

MỘT CẢI TIẾN VỀ ĐIỀU KHIỂN CHẤP NHẬN LẬP LỊCH DỰA TRÊN DỰ BÁO TỐC ĐỘ CHÙM ĐẾN KẾT HỢP ĐƯỜNG TRỄ FDL

Phạm Trung Đức¹, Võ Viết Minh Nhật², Đặng Thanh Chương¹

¹Trường Đại học Khoa học, Đại học Huế,

²Đại học Huế

phamtrungduc@hueuni.edu.vn, vvmnhat@hueuni.edu.vn, dtchuong@hueuni.edu.vn.

TÓM TẮT: Vấn đề điều khiển chấp nhận lập lịch đã được nghiên cứu trong [1][2][3][4], trong đó các tác giả trong [4] đã đề xuất một mô hình điều khiển chấp nhận lập lịch dựa trên dự báo tốc độ chùm đến nhằm giúp các chùm ưu tiên cao dành được nhiều tài nguyên hơn. Tuy nhiên cách làm này khiến cho tỷ lệ mất chùm không ưu tiên tăng lên. Bài viết này sẽ đề xuất một sự kết hợp đường trễ FDL trong điều khiển chấp nhận lập lịch nhằm cải thiện tỉ lệ mất chùm không ưu tiên. Kết quả phân tích và mô phỏng cho thấy rằng tỉ lệ mất chùm không ưu tiên được cải thiện đáng kể khi có sử dụng đường trễ và điều này càng hiệu quả hơn khi nhiều đường trễ khác nhau được sử dụng.

Từ khóa: mạng OBS, QoS, điều khiển chấp nhận lập lịch, dự đoán dựa trên tốc độ, đường trễ FDL.

I. GIỚI THIỆU

Mạng chuyển mạch chùm quang (mạng OBS, Optical Burst Switching) đã và đang phát triển hết sức mạnh mẽ nhằm đáp ứng sự phát triển nhanh chóng của lưu lượng Internet và việc triển khai ngày càng gia tăng các dịch vụ đa phương tiện mới như VoIP, video theo yêu cầu, điện toán đám mây,... Công nghệ OBS được chọn nhằm khai thác hiệu quả hơn băng thông của các sợi quang, tạo ra một cơ sở hạ tầng mạng linh hoạt, có thể cấu hình ở mức chùm (burst) và xử lý được các kiểu lưu lượng bursty được sinh ra bởi các dịch vụ đa phương tiện. Ngoài các yêu cầu về băng thông, các ứng dụng mới cũng đòi hỏi độ tương thích với chất lượng dịch vụ (QoS) ban đầu cần được đảm bảo: như tỉ lệ mất mát dữ liệu trong một giới hạn nhất định, độ trễ đầu cuối trong phạm vi cho phép,... do đó việc hỗ trợ QoS trở nên thiết yếu trong mạng OBS. Hơn nữa, bộ nhớ truy cập quang ngẫu nhiên như RAM chưa thể sản xuất được nên các chùm chỉ có thể được trì hoãn nhờ các đường trễ FDL (Fiber Delay Line). Vì vậy, việc phát triển các cơ chế điều khiển đảm bảo QoS cho mạng OBS trở nên thiết yếu.

Một đặc trưng tiêu biểu của mạng OBS là gói điều khiển chùm (Burst Control Packet - BCP) được tách rời với phần dữ liệu (Data Burst). Nói một cách khác, để thực hiện việc truyền một chùm trong mạng OBS, gói điều khiển BCP được tạo ra và được gửi đi trước một khoảng thời gian offset (offset-time). Thời gian offset này phải đủ lớn để đặt trước tài nguyên và cấu hình các chuyển mạch tại các nút trung gian dọc theo hành trình của chùm từ nguồn đến đích. Thêm vào đó, mạng OBS dành riêng một số kênh (bước sóng), được gọi là kênh điều khiển cho việc truyền gói điều khiển BCP, trong khi các kênh còn lại được dùng cho việc truyền chùm dữ liệu, nên được gọi là kênh dữ liệu. Như vậy việc truyền gói điều khiển BCP là tách rời hoàn toàn với truyền dữ liệu về mặt không gian và thời gian. Với cách truyền dữ liệu như vậy, rõ ràng mạng OBS không cần đến các bộ đệm quang để lưu tạm các chùm dữ liệu trong khi chờ đợi việc xử lý chuyển mạch tại các nút lõi, cũng như không yêu cầu các chuyển mạch ở tốc độ nano giây. Tuy nhiên, cách truyền tải này cũng đặt ra áp lực là làm thế nào để một gói điều khiển BCP kịp lập lịch đặt trước tài nguyên và cấu hình chuyển mạch tại các nút lõi, đảm bảo việc truyền tải chùm quang theo sau; đó chính là nhiệm vụ của hoạt động lập lịch đặt trước tài nguyên tại các nút lõi mạng. Vì vậy vấn đề lập lịch đóng vai trò rất quan trọng trong việc tối đa hiệu suất băng thông, giảm mất mát dữ liệu, đảm bảo chất lượng cho các dịch vụ khác nhau và nâng cao hiệu suất hoạt động của mạng OBS.

Trong mạng OBS, tranh chấp phát sinh khi hai hay nhiều chùm đến tranh chấp cùng tài nguyên trên một công ra. Nếu bước sóng được yêu cầu của một chùm đến bị chiếm giữ tại công ra, chùm có thể chuyển sang sử dụng bước sóng còn rỗi khác bằng cách sử dụng bộ chuyển đổi bước sóng. Trong trường hợp nếu tất cả các kênh bước sóng tại một công ra đều bị chiếm giữ, chùm đến có thể sử dụng đường trễ quang FDL hoặc định tuyến lệch hướng để giải quyết tranh chấp. Một hướng tiếp cận khác trong việc hạn chế tranh chấp tài nguyên là điều khiển chấp nhận lập lịch.

