

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

NGÔ THỊ BẾN

**CÁC KẾ HOẠCH QUẢN LÝ HÀNG ĐỘNG BLUE
CHO TRUYỀN THÔNG ĐA PHƯƠNG TIỆN**

LUẬN VĂN THẠC SĨ NGÀNH CÔNG NGHỆ THÔNG TIN

HÀ NỘI – 2016

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

NGÔ THỊ BẾN

**CÁC KẾ HOẠCH QUẢN LÝ HÀNG ĐỢI ĐỘNG BLUE
CHO TRUYỀN THÔNG ĐA PHƯƠNG TIỆN**

Ngành: Công nghệ thông tin

Chuyên ngành: Truyền dữ liệu và Mạng máy tính

Mã số: Chuyên ngành đào tạo thí điểm

LUẬN VĂN THẠC SĨ NGÀNH CÔNG NGHỆ THÔNG TIN

NGƯỜI HƯỚNG DẪN KHOA HỌC: PGS.TS. Nguyễn Đình Việt

HÀ NỘI – 2016

MỞ ĐẦU

Internet là một hệ thống kết nối mạng toàn cầu đảm bảo liên thông giữa các hệ thống máy tính và thiết bị trên diện rộng. Internet ngày càng phát triển không chỉ về số lượng kết nối mà còn sự đa dạng của các lớp ứng dụng, các dữ liệu được truyền đi không chỉ đơn thuần là dạng văn bản đơn giản, mà là dữ liệu đa phương tiện bao gồm cả âm thanh, hình ảnh tĩnh, động, ... Các ứng dụng đa phương tiện phổ biến có thể kể đến như điện thoại qua mạng (Internet telephony), hội thảo trực tuyến (video conferencing), xem video theo yêu cầu (video on demand),... đang ngày càng được sử dụng rộng rãi. Có thể nói mạng Internet đã làm một cuộc cách mạng thay đổi nhiều khía cạnh trong cuộc sống của chúng ta. Hơn nữa mạng Internet còn rẻ hơn nhiều so với các loại hình dịch vụ khác, do đó nó được sử dụng rộng khắp ở mọi quốc gia trên thế giới. Cùng với sự bùng nổ về nhu cầu sử dụng Internet và sự gia tăng của lưu lượng thông tin kéo theo vấn đề xảy ra tắc nghẽn trên Internet là không thể tránh khỏi. Vì vậy, để đảm bảo thông suốt đường truyền, kiểm soát tắc nghẽn tại nút mạng đóng một vai trò rất quan trọng cho Internet hoạt động hiệu quả và tin cậy với người sử dụng. Mô hình cung cấp dịch vụ kiểu cố gắng tối đa (best-effort) của Internet truyền thống đã không đủ để đáp ứng với những yêu cầu về chất lượng dịch vụ (QoS) khi có sự bùng nổ các luồng dữ liệu tham gia mạng và làm cho các nút mạng trung tâm nhanh chóng bị tắc nghẽn. Khi mạng IP ra đời đã thoả mãn được các yêu cầu cả về kỹ thuật lẫn chất lượng dịch vụ. Tuy nhiên để nâng cao chất lượng dịch vụ, đáp ứng được các yêu cầu của người sử dụng là một vấn đề thực sự khó khăn cho các nhà quản lý mạng, đặc biệt là trong hoàn cảnh hiện nay khi các luồng thông tin ngày càng đa dạng về chủng loại, đặc tính, mà yêu cầu chất lượng sử dụng thông tin thì ngày càng khắt khe. Việc yêu cầu chất lượng dịch vụ của người sử dụng cũng tạo ra sự cạnh tranh khốc liệt giữa các nhà cung cấp dịch vụ, yêu cầu các nhà cung cấp dịch vụ phải tìm ra các giải pháp mới để nâng cao chất lượng dịch vụ và tăng doanh thu cho mình.

Có nhiều hướng nghiên cứu để cải thiện chất lượng truyền dữ liệu đa phương tiện qua mạng, nhưng đối với một số kỹ thuật truyền thống chúng có hạn chế là gây ra gánh nặng truyền tải đối với âm thanh/hình ảnh video, do đó làm tiêu tốn thêm tài nguyên băng thông mạng. Khi có quá nhiều gói tin được đưa vào mạng, sẽ làm cho hiệu năng của mạng giảm đi vì các nút mạng không còn đủ khả năng lưu trữ, xử lý, truyền đi, chúng bắt đầu bị mất các gói tin dẫn đến sự tắc nghẽn trong mạng máy tính. Các nhà xây dựng mạng đã khéo léo đưa ra các mô hình mạng mới như mô hình mạng dịch vụ phân loại DiffServ và mạng dịch vụ tích hợp IntServ đồng thời kết hợp các mô hình mạng với nhau để lợi dụng ưu điểm của từng mạng và hạn chế nhược điểm của chúng. Bên cạnh đó các nhà thiết kế còn đi sâu vào tìm hiểu và thiết kế các phương pháp quản lý, giám sát các tiến trình truyền tin ngay bên trong bản thân của các thành phần nhỏ của mạng như router, chuyển mạch,... Điển hình các router được thiết kế theo cấu trúc CQS đã phần nào đơn giản hoá việc truyền tin và nâng cao chất lượng dịch vụ. Một trong những phương pháp đưa ra ở các router để cải thiện chất lượng dịch vụ trong mạng IP thông dụng nhất là phương pháp quản lý hàng đợi tích cực AQM [9, 10, 17]. Đặc trưng của các hàng đợi AQM là điều chỉnh xác suất đánh dấu hoặc loại bỏ gói tin tại các bộ đệm router để ngăn ngừa hiện tượng tắc nghẽn xảy ra. Mục tiêu chính của luận văn là tập trung nghiên cứu và đánh giá các kế hoạch quản lý hàng đợi động cho truyền thông đa phương tiện, nhằm đảm bảo chất lượng dịch vụ QoS. Tập trung nghiên cứu chiến lược BLUE, đánh giá và so sánh chiến lược quản lý hàng đợi BLUE với các chiến lược hàng đợi RED, A-RED. Ngoài việc đưa ra các chiến lược quản lý hàng đợi thích hợp thì một giải thuật cho phép quản lý kiểm soát tắc nghẽn dựa trên sự kiện mất gói dữ liệu và mức độ sử dụng đường truyền thay vì chiếm dụng hàng đợi sẽ đem lại hiệu quả cao nếu được áp dụng vào từng trường hợp cụ thể, xử lý một cách tối ưu việc vận chuyển thông tin trong mạng.

1. Tình hình nghiên cứu trong nước, nước ngoài về AQM

Internet là mạng kết nối mở lớn nhất trên thế giới, mạng của các mạng. Sự phát triển nhanh chóng của Internet dẫn đến Internet phải đối mặt với sự bùng nổ về số lượng máy tính kết nối và sự đa dạng của các lớp ứng dụng triển khai trên nó. Ngày nay khi cơ sở hạ tầng của mạng Internet được

nâng cao, đặc biệt là về băng thông, khả năng lưu trữ và xử lý của các máy chủ (servers) đã làm cho nhu cầu của các ứng dụng đa phương tiện qua mạng tăng lên nhanh chóng, các dịch vụ trên Internet không ngừng phát triển, xuất hiện trong mọi lĩnh vực như thương mại, chính trị, quân sự, nghiên cứu, giáo dục, văn hoá, xã hội...

Vì lưu lượng trên Internet có đặc tính bùng nổ nên hàng đợi (bộ đệm) tại các nút mạng (router) phải có kích thước đủ lớn, để đảm bảo cho các nút thực hiện chức năng store-and-forward một cách hiệu quả. Tuy nhiên, nếu thi hành chính sách phục vụ tại hàng đợi kiểu FIFO (Tail-Drop Queue) thì hàng đợi sẽ thường xuyên ở trạng thái đầy, làm tăng đáng kể thời gian trễ trung bình của các gói tin trong mạng. Do vậy, điều quan trọng là phải có các kỹ thuật để đảm bảo cho mạng đạt được thông lượng cao và thời gian trễ trung bình nhỏ. Quản lý hàng đợi tích cực AQM (Active Queue Management) là một trong các giải pháp quan trọng và hiệu quả cho điều khiển tránh tắc nghẽn trên Internet [17,21].

Thông thường có hai phương án để kiểm soát tránh tắc nghẽn là tăng hiệu suất các thiết bị phần cứng và dùng kỹ thuật phần mềm. Việc tăng hiệu suất các thiết bị là cần thiết, nhưng lại khá tốn kém, khó đồng bộ và hiệu quả chưa cao. Ngược lại, dùng kỹ thuật phần mềm để kiểm soát tắc nghẽn đã đem lại hiệu quả rất lớn. Trong kỹ thuật này có hai phương pháp được quan tâm và phát triển, đó là: cải tiến các giao thức điều khiển truyền thông và nâng cao các kỹ thuật quản lý hàng đợi tích cực AQM tại các nút mạng. Việc tăng hiệu năng của giao thức TCP thông qua các biến thể đã triển khai trên Internet và đã đem lại hiệu quả rất lớn. Tuy nhiên, do sự đa chuẩn của các loại mạng, sự phong phú các thiết bị kết nối và sự phức tạp các ứng dụng truyền thông nên điều quan trọng là cần có những cơ chế quản lý hàng đợi tích cực tại các nút mạng để hỗ trợ điều tiết lưu thông trên mạng, nhằm tránh và giải quyết tắc nghẽn.

Quản lý hàng đợi là một nhóm tổ hợp các phương pháp quản lý bộ đệm, đây là một trong những cơ chế cung cấp chất lượng dịch vụ (QoS). Quản lý hàng đợi quyết định việc phân phối bộ đệm và loại bỏ các gói đến theo một chính sách được quyết định trước.

Trong những năm gần đây, vấn đề nghiên cứu về chiến lược quản lý hàng đợi tích cực AQM trong mạng Internet đã phát triển mạnh mẽ và sôi động. Ở trong nước và nhiều nước trên thế giới cũng đã có nhiều công trình nghiên cứu tập trung vào nghiên cứu cải tiến các giao thức điều khiển từ đầu cuối đến đầu cuối (end-to-end) nhằm nâng cao hiệu năng của giao thức TCP, như: TCP NewReno, Vegas và các phương pháp quản lý hàng đợi tích cực, như: RED [12,13,14,26], ARED [28], ARIO [25], BLUE [13,24]... tại các nút mạng trung tâm. Thông qua các cơ chế đó, mỗi nút mạng đã kiểm soát được số lượng lớn các gói dữ liệu đến đồng thời trong hàng đợi của bộ định tuyến. Kết quả của những công trình nghiên cứu đã tập trung nghiên cứu một số giải pháp giải quyết vấn đề tránh tắc nghẽn duy trì tính ổn định của chất lượng mạng và hướng đến việc bảo đảm QoS trong môi trường mạng có mật độ gói tin dày đặc. Việc bảo đảm chất lượng dịch vụ liên quan mật thiết đến việc phân chia tài nguyên mạng (băng thông, bộ đệm). Tại mỗi nút mạng, việc phân chia băng thông, bộ đệm được thực hiện bằng bộ định trình lưu lượng theo một cơ chế định trình nhất định. Chất lượng dịch vụ toàn trình của mỗi ứng dụng phụ thuộc vào chất lượng dịch vụ tại mỗi nút mạng, và phụ thuộc vào gói tin của bộ định trình, thời gian gói tin bị trễ trong bộ đệm, khả năng mất gói tin do tràn bộ đệm. Có nhiều các kết quả khả thi từ việc nghiên cứu tăng cường khả năng bảo đảm chất lượng dịch vụ trong mạng IP nhằm ngăn ngừa hiện tượng tắc nghẽn xảy ra. Tuy nhiên qua khảo sát các công trình nghiên cứu trong và ngoài nước cho thấy các giải thuật AQM vẫn còn hạn chế khi ứng dụng đòi hỏi đáp ứng thời gian thực như truyền phát video trên mạng. Do đó việc đảm bảo QoS và đáp ứng yêu cầu nêu trên và kết hợp các cơ chế nhằm đem lại hiệu quả cao nhất trong môi trường mạng phức tạp như hiện nay.

2. Mục tiêu, kết quả cần đạt được của luận văn

Mục tiêu chính của Luận văn là tập trung nghiên cứu và đánh giá hiệu suất của thuật toán quản lý hàng đợi BLUE - một chiến lược điển hình của thuật toán quản lý hàng đợi tích cực dựa vào tải nạp.

