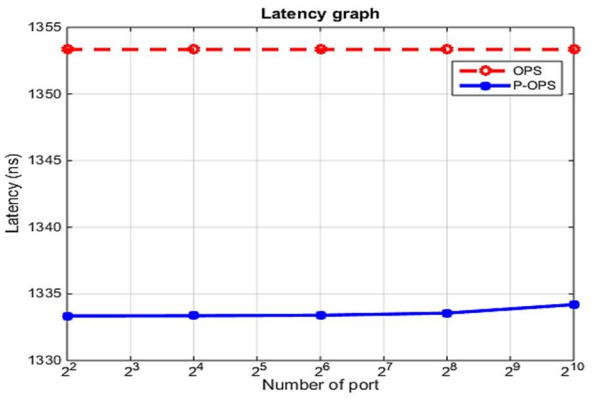
Hình 9. Kiến trúc chuyển mạch gói quang đề xuất

Như trong hình 9, mỗi mô-đun quang bao gồm một bộ xử lý mào đầu gói quang (OHP) dựa trên kỹ thuật điều chế vị trí xung (PPM) [6, 7], một khối chuyển mạch $(1 \times F)$ và bộ điều khiển để điều khiển chuyển mạch các gói đến các khối CRB tương ứng và đưa ra các cổng ra yêu cầu. Các mô-đun quang hoạt động như sau: Bộ xử lý mào đầu toàn quang sẽ tách mào đầu ra khỏi tải trọng gói và được xử lý để tách địa chỉ mào đầu và điều khiển các chuyển mạch trong khối SW để chuyển các gói đến cổng ra yêu cầu. Nhờ sử dụng bộ xử lý mào đầu này mà thời gian xử lý các gói trong kiến trúc giảm xuống đáng kể chỉ vài trăm ps [6, 7].

Kết quả so sánh hiệu năng của chuyển mạch gói quang đề xuất (P-OPS) và chuyển mạch gói quang Spanke cổ điển (OPS) như đưa ra trên hình 10 khi tốc độ bit truyền gói bằng 100Gb/s và số bước sóng sử dụng là M = 4. Hình 10a biểu thị thông lượng trung bình hệ thống thay đổi theo số cổng đầu vào trong hai kiến trúc chuyển mạch gói: OPS, P-OPS. Rõ ràng là kiến trúc chuyển mạch Spanke cải tiến kết hợp WDM sử dụng kỹ thuật xử lý mào đầu gói toàn quang (OHP) có thông lượng trung bình hệ thống cao hơn so với kiến trúc chuyển mạch Spanke cổ điển. Hình 10b biểu thị trễ của hệ thống thay đổi theo số cổng đầu vào trong hai kiến trúc chuyển mạch gói: OPS, P-OPS. Từ hình 10b cho thấy trễ của hệ thống trong kiến trúc chuyển mạch Spanke đề xuất nhỏ hơn so với trễ của hệ thống trong kiến trúc chuyển mạch Spanke cổ điển.



(a)



(b)

Hình 10. So sánh hiệu năng của chuyển mạch gói quang OPS và P-OPS:
(a)Thông lượng trung bình hệ thống thay đổi theo số cổng đầu vào
(b) Trễ của hệ thống thay đổi theo số cổng đầu vào

3. KẾT LUẬN

Trong bài báo, trên cơ sở phân tích đánh giá hiệu năng của kiến trúc chuyển mạch Benes và kiến trúc chuyển mạch Spanke, tác giả đã đề xuất một kiến trúc chuyển mạch gói toàn quang mới nhờ kết hợp kiến trúc Spanke cải tiến và kỹ thuật ghép phân chia theo bước sóng (WDM), đồng thời trong kiến trúc chuyển mạch sử dụng bộ xử lý mào đầu gói quang (OHP). Kiến trúc OPS đề xuất này có khả năng tăng dung lượng các cổng mà vẫn đáp ứng được các yêu cầu đặt ra cho các trung tâm dữ liệu lớn. Bên cạnh đó, với việc sử dụng bộ xử lý mào đầu gói toàn quang đã giảm được đáng kể thời gian xử lý các gói tại kiến trúc chuyển mạch và đã góp phần cải thiện được hiệu năng của hệ thống. Cụ thể làm giảm trễ hệ thống trung bình và cải thiện thông lượng chuyển mạch.

TÀI LIỆU THAM KHẢO

[1]. Shu, Y., et al., 2018. *Programmable OPS/OCS hybrid data centre network*. Optical Fiber Technology, 44, 102–114.

[2]. Toru Segawa, et al., 2016. *High-speed Optical Packet Switching for Photonic Datacenter Networks*. NTT Technical Review, Vol. 14, No. 1.

[3]. C. Chaintoutis, A. Bogris and D. Syvridis, 2018. P-Torus: Torus-based Optical Packet Switching Architecture for intra-Data Centre Networks. Photonics in Switching and Computing (PSC), pp. 1-3.

[4]. N. Calabretta, R. P. Centelles, S. Di Lucente and H. J. S. Dorren, 2013. *On the performance of a large-scale optical packet switch under realistic data center traffic.* in IEEE/OSA Journal of Optical Communications and Networking, vol. 5, no. 6, pp. 565-573.

[5]. S.Di Lucente, 2013. *Optical packet switching with distributed control for high performance data center networks*. Diss. Technische Universiteit Eindhoven.

[6]. Son H.Cao, Hong M.Nguyen, and Quy Q.Ho 2015. *Improving Performance of All-Optical Packet Switching Network with a Modified Pulse Position Modulation Routing T*able. International Journal of Electronics and Data Communications, Vol.3, No.3, pp.90-99, 11-2015.

[7]. Son H.Cao, 2017. *Multi-wavelength All-Optical Packet Switching Node Using Modified Pulse Position Modulation Header Processing*. Tạp chí Khoa học và công nghệ, Trường Đại học Công nghiệp Hà Nội số 43, trang 3-7.

AUTHOR INFORMATION

Cao Hong Son

Facuty of Telecommunication 1, Posts and Telecommunications Institute of Technology