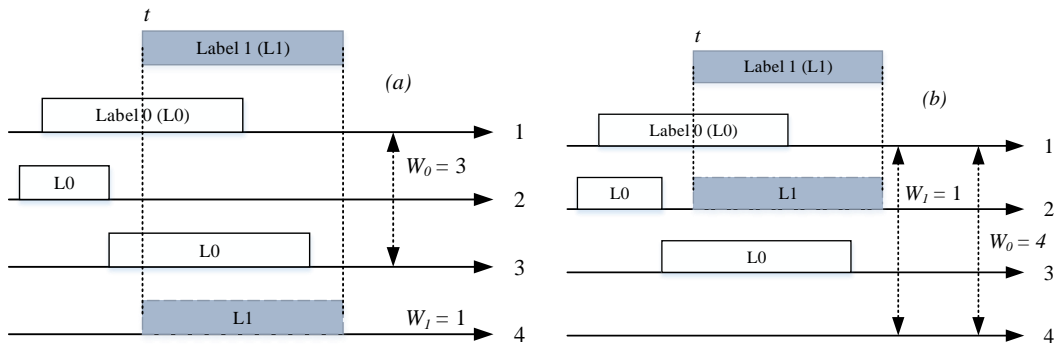
Điều khiển chấp nhận lập lịch (một cách ngắn gọn, điều khiển chấp nhận) có thể được thực hiện tại nút biên hay nút lõi [1]. Trong đa số các mô hình điều khiển chấp nhận đã được đề xuất, các chùm ưu tiên thấp luôn bị đánh rơi nhiều hơn nhằm để dành tài nguyên cho các chùm ưu tiên cao nếu có tranh chấp xảy ra. Việc lập lịch thường được thực hiện một cách tuần tự, theo kiểu *first come, first served*, nhưng trong trường hợp có phân biệt dịch vụ, việc lập lịch một chùm ưu tiên thấp đến có thể gây tắc nghẽn cho một burst ưu tiên cao đến sau. Do đó, lập lịch với điều khiển chấp nhận là cần thiết nhằm để dành nhiều tài nguyên cho chùm QoS cao, trong khi hạn chế tài nguyên đối với các chùm QoS thấp.

Việc điều khiển chấp nhận cũng có thể kết hợp với sử dụng đường trễ FDL nhằm giảm mất mát dữ liệu khi có tranh chấp xảy ra. Một FDL chỉ cho phép làm trễ chùm trong một khoảng thời gian cố định, vì vậy việc tích hợp thêm

FDL vào nút lõi OBS có thể xem như là một bộ đệm với kích thước hạn chế. Tuy nhiên, khác với các bộ đệm điện tử, một chùm không được làm trễ trong một khoảng thời gian không xác định; chùm sẽ bị đánh rơi sau khi ra khỏi đường trễ nếu không được phục vụ.

II. CÁC NGHIÊN CỨU LIÊN QUAN

Đã có một số kỹ thuật điều khiển được đề xuất trong mạng OBS có hỗ trợ QoS, như kỹ thuật điều khiển tĩnh SWG (Static Wavelength Grouping) [1], kỹ thuật điều khiển động DWG (Dynamic Wavelength Grouping) [1], kỹ thuật điều khiển dựa trên mức tải (Load-Level Admission Control, LLAC) [2] và kỹ thuật điều khiển dựa trên dự báo lưu lượng (Traffic Prediction based Admission Control, TPAC) [4]. Cụ thể, xét tại một cổng ra có W kênh bước sóng khả dụng, kỹ thuật SWG sẽ phân bổ W_0 ($W_0 < W$) bước sóng cho các chùm gắn nhãn QoS cao (Label 0) và W_1 ($W_1 < W$) bước sóng cho các chùm được gắn nhãn QoS thấp (Label 1), trong đó $W_0 > W_1$ và $W_0 + W_1 = W$. Như ví dụ được chỉ ra trong **Hình 1a**, với $W = 4$, kỹ thuật SWG phân bổ $W_0 = 3$ bước sóng (kênh số 1, 2 và 3) cho các chùm QoS cao và $W_1 = 1$ bước sóng còn lại (kênh số 4) cho các chùm QoS thấp; nên một chùm QoS thấp đến tại thời điểm t sẽ được lập lịch vì kênh 4 rồi.

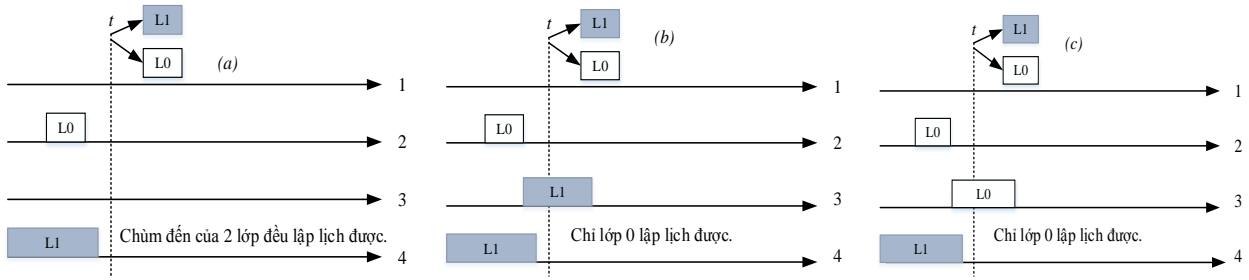


Hình 1. Ví dụ về điều khiển lập lịch của SWG (a) và DWG (b).

Ưu điểm của kỹ thuật SWG là đơn giản, đảm bảo dành riêng các kênh bước sóng cho các chùm QoS cao nên đảm bảo chất lượng dịch vụ. Tuy nhiên, sẽ xuất hiện một sự phân bố tải không đồng đều trên các kênh bước sóng nếu các chùm đến chỉ thuộc về một phía. Cụ thể, nhóm các kênh W_0 (hoặc W_1) sẽ quá tải, trong khi nhóm các kênh W_1 (hoặc W_0) là nhàn rỗi nếu các chùm đến chỉ thuộc về QoS cao (hoặc QoS thấp). Để linh hoạt hơn trong sử dụng các kênh bước sóng, các tác giả trong [1] đã đề xuất tiếp kỹ thuật DWG, trong đó các kênh không được chỉ định cụ thể cho các lớp QoS mà chỉ kiểm soát số lượng kênh được phép sử dụng tối đa cho mỗi loại chùm QoS cao và QoS thấp. Như ví dụ được chỉ ra trong **Hình 1b**, xét tại thời điểm đến (t) của một chùm đến, nếu số chồng lấp của chùm này với các chùm cùng loại (ω_0 và ω_1) mà nhỏ hơn lượng bước sóng được phân bổ ($\omega_0 < W_0$ và $\omega_1 < W_1$), việc lập lịch là được chấp nhận. Do đó một chùm QoS thấp đến tại thời điểm t vẫn có thể được lập lịch lên kênh 2 hoặc kênh 4 vì số chồng lấp đối với loại chùm này là $\omega_1 = 0$, trong khi $W_1 = 1$, nên $\omega_1 < W_1$ và trong số lượng một kênh được sử dụng tối đa này, có thể lập lịch trên bất cứ kênh nào thỏa mãn (kênh 2 và 4) chứ không cấp phát cố định kênh 4 như trong SWG.