Sau đó so sánh chiến lược này với các chiến lược quản lý hàng đợi khác như RED, A-RED, ARIO từ đó có những đánh giá, đưa ra các kết quả so sánh hiệu năng giữa các mô hình dựa trên các kết quả mô phỏng trên NS-2. Ngoài ra, vì mục đích cuối cùng là phải hướng tới người sử dụng, nên chúng tôi cũng đã dành một chương để trình bày tổng quan về truyền thông đa phương tiện trên mạng, đây là các dịch vụ ở mức ứng dụng, hiệu quả của nó phụ thuộc chặt chẽ vào các dịch vụ mức dưới.

Kết quả cần đạt được của luận văn: Nghiên cứu thuật toán RED, ARED, ARIO BLUE, tập trung nghiên cứu chiến lược quản lý hàng đợi BLUE. So sánh chiến lược này với các chiến lược quản lý hàng đợi khác từ đó có những đánh giá, đưa ra các kết quả so sánh hiệu năng giữa các mô hình.

3. Đối tượng và phạm vi nghiên cứu

Đề tài tập trung nghiên cứu lý thuyết về truyền thông đa phương tiện và các yêu cầu bảo đảm QoS đồng thời nghiên cứu một số chiến lược quản lý hàng đợi động, hiệu quả tại gateway, đi sâu nghiên cứu về BLUE – Một chiến lược quản lý hàng đợi dựa vào tải nạp, có thể được cài đặt để hỗ trợ Internet hoạt động hiệu quả hơn..

Đề tài sử dụng bộ công cụ mô phỏng mạng NS2 để nghiên cứu sâu về BLUE và đánh giá, so sánh hiệu suất của nó với các chiến lược quản lý hàng đợi RED, ARED.

4. Phương pháp nghiên cứu

Để đạt được các mục tiêu trên, phương pháp nghiên cứu trong luận văn được kết hợp chặt chẽ giữa nghiên cứu lý thuyết với cài đặt mô phỏng kiểm chứng. Về lý thuyết, luận văn nghiên cứu, khảo sát các công trình liên quan để tìm những tồn tại, lựa chọn những vấn đề sẽ giải quyết. Hệ thống những vấn đề cần giải quyết, đề xuất mô hình lý thuyết, sử dụng những công cụ hỗ trợ để phân tích. Luận văn thực hiện mô phỏng bằng phần mềm mô phỏng mạng NS2 (Network Simulator) được các nhà nghiên cứu khoa học tin dùng.

5. Bố cục của luận văn

Luận văn gồm phần mở đầu, 4 chương nội dung, kết luận. Cụ thể nội dung của các chương trong luận văn được trình bày như sau:

Chương 1: Trình bày về truyền thông đa phương tiện và các yêu cầu chất lượng dịch vụ QoS và các phương pháp đảm bảo chất lượng dịch vụ trong truyền thông đa phương tiện trên mạng.

Chương 2: Trình bày tổng quan về các chiến lược quản lý hàng đợi động AQM, tìm hiểu hai thuật toán tiêu biểu của AQM: RED, A-RED

Chương 3. Tập trung nghiên cứu sâu về chiến lược quản lý hàng đợi dựa vào tải nạp BLUE và đề xuất cải tiến giải thuật quản lý hàng đợi BLUE

Chương 4: Dựa trên bộ mô phỏng mạng NS để kiểm chứng các đánh giá hiệu suất đồng thời so sánh hiệu suất của chiến lược BLUE với các chiến lược quản lý hàng đợi khác: RED, A-RED

Phần kết luận nêu những kết quả chính của luận văn và hướng phát triển tiếp theo.

Chương 1. TỔNG QUAN VỀ TRUYỀN THÔNG ĐA PHƯƠNG TIỆN VÀ CÁC YÊU CẦU CHẤT LƯỢNG DỊCH VỤ

1.1. Các khái niệm cơ bản

1.1.1. Hệ thống truyền thông đa phương tiện

❖ Các khái niệm

- Media:

Media là phương tiện truyền đạt thông tin, đề cập đến các loại thông tin hay loại biểu diễn thông tin như dữ liệu văn bản, ảnh, âm thanh và video.

- Dữ liệu Multimedia:

Dữ liệu multimedia được chia thành hai lớp là các dữ liệu liên tục và các dữ liệu không liên tục. Các dữ liệu liên tục bao gồm các dữ liệu âm thanh, video thay đổi theo thời gian. Các dữ liệu không liên tục là các dữ liệu không phụ thuộc vào thời gian, các loại dữ liệu đặc trưng cho dạng này là các dữ liệu văn bản (có hoặc không có định dạng), hình ảnh tĩnh và các đối tượng đồ họa. Dữ liệu multimedia là dữ liệu ở các dạng thông tin khác nhau.

- Hệ thống truyền thông đa phương tiện:

Hệ thống truyền thông đa phương tiện (Multimedia Communication System) là hệ thống cung cấp tích hợp các chức năng lưu trữ, truyền dẫn và trình diễn các kiểu phương tiện mang tin rời rạc (văn bản, hình ảnh, đồ họa...) và liên tục (audio, video) trong một môi trường thông tin số.

Yêu cầu của truyền thông đa phương tiện:

- + Băng thông đủ lớn
- + Có khả năng phân chia lưu lượng cho từng loại dữ liệu, từng loại dịch vụ.
- + Có chính sách QoS với từng loại dữ liệu
- + Khả năng thích ứng với nhiều thiết bị người dùng
- + Khả năng quản lý tốt, dễ dàng mở rộng, nâng cấp

1.1.2. Hệ thống thời gian thực

Hệ thống thời gian thực - RTS (Real-Time System) là hệ thống mà trong đó sự đúng đắn của việc thực hiện các thao tác không chỉ phụ thuộc vào việc thu được kết quả đúng mà còn phải đưa ra kết quả đúng thời điểm. RTS khác biệt với các hệ thống khác ở tính quan trọng của thời điểm cho ra kết quả, điều đó có nghĩa là tính đúng đắn của hệ thống thời gian thực không chỉ phụ thuộc vào kết quả logic của thao tác mà còn phụ thuộc vào thời điểm tạo ra các kết quả.

Trong hệ thống thời gian thực chúng có các đặc điểm sau:

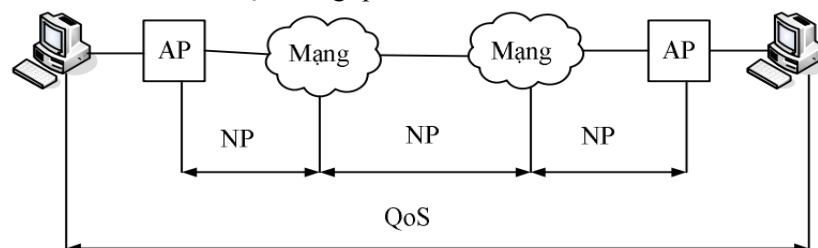
- + Các sự kiện bên trong và bên ngoài có thể xảy ra một cách định kỳ hoặc tự phát.
- + Sự đúng đắn của hệ thống còn phụ thuộc cả vào việc đáp ứng các ràng buộc thời gian.

1.1.3. Chất lượng dịch vụ QoS

1.1.3.1. Khái niệm QoS

Theo khuyến nghị E800 ITU-T, chất lượng dịch vụ là “Một tập các khía cạnh của hiệu năng dịch vụ nhằm xác định cấp độ thỏa mãn của người sử dụng đối với dịch vụ”. Như vậy QoS được xác định bằng các chỉ tiêu định tính và định lượng. Chỉ tiêu định tính thể hiện sự cảm nhận của khách hàng còn chỉ tiêu định lượng được thực hiện bằng các số đo cụ thể.

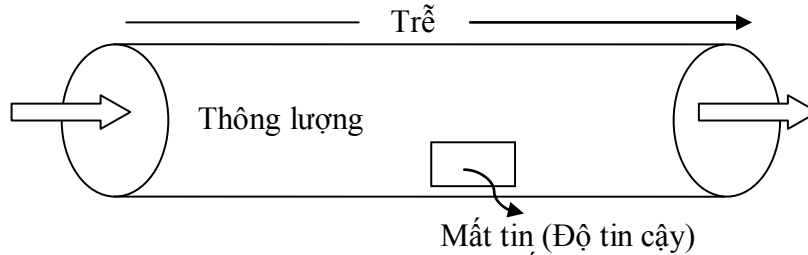
Dưới đây biểu diễn một mô hình QoS tổng quát:



Hình 1.1. Mô hình QoS tổng quát

1.1.3.2. Các tham số chính của QoS

Chất lượng dịch vụ bao gồm các tham số kỹ thuật như: độ trễ, thông lượng, tỷ số mất tin, jitter có thể được minh họa bằng Hình 1.1 dưới đây.



Hình 1.2. Các tham số QoS chính

1.1.3.3. Các mức QoS

Có ba mức dịch vụ:

- Dịch vụ cố gắng tối đa (Best-Effort Service)
- Dịch vụ phân loại (Differentiated Service)
- Dịch vụ đảm bảo (Guaranteed service)

1.1.3.4. Đảm bảo chất lượng dịch vụ (QoS) trong truyền thông đa phương tiện

Do tính đa dạng của dịch vụ, ứng dụng trên mạng, các yêu cầu về đảm bảo QoS cho các ứng dụng cũng hết sức đa dạng. Đối với truyền thông đa phương tiện về chất lượng dịch vụ có thể được phân thành các loại: chất lượng qua cảm nhận (nghe, nhìn) của người sử dụng, chất lượng dịch vụ của ứng dụng hay chất lượng dịch vụ truyền dữ liệu qua mạng.

1.2. Các ứng dụng đa phương tiện trên mạng Internet

1.2.1. Truyền video và audio đã được lưu trữ

1.2.2. Phát sóng trực tiếp của audio và video

1.2.3. Ứng dụng audio, video tương tác thời gian thực

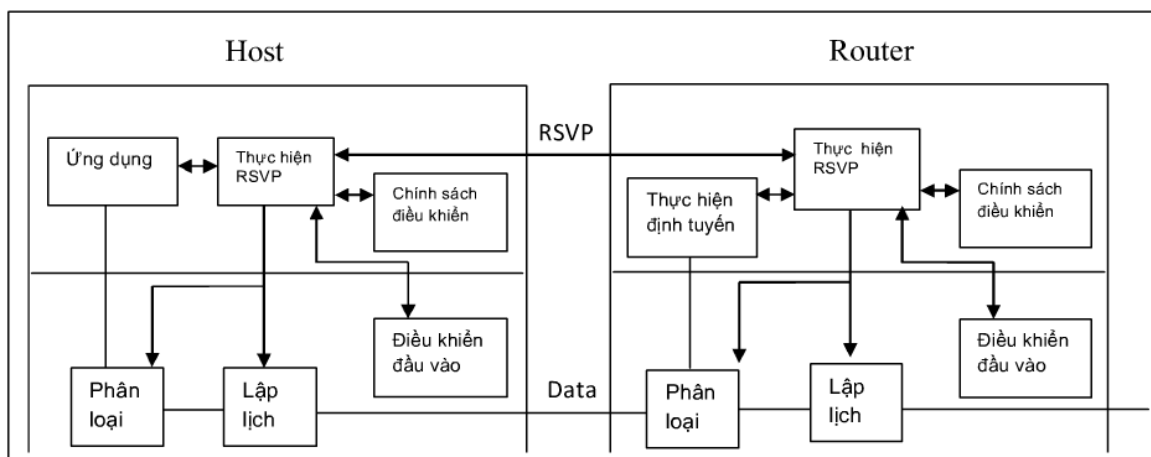
1.3. Các mô hình đảm bảo QoS cho truyền thông đa phương tiện

1.3.1. Mô hình dịch vụ tích hợp - IntServ

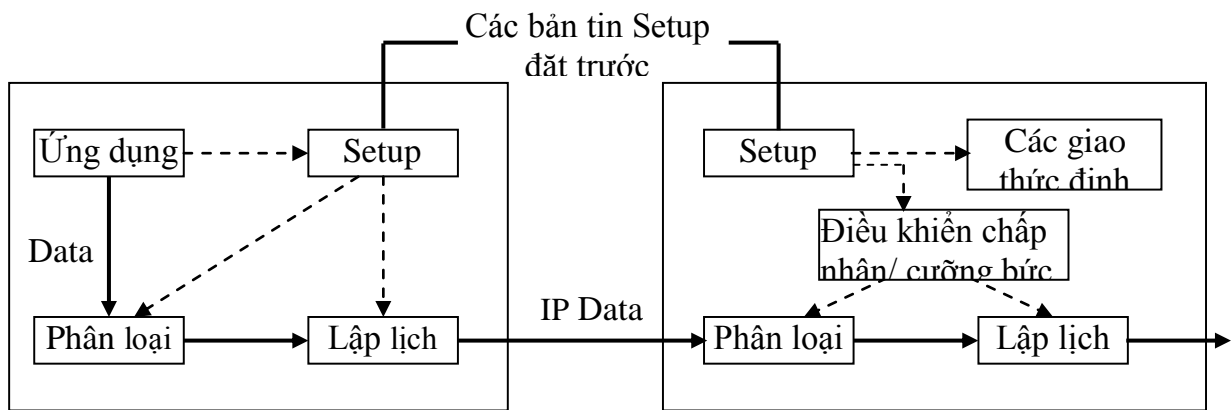
Mô hình IntServ được đưa ra bởi nhóm làm việc tại IETF với mục đích hỗ trợ chất lượng dịch vụ cho các ứng dụng từ đầu cuối tới đầu cuối. Mô hình này không những đáp ứng được các dịch vụ Best-Effort mà các dịch vụ thời gian thực cũng được thực thi qua mô hình này qua việc hỗ trợ chức năng dành trước băng thông trên Internet và các mạng tương tác. Các ứng dụng sẽ nhận được băng thông đúng yêu cầu và truyền đi trong mạng với độ trễ cho phép.