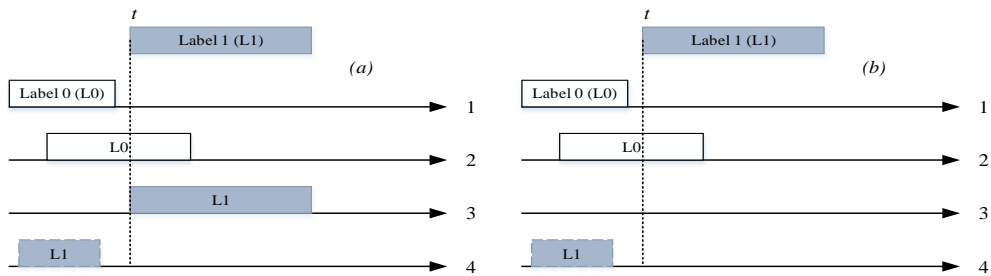
Ưu điểm của kỹ thuật DWG là phân bổ tải đồng đều trên các kênh hơn so với SWG nhưng cũng như SWG, kỹ thuật DWG cũng gây lãng phí tài nguyên đối với các kênh dành cho các chùm QoS cao nếu mật độ các chùm này đến thấp, trong khi các chùm QoS thấp bị đánh rơi vì thiếu tài nguyên. Cả hai kỹ thuật SWG và DWG đều yêu cầu lưu lại thông tin trạng thái của tất cả chùm đang được lập lịch trên các kênh nên cần nhiều bộ nhớ cho việc lưu trữ. Một hạn chế khác của hai kỹ thuật SWG và DWG là chúng chỉ hiệu quả khi biết trước lưu lượng các luồng chùm (ưu tiên thấp và cao) đến và các lưu lượng này không biến đổi đáng kể trong một thời gian dài; nhưng điều này ngược với thực tế vì lưu lượng trên mạng luôn thay đổi.

Để tăng hiệu quả việc phân phối các kênh bước sóng ra và giảm không gian lưu trữ, các tác giả trong [2] đã đề xuất kỹ thuật LLAC, trong đó chỉ có thông tin về mức tải của mỗi lớp và tổng số bước sóng bị chiếm là được lưu trữ. Các mức tải (l_0 và l_1) sẽ được qui định trước cho mỗi lớp QoS cao và thấp; nếu số chồng lấp của một chùm đến với bất kỳ loại chùm nào đã được lập lịch trên các kênh ra (ω) là nhỏ hơn mức tải ($\omega < l_0$ và $\omega < l_1$) chùm đến sẽ được lập lịch. Như ví dụ được chỉ ra trong **Hình 2a**, với hai mức tải $l_0 = 4$ và $l_1 = 1$ được thiết lập trước đối với 2 loại chùm QoS cao và thấp, khi có một chùm QoS thấp đến tại thời điểm t , nó được lập lịch vì $\omega < l_1$ ($\omega = 0$). Mặt khác, tại cùng thời điểm t , nếu chỉ một chùm QoS cao đến thì nó được lập lịch vì $\omega < l_0$ (lúc này $\omega = 0$). Tuy nhiên trong **Hình 2b**, với một chùm QoS cao đến tại thời điểm t , nó được lập lịch vì $\omega < l_0$ ($\omega = 1$). Trong khi tại cùng thời gian đến t , một chùm QoS thấp đến thì nó sẽ không được lập lịch do $\omega = l_1$ (lúc này $\omega = 1$). Tương tự với **Hình 2c**, một chùm QoS thấp đến tại thời điểm t sẽ không được lập lịch vì $\omega = l_1$ ($\omega = 1$), mặc dù chùm ưu tiên thấp này chồng lấp với chùm khác loại (ưu tiên), trong khi một chùm QoS cao đến tại cùng thời điểm t sẽ được lập lịch bởi vì $\omega = 1 < l_0$.



Hình 2. Các ví dụ về cách thức hoạt động của kỹ thuật LLAC.

Điều đó, khiến kỹ thuật LLAC có nhược điểm là nó có xu hướng đánh rơi các chùm QoS thấp nhiều hơn so với kỹ thuật DWG. Được khẳng định thông qua ví dụ như trong **Hình 3**, kỹ thuật DWG sẽ không đánh rơi chùm ưu tiên thấp khi nó đến tại thời điểm t (**Hình 3a**); trong khi kỹ thuật LLAC sẽ đánh rơi nó vì $\omega = l_1$ ($\omega = 1$) như **Hình 3b**. Ngoài ra, một nhược điểm khác của LLAC là không đưa ra bất cứ cách nhận biết nào về tải đến trong khi ý tưởng của đề xuất là giả định biết trước được tải lưu lượng sau đó phân bổ kênh dựa trên tải biết trước này.



Hình 3. Một so sánh về hiệu quả lập lịch giữa kỹ thuật DWG (a) và LLAC (b).

Một cải tiến của DWG là LDWG (Load-based Dynamic Wavelength Grouping) được đề xuất trong [3], trong đó việc phân bổ bước sóng là được dựa trên lưu lượng tải đến của mỗi lớp ưu tiên. Cụ thể, một hệ số phân bổ tài nguyên được định nghĩa là:

$$p_i = \frac{Tai_cua_tung_lop_i}{Tong_tai}$$

Như vậy, nếu có 60% lưu lượng tải ưu tiên cao và 40% lưu lượng tải ưu tiên thấp đến tại một liên kết ra có W bước sóng, tỉ lệ phân bổ bước sóng sẽ lần lượt sẽ là $0.6W$ và $0.4W$ cho lớp ưu tiên cao và ưu tiên thấp. Để sử dụng hiệu quả bước sóng đã được cấp phát và giảm mất burst, các tác giả trong [3] còn đề xuất rằng lưu lượng tải thuộc lớp ưu tiên cao có thể sử dụng các bước sóng dư thừa đã cấp phát cho lớp ưu tiên thấp và ngược lại. Tuy nhiên, tính hiệu quả của mô hình này chỉ được mô tả thông qua một ví dụ mà không có một đánh giá nào dựa trên mô phỏng.

Tóm lại, cả 4 kỹ thuật SWG, DWG, LLAC và LDWG đều không xét việc phân phối băng thông trong môi trường mạng mà ở đó lưu lượng chùm đến công ra thay đổi liên tục. Các tác giả trong [1] đã khắc phục vấn đề này bằng việc đề xuất mô hình điều khiển chấp nhận dựa trên phương pháp ước tính lưu lượng chùm đến (kỹ thuật TPAC) và dành thêm tài nguyên cho chùm ưu tiên nhằm phân phối các kênh ra một cách hiệu quả hơn. Cụ thể, với W kênh bước sóng khả dụng tại một liên kết ra, mà tại đó một chùm ưu tiên cao đến có thể được lập lịch lên một trong W kênh bước sóng; trong khi một chùm ưu tiên thấp đến chỉ được lập lịch trên một trong W_1 kênh, với $W_1 < W$. Cách làm này nhằm để dành tài nguyên nhiều hơn cho các chùm ưu tiên cao và việc cấp phát lại bước sóng chỉ được thực hiện cho lưu lượng các chùm ưu tiên thấp đến, trong đó băng thông được cấp phát cho luồng này (W_1) cần được tăng lên nếu các chùm ưu tiên thấp đến dày đặc và các chùm ưu tiên cao đến thưa thớt. Nói một cách khác, số bước sóng phân bổ cho các chùm ưu tiên thấp cần tỷ lệ với tốc độ đến của các burst ưu tiên thấp so với tổng tốc độ của cả 2 lớp ưu tiên, như Công thức (1).

$$\frac{W_1}{W} = \frac{\lambda_1}{\lambda_0 + \lambda_1} \Rightarrow W_1 = W \frac{\lambda_1}{\lambda_0 + \lambda_1} \tag{1}$$

trong đó λ_0 là tốc độ đến của các burst ưu tiên cao và λ_1 là tốc độ đến của các burst ưu tiên thấp.

Để xác định λ_0 và λ_1 cho việc phân bổ lại kênh bước sóng, phương pháp TW-EWMA [5] là được sử dụng dựa trên thống kê của tốc độ trung bình trong quá khứ và tốc độ hiện thời của các chùm đến. Cụ thể, tốc độ đến của các lớp ưu tiên (0 với lớp ưu tiên cao và 1 với lớp tiên thấp) là:

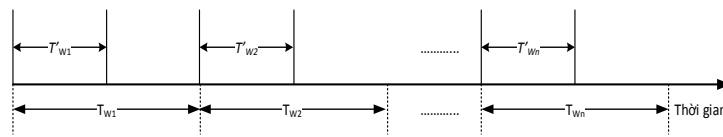
$$\lambda_i = (1 - \alpha) \times \lambda_i^{avg} + \alpha \times \lambda_i^{cur} \tag{2}$$

trong đó λ_i^{avg} là tốc độ chùng đến trong quá khứ, λ_i^{cur} là tốc độ chùng đến hiện thời; α là trọng số của chúng.

Trong [5], α được chọn bằng 0.3; nhưng theo đề xuất của các tác giả trong [6], α nên được điều chỉnh linh hoạt dựa trên tốc độ đến hiện thời và tổng tốc độ của các lớp như (3).

$$\frac{1 - \alpha}{\alpha} = \frac{\lambda_i^{avg}}{\lambda_i^{cur}} \Rightarrow \alpha = \frac{\lambda_i^{cur}}{\lambda_i^{cur} + \lambda_i^{avg}} \tag{3}$$

Một đặc điểm khác của TW-EWMA là cửa sổ quan sát để tính toán λ_i^{cur} là rời rạc thay vì liên tục để giảm chi phí tính toán (**Hình 4**). Thực tế, kích thước cửa sổ quan sát có một tác động đáng kể đến mức độ chính xác của việc dự đoán và chi phí tính toán: Nếu cửa sổ lớn thì việc dự đoán sẽ chính xác hơn, nhưng chi phí tính toán cao; trong khi nếu cửa sổ quan sát bé thì tính chính xác của dự đoán sẽ giảm, nhưng chi phí tính toán thấp [5]. Một thỏa hiệp giữa tính chính xác và khối lượng tính toán do đó cần được tính đến. Trong bài viết này, cửa sổ quan sát được chọn bằng một phần hai chu kỳ phân bố lại bước sóng (T_{W_i}).



Hình 4. Cửa sổ quan sát là rời rạc thay vì liên tục như trong TW-EWMA.

Từ (1), (2) và (3), số lượng bước sóng được phân bổ cho các burst ưu tiên thấp (W_i) thay đổi một cách thích nghi theo sự thay đổi bất thường của lưu lượng chùng đến.

Một kỹ thuật ưu tiên tuyệt đối đối với các chùng ưu tiên cao cũng được đề xuất trong mô hình điều khiển chấp nhận này. Các bước nhường lại tài nguyên cho burst ưu tiên cao được thực hiện như sau:

Khi một chùng ưu tiên cao đến một liên kết ra và không tìm thấy bước sóng cho việc lập lịch nó, tài nguyên bị chiếm dụng bởi một chùng ưu tiên thấp sẽ được giải phóng để lập lịch cho chùng ưu tiên cao này với điều kiện:

- Chùng ưu tiên cao chông lắp với chỉ một chùng ưu tiên thấp mà sẽ được gỡ ra; và
- Gói điều khiển của chùng ưu tiên thấp này chưa được gửi đến nút tiếp theo.

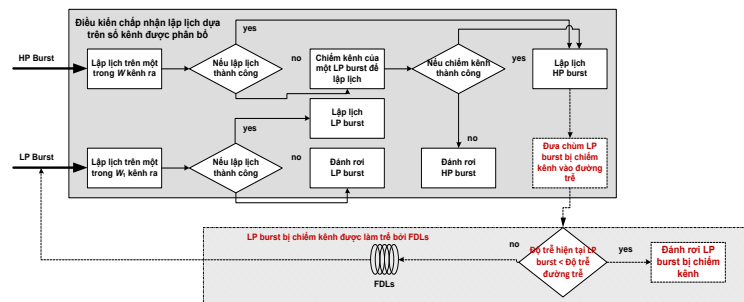
Nếu không, chùng ưu tiên cao mới đến bị đánh rơi.