➤ Nguyên lý hoạt động của mô hình tích hợp dịch vụ

IntServ sử dụng một giao thức đặc biệt RSVP để dành trước băng thông xác định trong mỗi bộ định tuyến dọc theo đường đi từ nguồn đến đích.



Hình 1.3. Nguyên lý hoạt động của mô hình dịch vụ tích hợp IntServ



Hình 1.4. Mô hình dịch vụ tích hợp IntServ

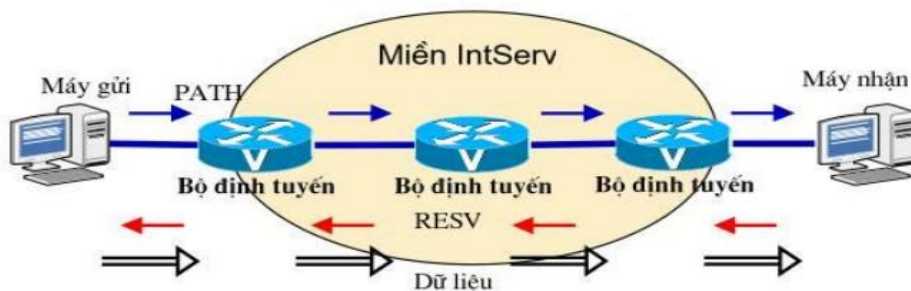
Giao thức dành trước tài nguyên(RSVP)

Giao thức dành tài nguyên RVSP (Resource Reservation Protocol) được sử dụng bởi IntServ được đặc tả trong RFC2205, các dịch vụ GS và CLS được mô tả trong RFC2210. RSVP có thể gửi yêu cầu đặt trước tài nguyên và đáp ứng tương ứng của thành phần chấp nhận luồng từ máy tính tới router, từ router tới router và từ router tới máy đích (hoặc nhiều một máy).

➤ Nguyên lý hoạt động của RSVP:

Một phiên làm việc của RSVP thường được xác định bởi 3 tham số sau:

- + Địa chỉ đích
- + Nhận dạng giao thức
- + Địa chỉ cổng đích



Hình 1.5. Nguyên lý hoạt động của giao thức dành trước tài nguyên RSVP

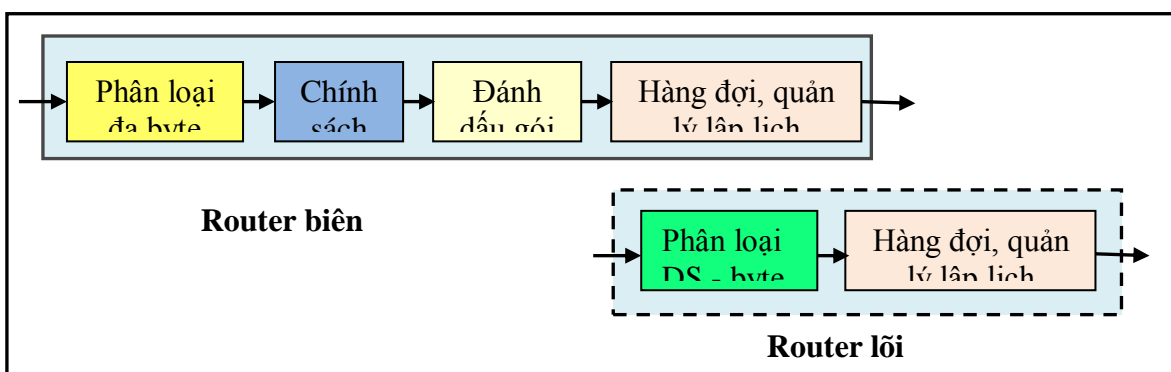
1.3.2. Mô hình dịch vụ phân loại - DiffServ

Mô hình dịch vụ phân loại được phát triển nhằm mục đích cung cấp các lớp dịch vụ khác nhau cho các lưu lượng trên Internet và nhằm đạt được tính linh động trong quá trình truyền thông.

Kiến trúc Diffserv bao gồm hai tập các thành phần chức năng:

- Tại biên của mạng, việc phân loại và điều khiển lưu lượng được thực hiện và các gói được phân vào các lớp.
- Tại lõi, một cơ chế phân loại đơn giản được thực hiện. Cơ chế hàng đợi dựa trên lớp được áp dụng.

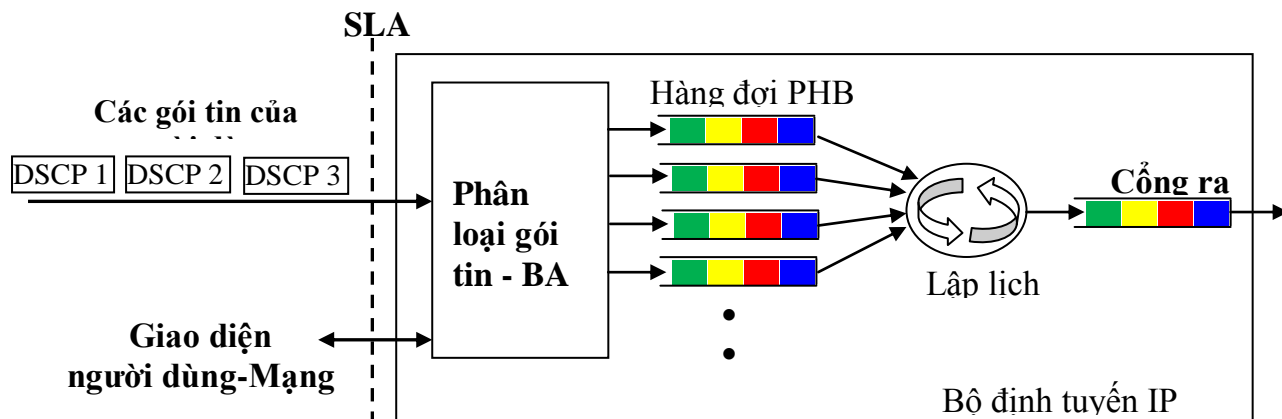
Sơ đồ khối kiến trúc DiffServ được mô tả cụ thể như sau:



Hình 1.6. Xử lý gói trong mô hình DiffServ

➤ Nguyên lý hoạt động của mô hình dịch vụ phân loại

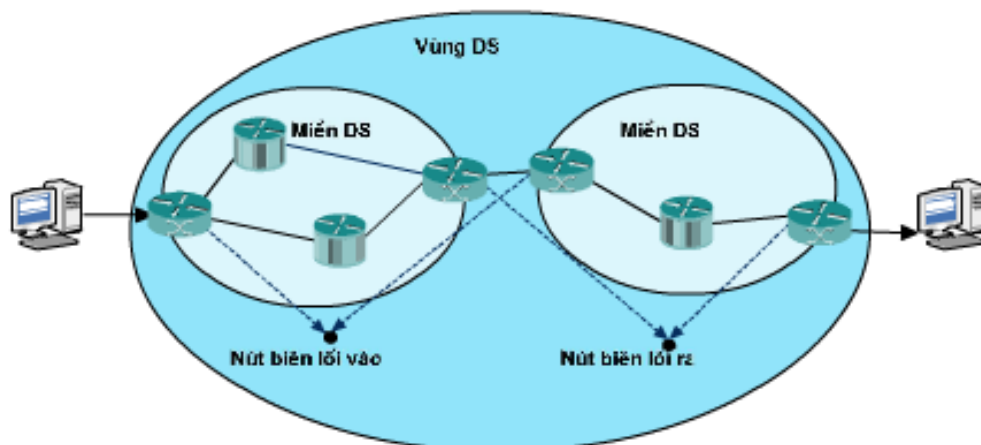
Hình vẽ dưới đây mô tả các bước cơ bản trong việc cung cấp các dịch vụ DiffServ.



Hình 1.7. Mô hình các bước dịch vụ phân loại DiffServ

1.3.2.1. Miền dịch vụ phân loại và điểm mã dịch vụ phân loại

Một miền dịch vụ phân loại – DS (Differentiated Service) gồm các nút DS (còn gọi là các bộ định tuyến hỗ trợ cơ chế dịch vụ phân loại) hoạt động với một chính sách cung cấp dịch vụ chung và thiết lập các nhóm hành vi theo chặng – PHB (Per-hop Behavior) được thực hiện trên mỗi nút.



Hình 1.8. Miền dịch vụ phân biệt DS

Các vùng DS có khả năng hỗ trợ các miền DS dọc theo đường định tuyến nối các miền trong vùng.

1.3.2.2. Hành vi theo chặng PHB (Per-hop Behavior)

Kiến trúc DiffServ định nghĩa hành vi theo chặng PHB cho việc xử lý chuyển tiếp gói tin tại mỗi node mạng áp dụng kết hợp hành vi BA. PHB liên quan đến các đặc tính về chất lượng dịch vụ như độ trễ, độ biến thiên độ trễ hay mất gói của gói tin khi đi qua node dịch vụ DiffServ.

- PHB chuyển tiếp nhanh - EF (Expedited Forwarding)
- PHB chuyển tiếp đảm bảo - AF (Assured Forwarding)
- PHB lựa chọn lớp - CS (Class Selectors)
- PHB mặc định (Default PHB)

Chương 2. CÁC CHIẾN LƯỢC QUẢN LÝ HÀNG ĐỢI ĐỘNG AQM

Quản lý hàng đợi tích cực là một trong các giải pháp cho điều khiển tránh tắc nghẽn đảm bảo truyền thông liên tục và hiệu quả trên mạng Internet. Có rất nhiều thuật toán được đưa ra trong kỹ thuật quản lý hàng đợi như các thuật toán lập lịch hay các thuật toán quản lý bộ đệm. Nội dung chính của chương này là trình bày về chiến lược quản lý hàng đợi động (AQM) và một số chiến lược quản lý hàng đợi tích cực dựa theo kích thước hàng đợi và dựa vào tải nạp.