Với trường hợp chùng ưu tiên thấp đến và tất cả tài nguyên đều bận, chùng này sẽ bị đánh rơi.

Tuy nhiên, việc ưu tiên dành nhiều tài nguyên cho chùng ưu tiên cao trong kỹ thuật điều khiển TPAC đã là tăng tỷ lệ mất chùng không ưu tiên. Phần tiếp theo sẽ trình bày một đề xuất mô hình điều khiển chấp nhận dựa trên dự đoán tốc độ chùng đến kết hợp đường trễ FDL nhằm làm giảm tỉ lệ mất chùng ưu tiên thấp.

III. ĐIỀU KHIỂN LẬP LỊCH KẾT HỢP ĐƯỜNG TRỄ FDL.

Như đã được phân tích trong Phần II, giải thuật TPAC [4] đã cải thiện đáng kể tỉ lệ mất chùng trên toàn mạng và ở lớp ưu tiên. Tuy nhiên, tỉ lệ mất dành cho lớp ưu tiên thấp vẫn còn cao. Để khắc phục vấn đề này chúng tôi đề xuất một giải thuật có tên là *iTPAC* (*improved TPAC*) là một cải tiến từ TPAC [4] chi tiết xem ở **Hình 5**.



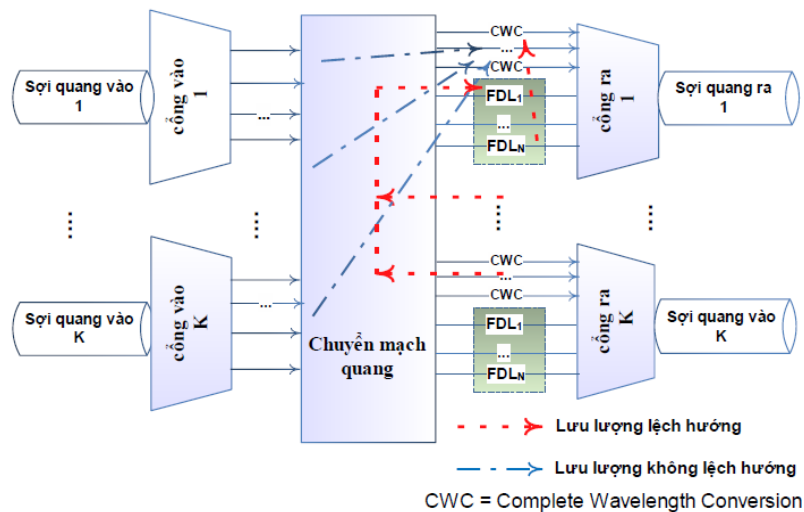
Hình 5. Mô tả cách thức sử dụng đường trễ FDL trong đề xuất.

Như có thể nhìn thấy trong **Hình 5** với giải thuật TPAC khi một chùng ưu tiên cao (HP burst) đến nếu lập lịch không thành công (do không còn tài nguyên để lập lịch) thì nó sẽ chiếm kênh của một chùng ưu tiên thấp (LP burst) để lập lịch. Nếu quá trình chiếm kênh thành công chùng ưu tiên cao sẽ được lập lịch và chùng ưu tiên thấp sẽ bị đánh rơi.

Đây chính là nguyên nhân làm cho tỉ lệ mất chùm của lớp ưu tiên thấp vẫn còn cao. Do đó, thay vì đánh rơi các chùm ưu tiên thấp, chúng sẽ được đưa vào đường trễ. Một cơ chế điều khiển việc làm trễ bằng FDL đối với các chùm ưu tiên thấp này được đề xuất như sau:

- nếu độ trễ cho phép (được xác định dựa trên thời gian offset hiện tại) của chùm ưu tiên thấp là bé hơn thời gian làm trễ của đường trễ FDL thì vì việc đưa chùm ưu tiên thấp vào đường trễ là không cần thiết (do chúng sẽ bị đánh rơi vì đã hết thời gian offset nhưng vẫn chưa đạt đến nút biên ra) và chùm ưu tiên thấp này sẽ bị đánh rơi;
- nếu độ trễ cho phép của chùm ưu tiên thấp lớn hơn độ dài đường trễ, chùm này sẽ được đưa vào đường trễ với hy vọng có thể tìm thấy tài nguyên để lập lịch khi ra khỏi đường trễ.

Do đường trễ quang dựa trên độ trễ lan truyền của sợi quang nên có nhiều hạn chế so với bộ đệm điện tử RAM nếu xét đến khả năng truy cập liên tục. Thực tế, trong bất kỳ kỹ thuật sử dụng bộ đệm quang nào, kích thước của các bộ đệm bị giới hạn rất nghiêm ngặt, không những bởi chất lượng tín hiệu mà còn bởi sự giới hạn về không gian vật lý. Nếu dung lượng bộ đệm lớn thì đòi hỏi số lượng và chiều dài của đường trễ quang càng tăng nên dễ gây tổn hao và việc sử dụng bộ đệm cũng không thể hoàn toàn giảm khả năng mất chùm. Trong bài báo này, chúng tôi chọn độ dài tối thiểu của đường trễ FDL là $50\mu\text{s}$ nhằm đảm bảo cho chùm ưu tiên thấp có kích thước lớn nhất có thể sử dụng được. Một mô tả kiến trúc đường trễ quang FDL truyền thẳng [7] được thể hiện như **Hình 6**.



Hình 6. Kiến trúc đường trễ FDL truyền thẳng tại nút lõi được sử dụng [7].