2.1. Cách tiếp cận truyền thống và hiệu quả

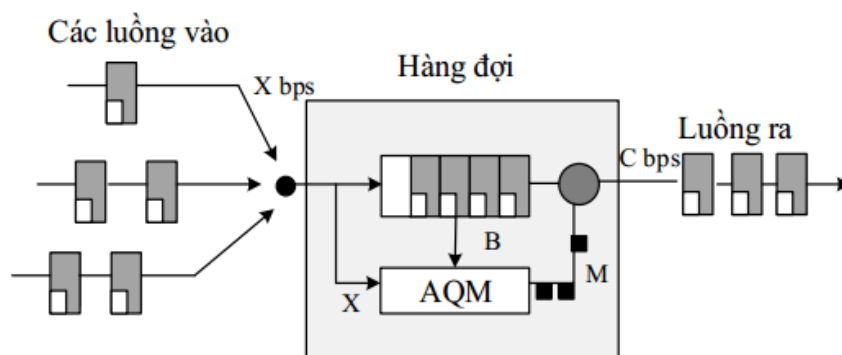
Kỹ thuật truyền thống và là kỹ thuật đơn giản nhất để quản lý kích thước hàng đợi là dựa vào cơ chế FIFO. Theo cơ chế này, tất cả các gói tin đến được xếp vào hàng đợi; khi hàng đợi đầy thì các gói tin đến sau đều bị loại bỏ; để chọn các gói tin truyền đi thì gói tin nào đến trước được phục vụ trước. Trong bộ mô phỏng NS, kỹ thuật này được cài đặt với tên gọi “DropTail”. Do tính đơn giản của nó, kỹ thuật này được sử dụng nhiều năm trên Internet, tuy nhiên nó có 2 nhược điểm cơ bản được trình bày dưới đây.

2.1.1. Hiện tượng Lock-Out và Global Synchronization

2.1.2. Hiện tượng Full Queues

2.2. Chiến lược quản lý hàng đợi động AQM

AQM là phương pháp chủ động thông báo với bên gửi khi mới bắt đầu có tắc nghẽn, trước khi xảy ra tràn bộ đệm. Bằng cách sử dụng cơ chế AQM, bên gửi được thông báo sớm về tắc nghẽn và có thể phản ứng phù hợp. Hình 2.1 trình bày mô hình quản lý hàng đợi tích cực trong mạng TCP/IP [27].



Hình 2.1. Mô hình quản lý hàng đợi tích cực

Nhìn chung, các chiến lược AQM đem lại các lợi ích sau:

2.2.1. Giảm số gói tin bị loại bỏ tại router

2.2.2. Giảm độ trễ

2.2.3. Tránh hiện tượng Lock-Out

2.3. Chiến lược RED

Thuật toán RED lần đầu tiên được đề xuất vào năm 1993 bởi Sally Floyd và Van Jacobson cho chức năng quản lý hàng đợi tích cực (AQM), sau đó nó được chuẩn hoá lại theo yêu cầu của IETF [14]. RED có khả năng chống hiện tượng đồng bộ toàn cục của các luồng TCP, duy trì khả năng đạt thông lượng qua hàng đợi RED cao cũng như độ trễ thấp cùng với cách đối xử công bằng giữa các kết nối TCP đi qua hàng đợi. RED gateway thực hiện loại bỏ gói tin trong hàng đợi theo chiến lược AQM, ngoài ra nó còn đánh dấu vào trường ECN trong gói tin TCP, để báo cho bên gửi biết về hiện tượng tắc nghẽn có dấu hiệu sắp xảy ra, cần có phản ứng tích cực (việc đánh dấu ECN là một tùy chọn của RED). Khi có dấu hiệu của tắc nghẽn xảy ra trong mạng, các bộ đệm của router được điền đầy và router bắt đầu loại bỏ các gói. Có hai vấn đề nan giải trong mạng: thứ nhất các gói bị mất sẽ phải được truyền lại, việc này làm tăng tải trong mạng đồng thời phát sinh ra trễ các luồng lưu lượng. Một vấn đề thứ hai là xảy ra hiện tượng đồng bộ toàn cục trên tất cả các luồng. Với một sự bùng nổ lưu lượng dạng bó (burst), các hàng đợi được điền đầy và các gói đến sau bị loại bỏ. Kết quả là kết nối nhiều kết nối TCP bị ảnh hưởng (mất gói tin) và chuyển sang chế độ khởi đầu chậm. Việc có nhiều kết nối TCP cùng chuyển sang chế độ khởi đầu chậm tại một thời điểm và cùng thoát khỏi chế độ này do đó sẽ gây ra thêm các bó lưu lượng lớn.

- Mục đích thiết kế thuật toán RED
- + Tránh tắc nghẽn
- + Tránh đồng bộ toàn cục
- + Giữ ổn định kích thước hàng đợi trung bình

2.3.1. Nguyên tắc hoạt động

Để đạt được những mục đích thiết kế của thuật toán như đã trình bày ở phần trên thì RED gateways phải thực hiện các công việc sau:

- Phát hiện sớm tắc nghẽn, bằng cách thường xuyên giám sát kích thước hàng đợi. Tránh tắc nghẽn bằng cách loại bỏ các gói tin trong hàng đợi theo một số quy tắc nhất định, để giữ cho kích thước hàng đợi trung bình đủ nhỏ, làm cho mạng hoạt động ở vùng có độ trễ thấp và thông lượng cao, trong khi vẫn cho phép kích thước hàng đợi dao động trong một miền nhất định để hấp thụ các thăng giáng lưu lượng ngắn hạn.

- Báo hiệu sớm tắc nghẽn tới nguồn phát. Khi có dấu hiệu của tắc nghẽn, ngoài việc dựa trên biện pháp loại bỏ ngẫu nhiên các gói tin nêu trên, cần áp dụng biện pháp đánh dấu vào trường ECN của gói tin với một xác suất nhất định. Các gói tin này được lựa chọn ngẫu nhiên để cho phép truyền đi cùng với dấu hiệu tắc nghẽn được đánh dấu để thông báo cho thực thể gửi TCP biết nhằm giảm lưu lượng đưa vào mạng (thông tin ECN được bên đích gửi cho bên nguồn trong gói tin ACK). Việc đánh dấu được thực hiện ngẫu nhiên để tránh hiện tượng đồng bộ toàn cục và không chống lại các dòng lưu lượng có giá trị trung bình thấp nhưng độ thăng giáng cao.

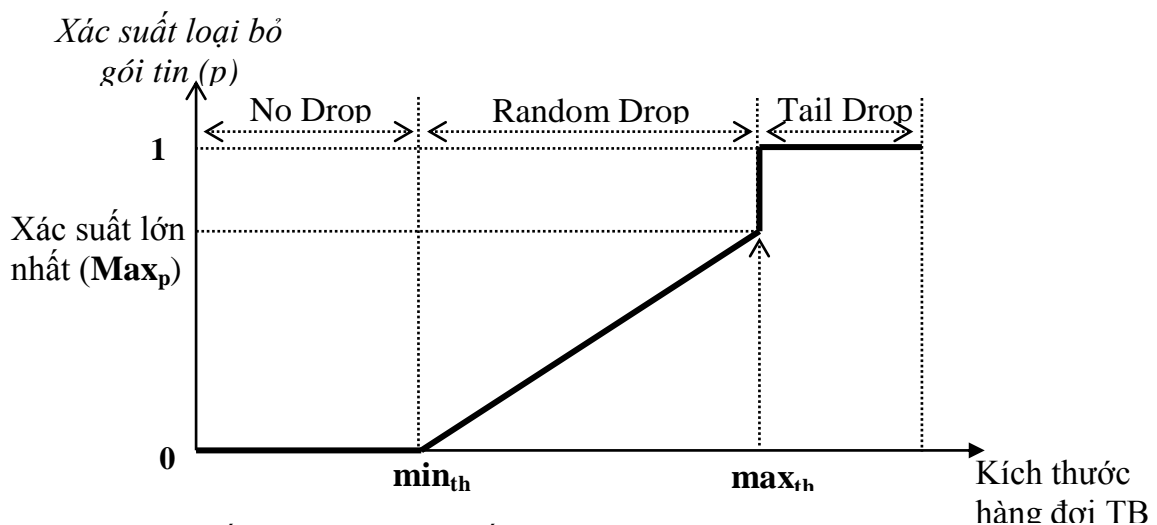
2.3.2. Giải thuật RED

RED sẽ tính toán kích thước hàng đợi trung bình dựa trên bộ lọc thông thấp (Low-Pass Filter), giá trị trung bình này còn được gọi là trung bình dịch chuyển có trọng số tăng theo hàm mũ - EWMA (Exponential Weighted Moving Average). Kích thước hàng đợi trung bình được so sánh với hai giá trị ngưỡng: ngưỡng dưới min_{th} và ngưỡng trên max_{th} để quyết định việc đánh dấu hoặc loại bỏ các gói tin trong hàng đợi. Hoạt động của RED được mô tả bởi ba quy tắc để xác định vị trí của mỗi gói tin gửi đến:

- + Khi kích thước hàng đợi trung bình nhỏ hơn ngưỡng dưới thì không có gói tin nào bị đánh dấu hay loại bỏ (hay gán xác suất loại bỏ gói bằng 0). Đây là trường hợp hoạt động bình thường.

- + Khi kích thước hàng đợi trung bình lớn hơn ngưỡng trên thì tất cả các gói đến đều bị loại bỏ. Khi các gói bị loại bỏ hoặc nếu tất cả các nguồn cùng hợp tác với nhau thì kích thước hàng đợi trung bình sẽ không vượt quá giá trị ngưỡng trên.

- + Khi kích thước hàng đợi trung bình biến thiên từ min_{th} đến max_{th} thì mỗi gói tin đến được đánh dấu hoặc loại bỏ một cách ngẫu nhiên tùy theo một hàm xác suất p .



Hình 2.2. Mối quan hệ giữa xác suất loại bỏ gói và kích thước hàng đợi trung bình.
Giải thuật tổng quát của RED gateway được mô tả như ở hình 2.3.

```

For each packet arrival
  Caculate the average queue size  $avg$ 
  If  $min_{th} \leq avg < max_{th}$ 
    Caculate propability  $p_a$ 
    With propability  $p_a$ :
      mark the ariving packet
  else if  $max_{th} \leq avg$ 
    Mark or drop the ariving packet
  Else
    Accept the arriving packet.

```

Hình 2.3. Giải thuật tổng quát của RED

Giải thuật chi tiết của RED tại gateway được mô tả như hình 2.4.

```

Khởi tạo:
   $avg \leftarrow 0$ 
   $count \leftarrow -1$ 
for mỗi gói tin đến
  Tính kích thước hàng đợi trung bình  $avg$ :
    if hàng đợi không rỗng
       $avg \leftarrow (1-w_q) avg + w_q * q$ 
    else
       $m \leftarrow f(time - q\_time)$ 
       $avg \leftarrow (1-w_q)^m avg$ 
    if  $min_{th} \leq avg < max_{th}$ 
       $count ++$ 
    Tính toán xác suất  $p_a$ :
       $p_b \leftarrow max_p (avg - min_{th}) / (max_{th} - min_{th})$ 
       $p_a \leftarrow p_b / (1 - count * p_b)$ 
    Với xác suất  $p_a$ :
      đánh dấu gói tin đến
       $count \leftarrow 0$ 
    else if  $avg \leq max_{th}$ 
      đánh dấu gói tin đến
       $count \leftarrow 0$ 
    else  $count \leftarrow -1$ 
  Khi hàng đợi trở nên rỗng
     $q\_time \leftarrow time$ 

```

Hình 2.4. Giải thuật chi tiết của RED

2.3.3. Các tham số của RED

a. Trọng số hàng đợi w_q

b. Các giá trị ngưỡng min_{th} và max_{th}

c. Xác suất loại bỏ tối đa max_p

2.3.4. Một số đánh giá về RED

❖ Ưu điểm:

RED là một điển hình của các chiến lược quản lý hàng đợi động AQM, do vậy RED có đầy đủ các ưu điểm chung của chiến lược AQM, ngoài ra RED còn có một số ưu điểm khác biệt sau: *Tránh tắc nghẽn, Tránh đồng bộ toàn cục, Đơn giản, Cục đại hoá công suất toàn cục, Tính công bằng*

❖ **Nhược điểm:**

- Một trong những vấn đề cơ bản của RED là nó dựa vào độ dài hàng đợi để đánh giá sự tắc nghẽn, trong khi sự tắc nghẽn chỉ xảy ra ở hàng đợi cố định và độ dài hàng đợi đem lại rất ít thông tin về tắc nghẽn.

- Việc cài đặt các tham số phù hợp cho RED khi thực thi ở những môi trường mạng khác nhau là rất khó. Để RED có thể hoạt động lý tưởng, cần phải có một số lượng đủ không gian hàng đợi và giá trị các tham số phù hợp.

- Do phép tính xác suất loại bỏ gói của RED được hình thành nên cơ sở mô hình tuyến tính, nên RED không đáp ứng bản chất phi tuyến của mạng. Vì vậy, cần có những thay đổi cho RED vì lưu lượng trên mạng đi theo từng đợt, gây ra những dao động quá nhanh của hàng đợi trong nút mạng.

- Cơ chế RED hoạt động phụ thuộc rất nhiều vào min_{th} và max_{th} trong khi tình trạng mạng luôn biến động bởi lưu lượng gói tin từ các tuyến khác nhau đến nút mạng. Ngoài ra, RED cũng không đảm bảo sự công bằng giữa các luồng, RED loại bỏ hay nhận gói nhưng không quan tâm đến băng thông của các luồng và cũng không hạn chế được luồng không thích nghi gây ảnh hưởng xấu đến luồng thích nghi.

2.4. Chiến lược A-RED

2.4.1. Hoạt động của thuật toán A-RED

A-RED điều chỉnh thích ứng giá trị max_p để giữ cho kích thước hàng đợi trung bình nằm trong khoảng giá trị min_{th} và max_{th} . Để đạt được mục tiêu này có 4 cách:

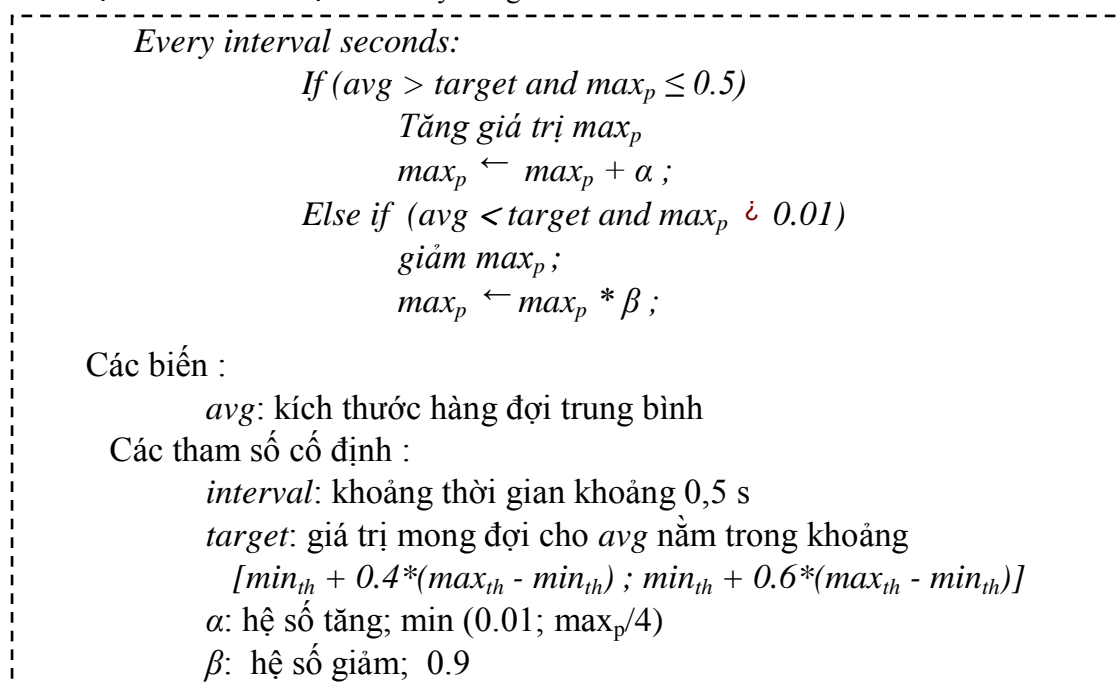
- Max_p được hiệu chỉnh không chỉ giữ cho kích thước hàng đợi trung bình nằm giữa hai giá trị min_{th} và max_{th} mà còn giữ cho kích thước hàng đợi trung bình nằm trong một giải cho phép trong khoảng min_{th} và max_{th} .

- Max_p được điều chỉnh thích nghi chậm, sau những khoảng thời gian vượt quá thời gian khứ hồi (round-trip time), và được thực hiện với chi phí thấp.

- Giá trị max_p được duy trì trong miền $[0.01 ; 0.5]$ (tương ứng với $[1\%, 50\%]$).

- Thay cho việc tăng theo cấp số nhân và giảm giá trị max_p thuật toán thực hiện chế độ tăng theo cấp số cộng giảm theo cấp số nhân (AIMD). Điều đó có nghĩa là khi tăng thì cộng thêm một lượng đủ nhỏ (α), khi giảm thì nhân với một giá trị nhỏ hơn 1 (β). Các giá trị α , β được chọn sao cho kích thước hàng đợi trung bình quay trở lại miền mục tiêu (khoảng (q_{low}, q_{high})) trong thời gian không quá 25s.

Thuật toán A-RED được trình bày trong hình 2.5



Hình 2.5. Thuật toán A-RED

2.4.2. Các tham số của A-RED

a) Phạm vi của max_p

b) Tham số α, β

c) Thiết lập các tham số max_{th} và w_q

2.4.3. Một số đánh giá về A-RED

❖ Ưu điểm

- Tính ổn định của A-RED không phụ thuộc vào tải mạng.
- A-RED có thể dự đoán trước độ trễ hàng đợi trung bình.
- A-RED tự động việc thiết lập các thông số của nó để đáp ứng thay đổi tải nạp.

❖ Nhược điểm

A-RED không làm rõ được rằng đó là chính sách tốt nhất và tối ưu của các thay đổi tham số.

2.4.4. So sánh thuật toán RED và A-RED

Thuật toán A-RED là thuật toán nâng cao của thuật toán RED do đó A-RED khắc phục được mặt hạn chế của RED:

- RED quản lý hàng đợi dựa trên kích thước trung bình của hàng đợi nên kích thước trung bình hàng đợi thay đổi theo các mức tắc nghẽn và quá trình thiết lập các tham số. Điều này được thể hiện bằng việc khi tắc nghẽn xảy ra nhẹ hay max_p cao thì kích thước hàng đợi gần tới giá trị min_{th} . Khi tắc nghẽn trong mạng nặng hay kích thước hàng đợi trung bình bằng hoặc lớn hơn max_{th} . Kết quả trễ hàng đợi trong thuật toán RED phụ thuộc vào tải lưu lượng và các tham số, do đó mà trễ hàng đợi không thể đoán trước.

- Một nhược điểm nữa của RED là khả năng thông qua trong thuật toán này cũng phụ thuộc nhiều vào tải lưu lượng và các tham số.

Do thuật toán A-RED quản lý kích thước trung bình của hàng đợi dựa trên việc tương thích giá trị max_p sao cho kích thước trung bình hàng đợi thay đổi trong khoảng min_{th} và max_{th} nên khắc phục được sự phụ thuộc của trễ hàng đợi và thông lượng của hàng đợi vào các tham số và tải lưu lượng.

2.5. Thuật toán A-RIO

2.5.1. Giới thiệu

2.5.2. Quản lý hàng đợi động trong kiến trúc DiffServ

2.5.3. Thuật toán quản lý hàng đợi A-RIO

A-RIO là một mở rộng trực tiếp cả hai thuật toán A-RED và RIO-C. A-RIO theo cách tiếp cận của A-RED, thực hiện một hiệu chỉnh tự động on-line đối với xxx nhằm đạt được một hiệu năng có thể dự đoán trước. Có nhiều cách tiếp cận đã được đưa ra nhằm hiệu chỉnh RED, ở đây A-RED được chọn vì tính đơn giản và hiệu quả của nó (cả về tư tưởng và cài đặt).

Cũng như A-RED ở chế độ tự động, A-RIO chỉ cần một tham số đầu vào là độ trễ đích, nó sẽ tự động ánh xạ sang tập các tham số router. Đặc trưng này rất có ý nghĩa đối với nhà cung cấp dịch vụ phân loại: cấu hình router theo độ trễ - một độ đo QoS liên quan trực tiếp đến đặc tả dịch vụ và đặc tả yêu cầu của người dùng- chắc chắn dễ hiểu hơn nhiều so với các tham số trừu tượng như ngưỡng hàng đợi, xác suất loại bỏ, hoặc trọng số trung bình...

Thuật toán A-RIO được thể hiện như ở hình 2.6.

```

for mỗi gói tin đến với độ ưu tiên  $i$ ,
  for mỗi mức ưu tiên  $j = i, i + 1, \dots, n$ 
    cập nhật  $avg^{(j)}$ :  $avg^{(j)} \leftarrow avg^{(j)} * (1 - w_q) + q^{(j)} * w_q$ 
    với mỗi đơn vị thời gian  $interval$  cập nhật  $max_p^{(j)}$ :
      if  $avg^{(j)} > q_{high}$  and  $max_p^{(j)} < 0.5$ 
        tính hệ số tăng:  $\alpha \leftarrow \min(0.01, max_p^{(j)} / 4)$ 
        tăng  $max_p^{(j)}$ :  $max_p^{(j)} \leftarrow max_p^{(j)} + \alpha$ 
        if  $j < n$  then:  $max_p^{(j)} \leftarrow \min(max_p^{(j)}, max_p^{(j+1)})$ 
      else if  $avg^{(j)} < q_{low}$  và  $max_p^{(j)} > 0.01$ 
        giảm  $max_p^{(j)}$ :  $max_p^{(j)} \leftarrow max_p^{(j)} * \beta$ 
        if  $j > 0$  then:  $max_p^{(j)} \leftarrow \max(max_p^{(j)}, max_p^{(j-1)})$ 
    if  $min_{th} < avg^{(i)} \leq max_{th}$ 
      tính  $p^{(i)}$  như trong A-RED
      loại gói tin này với xác suất  $p^{(i)}$ 
    else if  $max_{th} < avg^{(i)} \leq 2 * max_{th}$ 
      tính  $p^{(i)_{gentle}}$  như trong A-RED
      loại gói tin này với xác suất  $p^{(i)_{gentle}}$ 
    else if  $avg^{(i)} > 2 * max_{th}$ 
      loại gói tin này

```

Hình 2.6. Thuật toán A-RIO

Các biến và các tham số cố định:

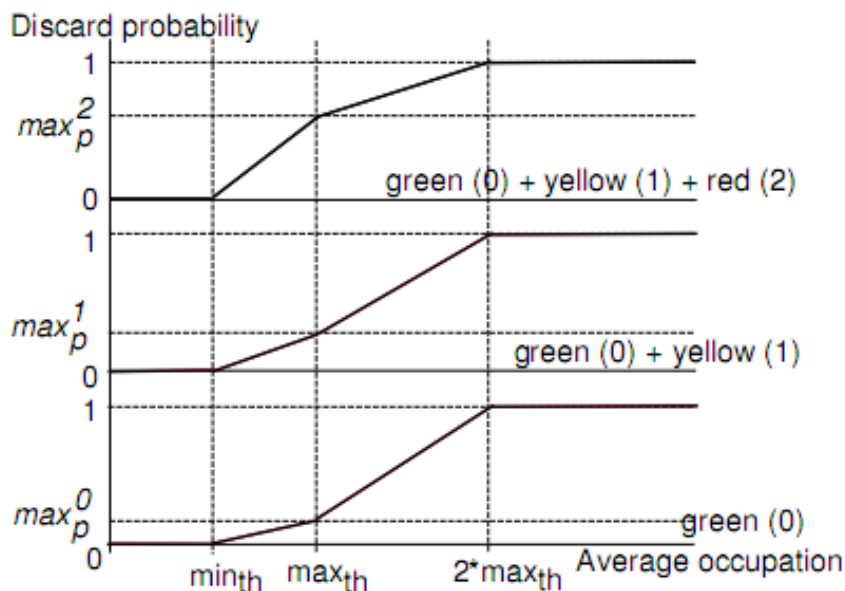
$avg^{(i)}$: kích thước trung bình ứng với mức ưu tiên i (là tổng số gói tin của các mức ưu tiên từ 1 đến i)

$max_p^{(i)}$: xác suất loại bỏ tối đa ứng với mức ưu tiên i (khi $avg^{(i)} = max_p^{(i)}$)