Thuật toán *iTPAC* được cải tiến từ *TPAC* được mô tả như sau:

Thuật toán *iTPAC*:

Đầu vào:

- $W = \{1, 2, \dots, n\}$: số kênh ra;
- T_c ; $//T_c = 10000$

Đầu ra:

Xử lý:

```

1  while chùm ub đến do
2      iBF-VF(ub, Wi, W);
3      if (đạt đến ngưỡng Tc) then
4          λicur := tỷ lệ hiện tại của các lớp HP và LP đến trong của số quan sát;
5          αi := λicur / (λicur + λiavg);
6          λiavg := (1 - αi) × λiavg + αi × λicur;
7          W1 := W0 ×  $\frac{\lambda_1^{avg}}{\lambda_0^{avg} + \lambda_1^{avg}}$ ;
8      end if
9  end while;
```

Thuật toán *iBF-VF*:

Đầu vào:

- $ub = \{s_{ub}, e_{ub}, prio_{ub}, delay_{ub}\}$, trong đó s_{ub} và e_{ub} là thời gian đến và kết thúc, $prio_{ub} =$

```

{0,1} là mức QoS cao (0) và thấp (1) của chùm đến  $ub$ ,  $delay_{ub}$  là độ trễ của chùm  $ub$ ;
-  $L_{delay}$ ; // Độ dài tối đa của đường trễ.
-  $W_1$ ;
-  $W$ .

```

Đầu ra: - $state$; // $state = -1$ nếu việc lập lịch không thành công.

Xử lý:

```

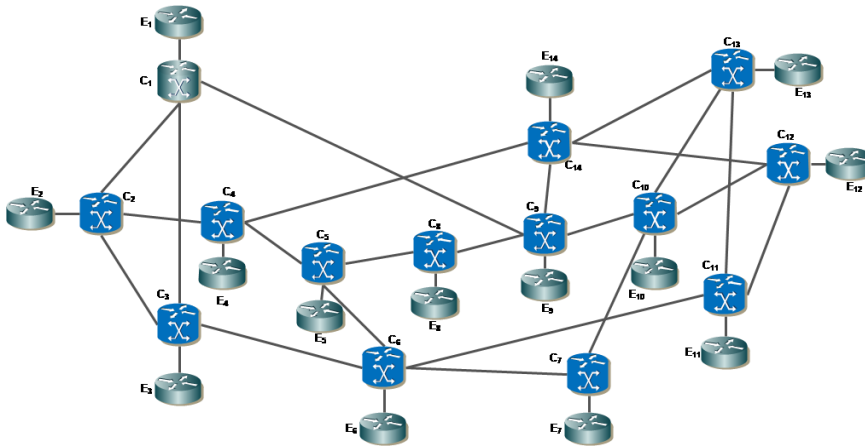
1  if ( $prio_{ub} = 1$ ) then
2      overlap := tổng chồng lấp của chùm  $ub$  với chùm LP đã lập lịch;
3      if overlap  $\geq W_1$  then return -1;
4  end if;
5  best_channel := BFVF( $ub$ ,  $W$ );
6  if ( $best\_channel = -1$ )  $\wedge$  ( $prio_{ub} = 0$ ) then
7      best_burst := Replace( $ub$ ,  $W$ ); //dành tài nguyên không ưu tiên
8      if ( $prio_{ub} = 1$ )  $\wedge$  ( $delay_{ub} < L_{delay}$ ) then
9          UseFDL( $ub$ ,  $W_1$ ); //kết hợp sử dụng đường trễ FDL nếu thỏa điều kiện.
10     end if
11 end if

```

Độ phức tạp của giải thuật **iBF-VF** được xác định trong [8] là $O(W \log(m))$, trong đó W là tổng số bước sóng tại liên kết ra và m là số chùm trung bình đã lập lịch trên mỗi kênh. Do tốc độ chùm đến và tốc độ xử lý trung bình tại mỗi nút lõi OBS là rất nhanh, nên chỉ thường 2 chùm đã lập lịch cuối cùng là được xem xét; độ phức tạp của **iBF-VF** do đó giảm còn $O(W)$. Hàm **Replace** tìm trong W bước sóng và chọn bước sóng đầu tiên mà trên đó một chùm ưu tiên thấp đã lập lịch có thể được xóa để truyền tải tài nguyên cho việc lập lịch ub , nên có độ phức tạp của nó là $O(W)$. Hàm **UseFDL** làm trễ chùm ub bằng cách sử dụng đường trễ FDL. Do **iBF-VF**, **Replace** và **UseFDL** là các hàm độc lập nhau nên thuật toán **iTPAC** có độ phức tạp là $O(W)$.

IV. MÔ PHỎNG VÀ PHÂN TÍCH KẾT QUẢ

Mô phỏng được cài đặt trên máy tính có CPU 2,4 GHz Intel Core 2,2G RAM. Mô hình mạng được xem xét là NSFNET (**Hình 7**) trong đó băng thông trên mỗi liên kết là 10Gb/s. Các gói tin đến tại các nút biên vào có phân phối Poisson giả sử chỉ thuộc về một trong 2 lớp (ưu tiên cao và ưu tiên thấp); Các chùm được sinh ra do đó cũng có phân phối Poisson và thuộc về một trong 2 lớp này. Lưu lượng tải chuẩn hóa (tải chuẩn hóa được định nghĩa là bằng tốc độ đến trên khả năng đáp ứng của băng thông) được xem xét thay đổi từ 0.1 đến 0.9 theo 3 tỉ lệ 3:7, 5:5 và 7:3 giữa luồng ưu tiên cao và ưu tiên thấp. Số bước sóng trên liên kết ở cổng ra $W = 12$, cửa sổ quan sát $T_w = 5ms$. Dữ liệu được trích xuất từ NS2 với gói hỗ trợ mạng OBS là obs-0.9a trong thời gian mô phỏng 10s.



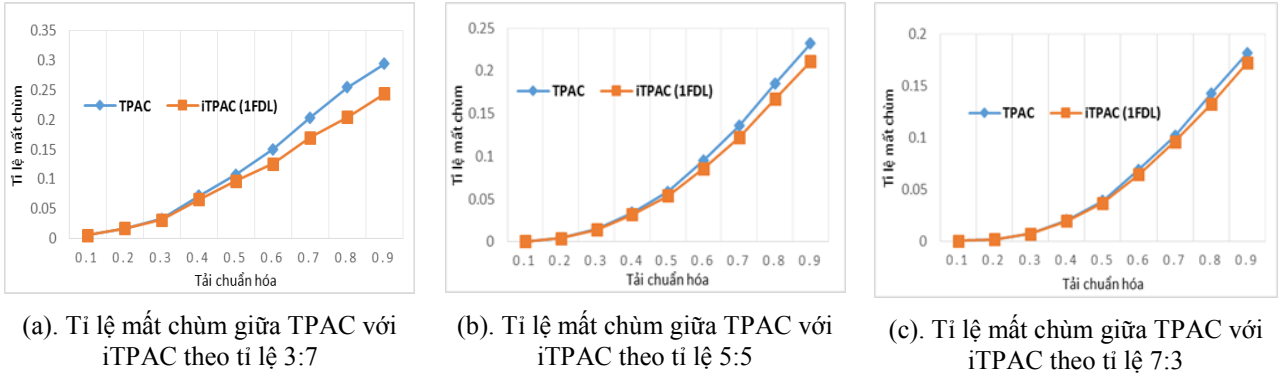
Hình 7. Mô hình mạng mô phỏng NSFNET.