$p^{(i)}$: xác suất loại bỏ ứng với mức ưu tiên i

$p^{(i)_{gentle}}$: xác suất loại bỏ ứng với mức ưu tiên i trong vùng *gentle*

$interval$: 0.5 s; β (hệ số giảm): 0.9



Hình 2.7. A-RIO với ba mức ưu tiên

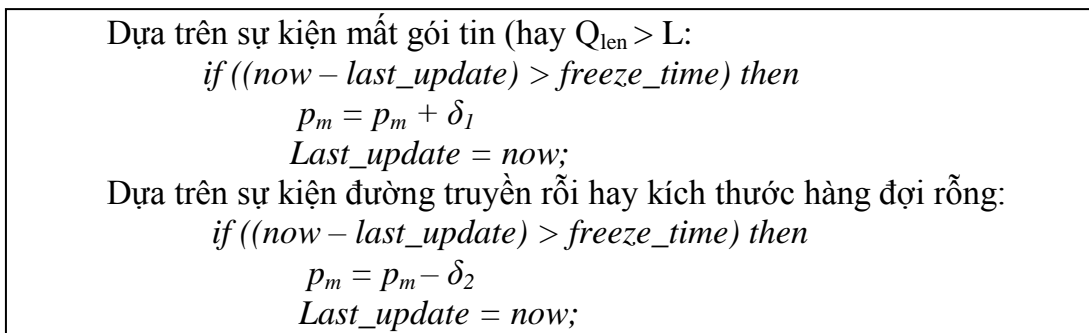
CHƯƠNG 3. CHIẾN LƯỢC BLUE VÀ ĐỀ XUẤT CẢI TIẾN GIẢI THUẬT QUẢN LÝ HÀNG ĐỢI BLUE

Như đã trình bày ở trên về chiến lược RED và A-RED, ta thấy các thuật toán được thiết kế với mục tiêu giảm thiểu mất gói tin và độ trễ của hàng đợi, tránh hiện tượng đồng bộ hóa toàn cầu về nguồn, duy trì mức độ sử dụng đường truyền cao. Nhưng các chiến lược này lại không hiệu quả trong việc ngăn ngừa tỉ lệ mất gói tin cao. Một thuật toán quản lý hàng đợi tích cực khác được đề xuất đó là thuật toán BLUE. Bằng cách sử dụng cả mô phỏng và thử nghiệm, ta thấy rằng BLUE khắc phục được nhiều thiếu sót của RED, nó cải thiện hiệu suất của RED trong tất cả các khía cạnh, ngay cả khi được sử dụng với kích thước không gian bộ đệm tối thiểu. Điều này làm giảm độ trễ end-to-end qua mạng. BLUE là cơ chế quản lý hàng đợi tích cực dựa trên tải nạp để dự đoán khả năng sử dụng đường truyền, xác định tắc nghẽn và đưa ra cách xử lý hiệu quả hơn phương pháp dựa vào kích thước hàng đợi trung bình. Nó là một giải thuật cho phép quản lý kiểm soát tắc nghẽn dựa trên sự kiện mất gói dữ liệu và mức độ sử dụng đường truyền thay vì chiếm dụng hàng đợi.

3.1. Giải thuật BLUE

Năm 2002, Wu-chang Feng và cộng sự đề xuất cơ chế BLUE [13,24]. Ý tưởng chính đằng sau thuật toán quản lý hàng đợi BLUE là dựa trực tiếp trên sự mất gói tin và việc sử dụng các liên kết hơn là trên các độ dài trung bình hàng đợi tức thời. Điều này tương phản sắc nét với tất cả các kỹ thuật quản lý hàng đợi đã được sử dụng trong điều khiển tắc nghẽn trước đó. Thuật toán quản lý hàng đợi Blue sử dụng độ mất gói và độ khả dụng liên kết để quản lý tắc nghẽn bằng cách phát hiện và điều chỉnh tốc độ của các gói bị loại bỏ hoặc bị đánh dấu.

BLUE sử dụng một biến tham số xác suất p_m để đánh dấu các gói tin khi chúng vào hàng đợi. Xác suất này tăng/giảm một cách tuyến tính tùy thuộc vào tỉ lệ rơi (loại bỏ) gói tin hay mức độ sử dụng đường truyền. Nếu như hàng đợi liên tục hủy bỏ các gói tin vì nguyên nhân tải nạp lớn làm tràn bộ nhớ đệm thì BLUE sẽ tăng xác suất đánh dấu p_m , do đó tăng tốc độ gửi lại thông báo tắc nghẽn hoặc loại bỏ các gói tin. Ngược lại, nếu như hàng đợi trở nên trống rỗng hoặc đường truyền rỗi, BLUE lại giảm xác suất đánh dấu (hay loại bỏ) gói tin của nó. Điều này cho phép BLUE tự điều chỉnh tốc độ cần thiết để gửi thông báo tắc nghẽn trở lại nơi gửi hoặc cho rơi gói tin. Thuật toán đánh dấu (loại bỏ) các gói tin với xác suất p_m được trình bày như hình 3.1.



Hình 3.1. Giải thuật BLUE

Hoạt động của thuật toán trên có thể được mô tả theo 4 bước sau:

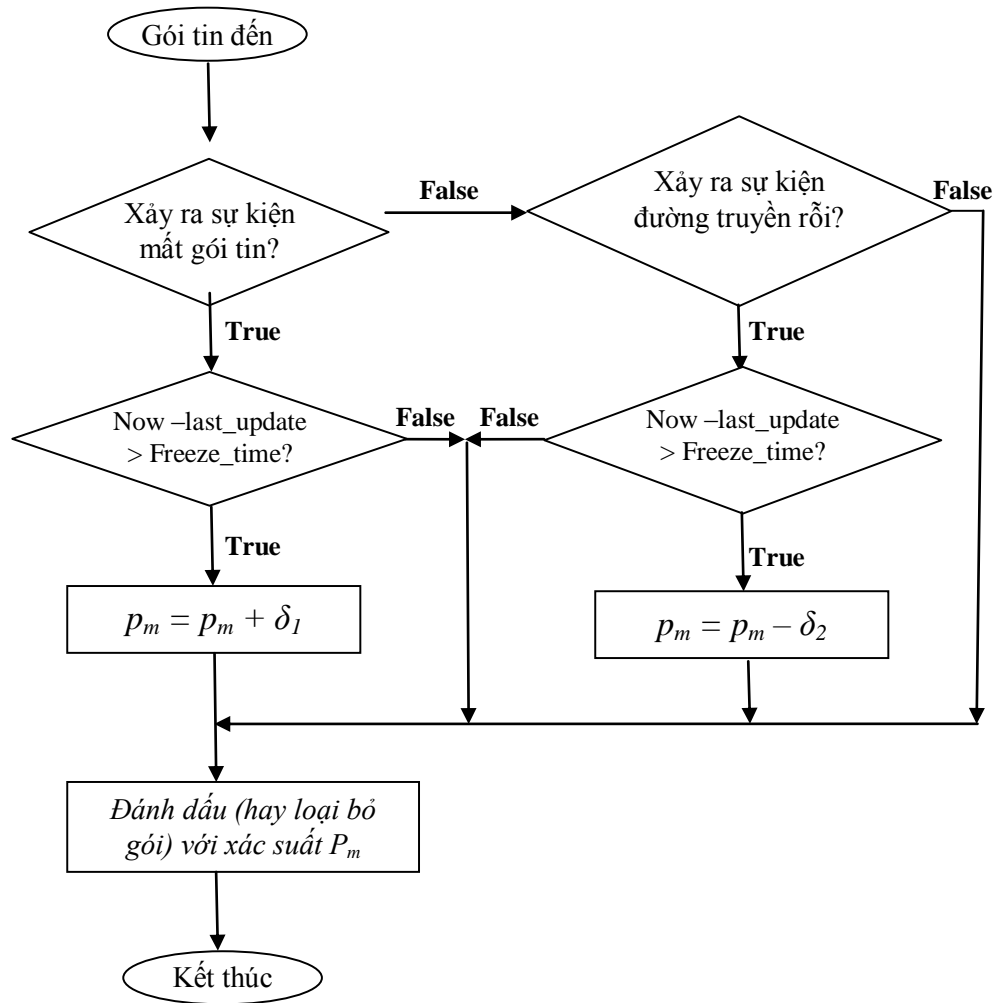
Bước 1: Kiểm tra nếu xảy ra sự kiện mất gói tin thì qua bước 2, nếu không, kiểm tra xem nếu sự kiện đường truyền rỗi thì qua bước 3, ngược lại qua bước 4.

Bước 2: Kiểm tra, nếu khoảng thời gian từ lần cập nhật cuối cùng đến thời điểm hiện tại mà lớn hơn ngưỡng cho phép thì tăng xác suất đánh dấu (hoặc loại bỏ) gói p_m lên. Qua bước 4.

Bước 3: Kiểm tra, nếu khoảng thời gian từ lần cập nhật cuối cùng đến thời điểm hiện tại mà lớn hơn ngưỡng cho phép thì xác suất đánh dấu (hoặc hủy bỏ) gói p_m xuống. Qua bước 4.

Bước 4: Đánh dấu (hoặc loại bỏ) gói tin đến với xác suất p_m .

Giải thuật quản lý hàng đợi BLUE cũng có thể trình bày dạng lưu đồ như trong hình 3.2.



Hình 3.2 Lưu đồ giải thuật BLUE

BLUE có cơ chế quản lý hàng đợi theo tải nạp hiệu quả hơn phương pháp dựa vào chiều dài hàng đợi trung bình. Điều này tương phản một cách rõ ràng với tất cả các thuật toán điều khiển hàng đợi tích cực đã biết, bởi các thuật toán này sử dụng không gian của hàng đợi như là một tiêu chuẩn trong việc điều khiển tránh tắc nghẽn. Nhưng ở chiến lược BLUE thì chúng ta thấy rằng cơ chế này đánh dấu gói tin đến mà không quan tâm đến luồng truyền gói tin đó. Điều này dẫn đến việc đánh rơi gói tin một cách ngẫu nhiên trong tất cả các luồng hoạt động, không đảm bảo sự công bằng giữa các luồng. Đồng thời, việc thiết lập giá trị cho các tham số *freeze_time*, δ_1 và δ_2 hợp lý cho từng môi trường mạng gặp rất nhiều khó khăn. Chính bởi điều này BLUE còn tồn tại một số vấn đề cần phải được cải tiến sau:

- Tham số *freeze_time* cần phải được thiết lập một cách tự động dựa trên thời gian khứ hồi hiệu quả nhằm cho phép bất kỳ sự thay đổi nào trong việc gán xác suất sẽ được phản ánh lại đến nơi gửi trước khi sự thay đổi tiếp theo xảy ra. Đối với đường truyền dài với độ trễ lớn như các đường truyền vệ tinh, *freeze_time* cần phải được tăng lên để phù hợp với thời gian khứ hồi lớn hơn.

- Tham số δ_1 và δ_2 phải được thiết lập phù hợp với tình trạng mạng, cho phép đường truyền có khả năng thích nghi hiệu quả với những thay đổi vĩ mô trong lưu lượng truyền đi qua đường kết nối. Đối với các đường truyền mà tại đó trung bình trong vài phút xảy ra sự thay đổi lưu lượng truyền thì δ_1 và δ_2 phải được thiết lập kết hợp với *freeze_time* để cho phép p_m thay đổi giá trị trung bình trong vài phút. Ngược lại, trong những môi trường mạng thay đổi lưu lượng gói tin đến nút mạng theo giây thì phải cập nhật các tham số *freeze_time*, δ_1 và δ_2 để p_m thích nghi từng giây.

3.2. Đánh giá về thuật toán BLUE:

❖ Ưu điểm:

- BLUE làm mất ít gói dữ liệu hơn
- BLUE sử dụng không gian bộ đệm nhỏ
- BLUE duy trì chiều dài hàng đợi ổn định hơn
- Loại bỏ những tác nhân chống lại các nguồn bùng phát.

❖ **Nhược điểm:**

- Không phát hiện tắc nghẽn sớm (các gói tin bị loại bỏ được cập nhật trong hàng đợi trên các luồng hoặc các sự kiện liên kết nhàn rỗi)
- Phản ứng chậm và phụ thuộc vào lịch sử
- Khi tất cả các gói tin được đánh dấu, nhưng các nguồn vẫn đang bị quá tải tại các liên kết nút cổ chai.