Các mục tiêu mô phỏng bao gồm:

1. So sánh tỉ lệ mất chùm của lớp ưu tiên thấp giữa TPAC với *iTPAC* khi tiến hành thay đổi độ dài đường trễ từ $50\mu s$ đến $300\mu s$.
2. So sánh tỉ lệ mất chùm của lớp ưu tiên thấp giữa TPAC với *iTPAC* khi sử dụng 1, 2 và 3 đường trễ.

4.1. So sánh tỉ lệ mất chùm của lớp ưu tiên thấp giữa TPAC với *iTPAC* khi thay đổi độ dài đường trễ từ $50\mu s$ đến $300\mu s$.

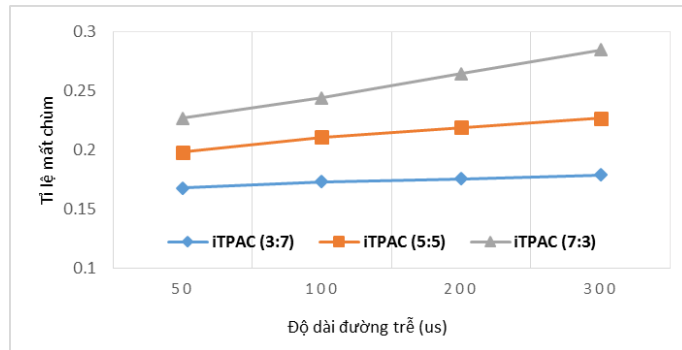
Đầu tiên chúng tôi tiến hành so sánh tỉ lệ mất chùm của lớp ưu tiên thấp giữa TPAC với *iTPAC* khi sử dụng một đường trễ với độ dài đường trễ là $100\mu s$ theo 3 tỉ lệ 3:7, 5:5 và 7:3, kết quả như được chỉ ra ở **Hình 8**.



Hình 8. Tỷ lệ mất chùm của lớp ưu tiên thấp giữa TPAC và iTPAC

Hình 8 cho thấy rằng tỉ lệ mất chùm của lớp ưu tiên thấp đã giảm đi đáng kể ở cả 3 tỉ lệ 3:7, 5:5 và 7:3. Cụ thể ở tỉ lệ 3:7 ở tải 0.9 iTPAC giảm gần 20%, 15% ở tỉ lệ 5:5 và 10% ở tỉ lệ 7:3. Nguyên nhân là ở tỉ lệ 3:7 số lượng chùm của lớp ưu tiên thấp đến nhiều nên việc sử dụng đường trễ cải thiện được nhiều tỉ lệ mất hơn so với 5:5 và 7:3. Việc làm giảm tỉ lệ mất chùm ở iTPAC không làm ảnh hưởng đến tỉ lệ mất chùm của lớp ưu tiên cao do không có thay đổi nào về chính sách đối với chùm ưu tiên cao trong iTPAC so với TPAC.

Để làm rõ việc thay đổi độ dài đường trễ có ảnh hưởng đến tỉ lệ mất chùm của lớp ưu tiên thấp hay không, chúng tôi tiến hành thay đổi giá trị đường trễ từ 50µs đến 300µs theo tải chuẩn hóa 0.9, kết quả thu được như được chỉ ra ở Hình 9.

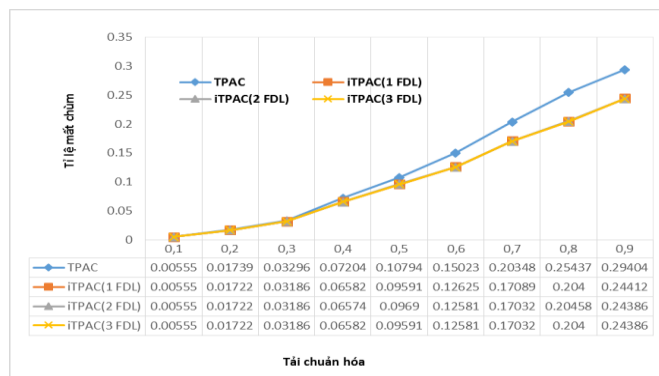


Hình 9. Tỷ lệ mất chùm ở lớp không ưu tiên của iTPAC khi thay đổi độ dài đường trễ (FDL) ở tải chuẩn hóa 0.9.

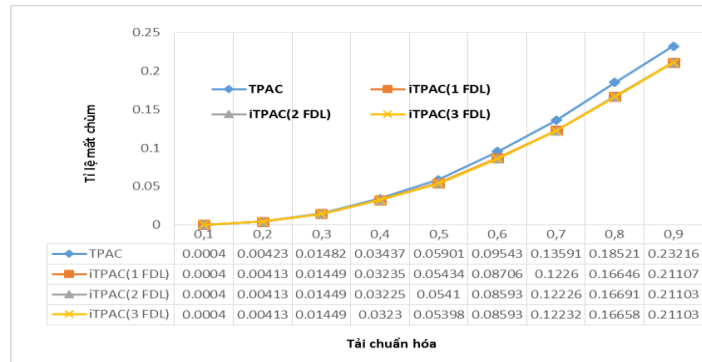
Hình 9 cho nhận thấy rằng khi tăng độ dài đường trễ thì tỉ lệ mất chùm có xu hướng tăng ở mọi tỉ lệ 3:7, 5:5 và 7:3, do khi tăng độ dài đường trễ thì số lượng các chùm được đưa vào đường trễ sẽ giảm đi, như mô tả ở trong Hình 5, vì chỉ các chùm chịu được độ trễ tăng thêm mới được đưa vào đường trễ này. Lưu ý rằng, giá trị đường trễ tối thiểu phải lớn hơn so với kích thước chùm để chùm có thể làm trễ trong đường trễ.