3.3. So sánh thuật toán RED và thuật toán Blue

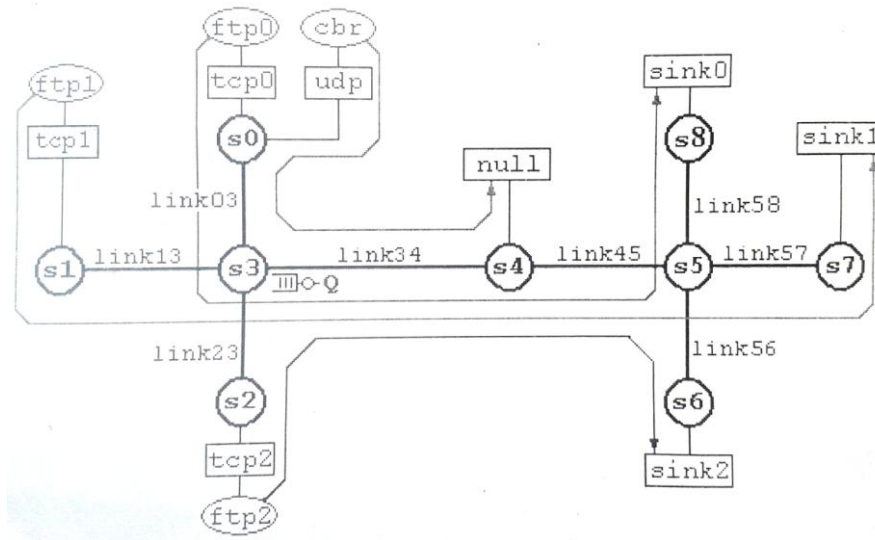
Một vấn đề quan trọng trong quản lý hàng đợi bằng thuật toán Blue là điều khiển tắc nghẽn có thể được thực hiện bởi kích thước hàng đợi nhỏ nhất. Trong khi đó thuật toán RED thì lại yêu cầu kích thước hàng đợi lớn hơn cho cùng một mục đích. Do có kích thước bộ đệm nhỏ hơn có trễ đầu cuối qua mạng nhỏ hơn so với thuật toán RED, do đó nó cải thiện được nhược điểm của thuật toán điều khiển tắc nghẽn. Thêm vào đó các yêu cầu kích thước bộ đệm nhỏ hơn cho phép có nhiều bộ nhớ hơn để phân phối cho các gói có độ ưu tiên cao và giải phóng bộ nhớ trong router để cho các chức năng khác như lưu trữ các bảng định tuyến lớn. Blue cho phép các router thế hệ sau hoạt động tốt thậm chí cả trong trường hợp tài nguyên bộ nhớ bị giới hạn. Tuy nhiên thuật toán Blue ít nhạy cảm với các lựa chọn tham số hơn là đối với giải thuật RED.

CHƯƠNG 4. ĐÁNH GIÁ HIỆU SUẤT CÁC CHIẾN LƯỢC QUẢN LÝ HÀNG ĐỢI RED, ARED VÀ BLUE BẰNG BỘ MÔ PHÒNG

4.1 Đánh giá hiệu suất của chiến lược quản lý hàng đợi Red

Trước tiên chúng tôi tiến hành thực hiện lại các mô phỏng trong bài báo [1] (không trình bày ở đây), mục đích là kiểm nghiệm lại các đánh giá của các tác giả, và để làm cơ sở cho việc chọn các tham số RED phục vụ cho các mô phỏng sau này của chúng tôi. Sau khi thực hiện các mô phỏng này, chúng tôi nhận thấy các kết quả mà các tác giả đưa ra là hoàn toàn chính xác và có thể tin cậy. Sau đây là phần trình bày mô phỏng của chúng tôi nhằm so sánh, đánh giá hiệu suất các chiến lược RED và DropTail.

4.1.1 Cấu hình mạng mô phỏng



Hình 4.1. Topo mạng mô phỏng

Cho một mạng mô phỏng có cấu hình, các thực thể gửi/nhận và các nguồn sinh lưu lượng ... như hình vẽ trên. Mạng mô phỏng gồm 9 nút được đánh số từ s0 đến s8. Đường truyền giữa các node đều là full-duplex, không có lỗi:

Link03, link13, link23: 10 Mbps, 1ms

Link34, link45: 1.5 Mbps, 10 ms

Link 56, link57, link58: 10 Mbps, 1ms

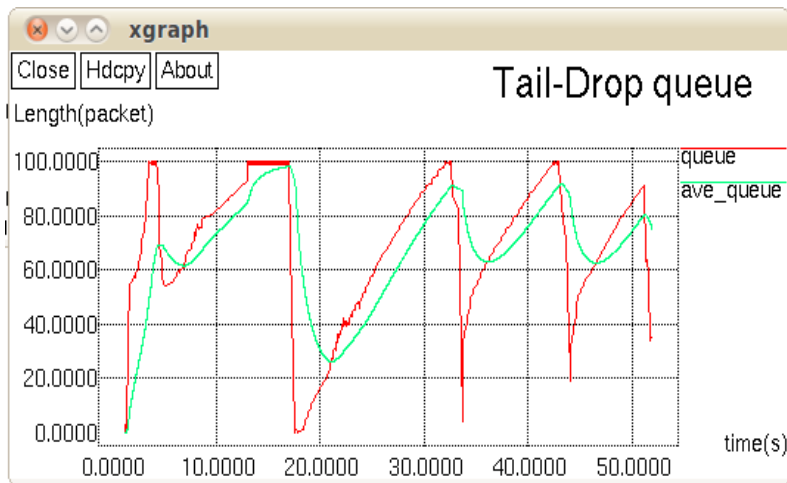
Các thực thể gửi gắn với ba luồng tcp là các nguồn FTP (ftp0, ftp1, ftp2), còn nguồn phát cho udp là nguồn CBR (nguồn sinh lưu lượng với tốc độ không đổi, ở đây ta chọn tốc độ phát cho cbr là 1.8Mbps). Các luồng tcp gửi các gói tin với kích thước 1000 bytes, kích thước cửa sổ tối đa là 64. Các thực thể gửi đưa lưu lượng vào mạng trong các khoảng thời gian như sau: ftp0:1.1s-51.1s; ftp1:1.5s-51.5s; ftp2:1.9s-51.9s; cbr: 13.0s-17s. Hàng đợi Q được đặt giữa nút s3 và nút s4 có kích thước bằng 100 gói tin. Tổng thời gian mô phỏng là 52s. Chúng ta sẽ thay đổi chính sách quản lý tại hàng đợi Q lần lượt là DropTail và RED và so sánh các kết quả.

Với mỗi mô phỏng chúng tôi đưa ra 3 đồ thị: kích thước hàng đợi trung bình, thông lượng sử dụng của mỗi kết nối và kích thước cửa sổ để đánh giá các đại lượng liên quan như độ trễ trung bình, thông lượng của từng kết nối, và nghiên cứu hiện tượng đồng bộ toàn cục. Sau đây là chi tiết về kết quả thu được từ các mô phỏng.

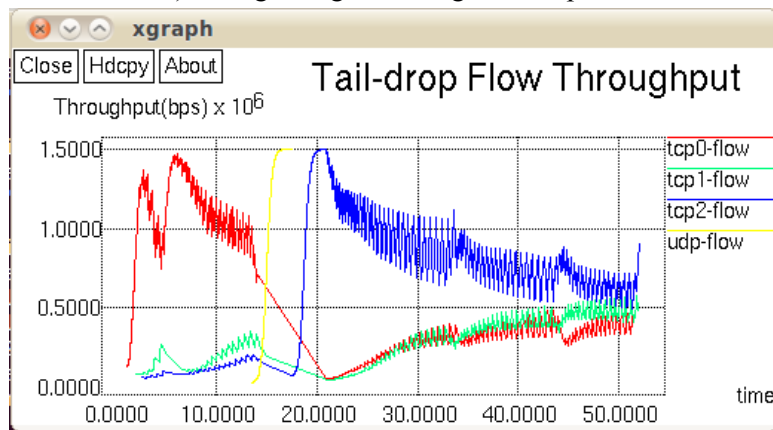
4.1.2 Mô phỏng với chính sách quản lý hàng đợi DropTail:

Sử dụng các tham số cấu hình mạng mô phỏng như mục trên, với việc thiết lập tất cả các hàng đợi đều là DropTail, chúng tôi nhận được các kết quả như ở hình 4.2.

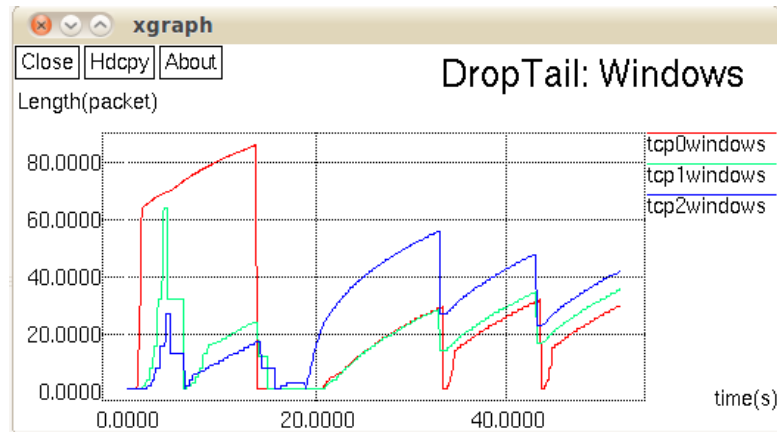
a) Kích thước hàng đợi trung bình của DropTail



b) Thông lượng các luồng của DropTail



c) Kích thước cửa sổ của các luồng của DropTail và RED



Hình 4.2. Các kết quả mô phỏng 1 với hàng đợi DropTail

Nhận xét:

- Trong các khoảng có đột biến lưu lượng (13.0s-17.0s), nguồn cbr có tốc độ cao hơn dung lượng đường truyền ra tại nút 3, gây ra hiện tượng lock-out tại hàng đợi Q, dẫn tới hiện tượng global synchronization, thông lượng của các kết nối tcp đồng loạt giảm kích thước cửa sổ phát (hình 4.2c), dẫn tới thông lượng của các kết nối tcp giảm xuống rất nhỏ (hình 4.2b), cực tiểu khoảng 0.12Mbps. Do cơ chế rút lui theo hàm mũ (exponential backoff) của TCP, trạng thái này còn kéo dài sau khi nguồn cbr ngừng hoạt động. Mặt khác trong thời gian này kích thước hàng đợi hầu như đầy (hình 4.2a), dẫn tới độ trễ hàng đợi cao.

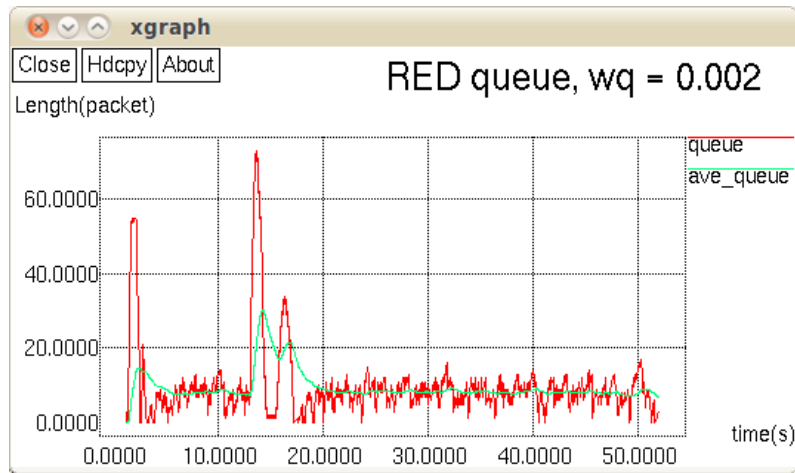
- Ngay cả khi nguồn cbr không hoạt động (từ 20s đến 51.9s), thông lượng của các kết nối tcp0, tcp1 và tcp2 cũng thăng giáng trong một miền rất rộng. Và cũng trong khoảng thời gian này hiện tượng đồng bộ toàn cục vẫn xuất hiện (ở các thời điểm 22.5s, 42.5s), các kết nối cùng tăng kích thước cửa sổ

cho đến khi đạt đến ngưỡng thì đồng thời giảm xuống; kích thước hàng đợi vì thế cũng dao động trong một miền rất rộng (khoảng 20% - 30%) so với giá trị trung bình. Ngoài ra chiều dài hàng đợi trung bình (ave_queue) thường xuyên ở mức cao (cỡ 75 ± 15 packet).