4.2. So sánh tỉ lệ mất chùm của lớp ưu tiên thấp giữa TPAC với iTPAC khi sử dụng 1, 2 và 3 đường trễ

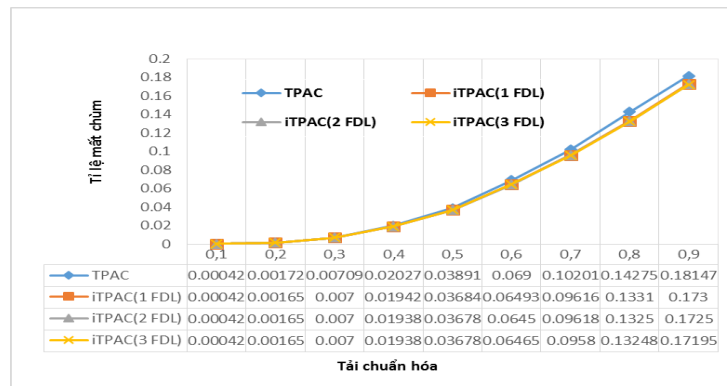
Chúng tôi tiếp tục so sánh tỉ lệ mất chùm của lớp ưu tiên thấp trong 4 trường hợp: (1) sử dụng giải thuật TPAC; (2) sử dụng giải thuật iTPAC với sử dụng một đường trễ có độ dài là 100µs; (3) sử dụng giải thuật iTPAC với sử dụng 2 đường trễ 100µs và 200µs; (4) sử dụng giải thuật iTPAC với việc sử dụng 3 đường trễ 100µs, 200µs và 300µs. Tỉ lệ các chùm ưu tiên và không ưu tiên đến được xem xét là 3:7, 5:5 và 7:3. Kết quả mô phỏng thu được như được chỉ ra ở Hình 10, 11 và 12.



Hình 10. Tỷ lệ mất chùm của lớp ưu tiên thấp khi sử dụng đồng thời 1, 2 và 3 đường trễ theo tỉ lệ 3:7.



Hình 11. Tỷ lệ mất chùm của lớp ưu tiên thấp khi sử dụng đồng thời 1, 2 và 3 đường trễ theo tỉ lệ 5:5.



Hình 12. Tỷ lệ mất chùm của lớp ưu tiên thấp khi sử dụng đồng thời 1, 2 và 3 đường trễ theo tỉ lệ 7:3.

Từ các Hình 10, 11, 12 chúng ta nhận thấy rằng tỉ lệ mất chùm khi sử dụng càng nhiều đường trễ cùng lúc thì tốt hơn ở cả 3 tỉ lệ. Tuy nhiên, tỉ lệ chênh lệch là không cao, do số lượng các chùm của lớp ưu tiên thấp đi vào đường trễ thứ 2 và thứ 3 ít đi, vì ở đường trễ này độ trễ của nó là khá lớn, một số vượt quá độ trễ tối đa cho phép của các chùm.

V. KẾT LUẬN

Trong bài báo này, chúng tôi đã đề xuất giải thuật iTPAC, là một bước cải tiến của TPAC nhằm tối ưu tỉ lệ mất chùm của lớp ưu tiên thấp bằng cách sử dụng đường trễ. Việc làm trễ các chùm ưu tiên thấp là nhằm trao cơ hội thứ tự lập lịch lại sau một khoảng trễ. Như được chỉ ra ở các Hình 8 đến Hình 12 việc sử dụng đường trễ đã làm cho tỉ lệ mất chùm ở lớp ưu tiên thấp giảm đi đáng kể. Tuy nhiên, chi phí phải bỏ ra để thiết lập đường trễ là không hề nhỏ. Việc sử dụng nhiều đường trễ cùng lúc sẽ làm giảm tỉ lệ mất chùm, nhưng độ trễ gia tăng cũng làm giảm đáng kể tỉ lệ sử dụng càng nhiều đường trễ này. Do đó một sự thỏa hiệp giữa hiệu quả lập lịch và chi phí sử dụng đường trễ cần có các nghiên cứu sâu.

TÀI LIỆU THAM KHẢO

- [1] Q. Zhang, V. M. Vokkarane, J. P. Jue, and B. Chen, "Absolute QoS differentiation in optical burst-switched networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 9, pp. 1781-1795, 2004.
- [2] I. M. Moraes, D. D. O. Cunha, M. D. D. Bicudo, R. P. Laufer, and O. C. M. B. Duarte, "An Admission Control Mechanism for Providing Service Differentiation in Optical Burst-Switching Networks," *Tech. Rep.*, 2010.
- [3] P. R. Reddy, A. Nagarajan, K. Ramanujam, and S. Talabathula, "Reducing Burst Loss Probability of Service Differentiated Optical Burst - Switched Networks," pp. 2-4, 2013.
- [4] P. T. Duc, V. V. M. Nhat, D. T. Chuong, "A Model of Traffic Prediction based Admission Control in OBS Nodes.", accepted RIVF Conference, 20 March 2019.
- [5] K. Salah and F. Haidari, "On the performance of a simple packet rate estimator," *AICCSA 08 - 6th IEEE/ACS Int. Conf. Comput. Syst. Appl.*, pp. 392-395, 2008.
- [6] V. Minh, N. Vo, V. H. Le, and H. S. Nguyen, "A model of optimal burst assembly for delay reduction at ingress OBS nodes," pp. 3970-3982, 2017.
- [7] C. McArdle, D. Tafani, and L. P. Barry, "Analysis of a buffered optical switch with general interarrival times," *J. Networks*, vol. 6, no. 4, pp. 536-548, 2011.
- [8] M. Nandi, A. K. Turuk, D. K. Puthal, and S. Dutta, "Best Fit Void Filling Algorithm in Optical Burst Switching Networks," pp. 609-614, 2009.

AN IMPROVED ABOUT RATE PREDICTION BASED SCHEDULING ADMISSION CONTROL COMBINED FDL DELAY LINE

Pham Trung Duc, Vo Viet Minh Nhat, Dang Thanh Chuong

ABSTRACT: *The issue of scheduling admission control researched in [1][2][3][4]. The authors in [4] is proposed a model of traffic prediction based scheduling admission control, in order to help the high priority burst take more resources. However this method make burst loss rate of low priority high. This article will propose a combination line delay FDL in scheduling admission control to improved low priority burst loss rate. The result of analysis and simulation show that low priority burst lost rate improved significantly when using delay line and this usage is more effective when many different delay lines are used.*