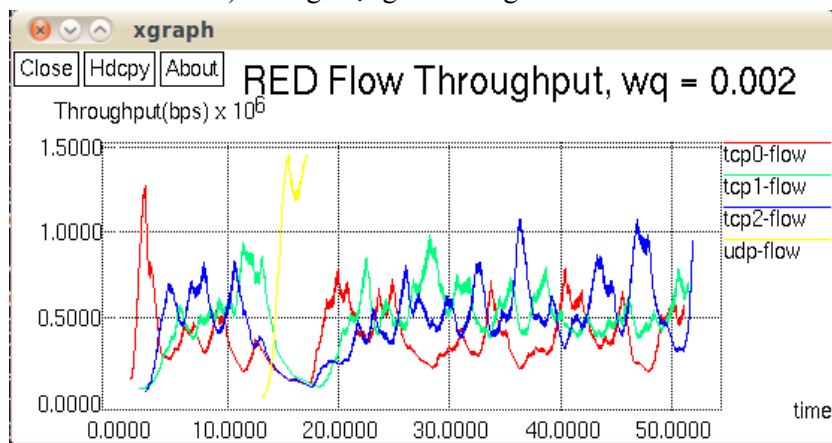
4.1.3 Mô phỏng với chính sách RED:

Các tham số được thiết lập cho RED như sau: minth = 5, maxth = 15, maxp = 0.1 và wq = 0.002. Kết quả mô phỏng được thể hiện ở các đồ thị trên hình 4.3.

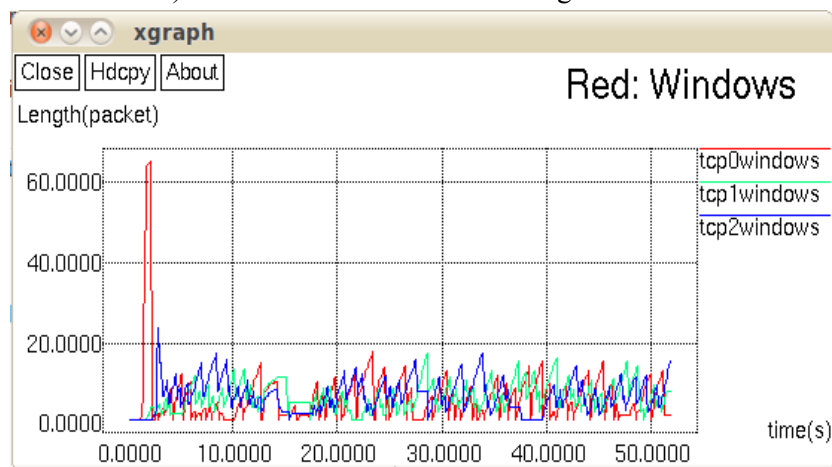
a) Kích thước hàng đợi trung bình của DropTail và RED



b) Thông lượng các luồng của RED



c) Kích thước cửa sổ của các luồng của RED



Hình 4.3. Các kết quả mô phỏng 1 với hàng đợi RED

Nhận xét:

- Chúng ta có thể thấy rằng, trong các giai đoạn đưa lưu lượng đột biến cbr vào, kích thước cửa sổ các kết nối tcp giảm xuống rất nhanh (hình 4.3c), kéo theo thông lượng của chúng giảm theo (hình 4.3b),

nhưng sau khi ra khỏi giai đoạn đó thì các kết nối này nhanh chóng tăng kích thước cửa sổ lên, thông lượng vì thế nhanh chóng được hồi phục; mặt khác kích thước hàng đợi tăng lên nhưng nhanh chóng được kéo xuống;

- Trong giai đoạn không có đột biến (từ 20s trở đi) thì RED luôn duy trì được kích thước hàng đợi trung bình (ở khoảng 10 gói tin), kích thước hàng đợi hiện tại dao động ở mức nhỏ (10 ± 2).

Ngoài ra chúng tôi cũng đã so sánh độ trễ trung bình và độ lệch chuẩn của độ trễ trong toàn thời gian mô phỏng với hàng đợi DropTail và RED. Kết quả mô phỏng với hàng đợi DropTail được thể hiện trên bảng 4.1.

<i>Kết nối</i>	<i>Độ trễ trung bình (Mean delay)</i>	<i>Độ lệch chuẩn của độ trễ (Standard deviation of delay)</i>
TCP (s0-s8)	0.620883004129387	68.041217909686
TCP (s1-s7)	0.723091946788987	49.713802901559
TCP (s2-s6)	0.411800912368419	23.8705707028506
UDP (s0-s4)	0.410905662222221	7.18025976022592

Bảng 4.1. So sánh độ trễ trung bình và độ lệch chuẩn của độ trễ với hàng đợi DropTail
Kết quả mô phỏng với hàng đợi RED được thể hiện trên bảng 4.2 dưới đây.

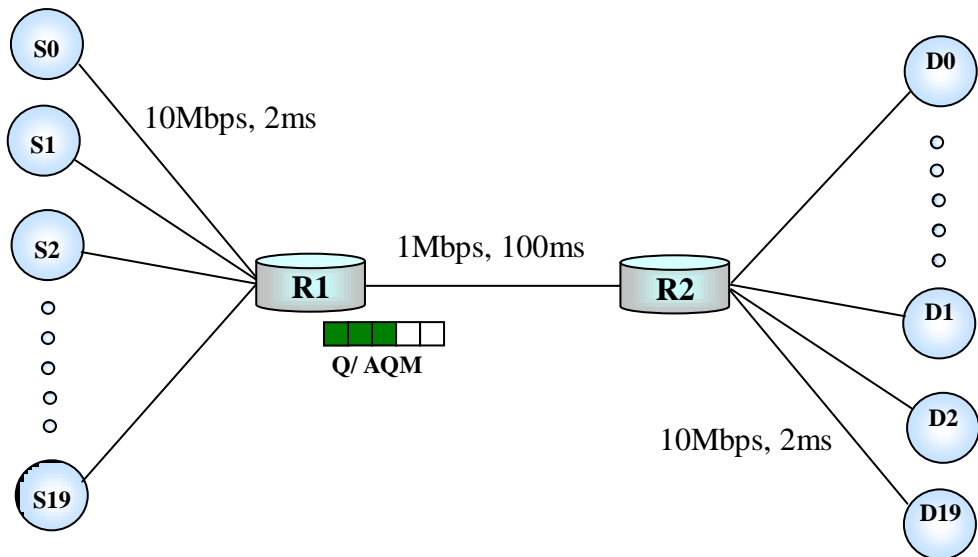
<i>Kết nối</i>	<i>Độ trễ trung bình (Mean delay)</i>	<i>Độ lệch chuẩn của độ trễ (Standard deviation of delay)</i>
TCP (s0-s8)	0.105534121971595	4.83272185887846
TCP (s1-s7)	0.0922332979999994	5.99194460746312
TCP (s2-s6)	0.0860523832394832	3.66091765635048
UDP (s0-s4)	0.11072745	3.69114334991382

Bảng 4.2. So sánh độ trễ trung bình và độ lệch chuẩn của độ trễ với hàng đợi RED

Ở phần trên chúng ta đã thực hiện mô phỏng khi hệ thống trong trạng thái bình thường. Bây giờ chúng ta sẽ xem xét hệ thống mạng khi có tắc nghẽn xảy ra để đánh giá khả năng hấp thụ các lưu lượng đột biến của RED và DropTail.

Chúng tôi đã xây dựng mạng mô phỏng có số lượng các thực thể gửi và nhận lớn như trên hình 4.6. Bằng cách thay đổi hàng đợi DropTail, RED, A-RED, BLUE chúng ta sẽ so sánh hiệu năng của các chiến lược.

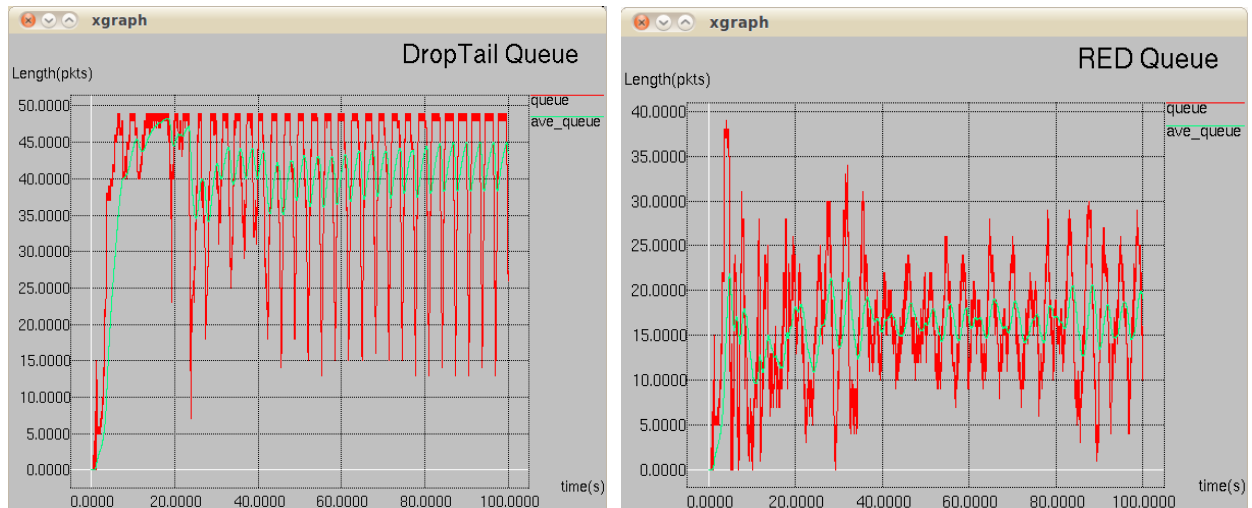
Mạng mô phỏng có tổng số nút là 40 bao gồm 20 nút nguồn (nút gửi dữ liệu) từ S0-S19 và 20 nút nhận D0-D19. Các thực thể gửi đều là TCP, kích thước cửa sổ gửi tối đa là 32 gói tin và đều xuất phát từ các nguồn sinh lưu lượng FTP. Các đường truyền đều là duplex-link, không lỗi; đường truyền từ các nút nguồn đến R1 có băng thông 10Mbps, độ trễ 2ms; đường truyền từ R2 đến các nút nhận đều là 10Mbps, 2ms. Các kết nối cùng chia sẻ đường truyền chung R1- R2 có băng thông, độ trễ lần lượt là 1Mbps, 100ms. Hàng đợi Q được đặt giữa R1-R2 có kích thước tối đa 50 gói tin, các luồng TCP gửi các gói tin có kích thước 1000 bytes. Tổng thời gian mô phỏng là 100s.



Hình 4.6. Cấu hình mạng mô phỏng RED/ A-RED/ BLUE

4.1.4. Khả năng hấp thụ các lưu lượng đột biến của RED

Sử dụng cấu hình mạng như hình 4.6: Các luồng ftp được đưa vào mạng, mỗi luồng cách nhau 2s. Chúng ta sẽ xem kết quả ở hình 4.7.



Hình 4.7. Kết quả mô phỏng 2 so sánh DropTail và RED

Nhìn vào các đồ thị trên hình 4.7 ta thấy rằng: khi ta đưa luồng lưu lượng đột biến vào mạng trong khoảng thời gian mô phỏng thì cả hai chiến lược DropTail và RED đều làm cho kích thước hàng đợi tăng. Nhưng với DropTail mỗi khi ta đưa luồng lưu lượng đột biến vào thì kích thước hàng đợi tăng đột ngột đến ngưỡng rồi giảm xuống rất nhanh: khoảng 20s đầu kích thước hàng đợi và kích thước hàng đợi trung bình dao động mạnh từ 0 đến 50 packet. Trong các khoảng thời gian còn lại của mô phỏng thì kích thước hàng đợi dao động ở ngưỡng từ 15 đến 50 packet, kích thước hàng đợi trung bình dao động trong khoảng 40 ± 5 gói tin. Trong khi đó với RED thì kích thước hàng đợi dao động ổn định từ 5 đến 25 packet, kích thước hàng đợi trung bình dao động trong khoảng 16 ± 3 gói tin.

Ngoài ra chúng tôi cũng đã thống kê một số giá trị trung bình trên toàn bộ thời gian mô phỏng và nhận được kết quả như trong bảng dưới 4.3 dưới đây.

Chiến lược	Kích thước hàng đợi trung bình (gói tin)	Độ trễ hàng đợi trung bình (ms)	Hệ số sử dụng đường truyền (%)
DropTail	40	213.33	95.51
RED	16	85.33	95.36

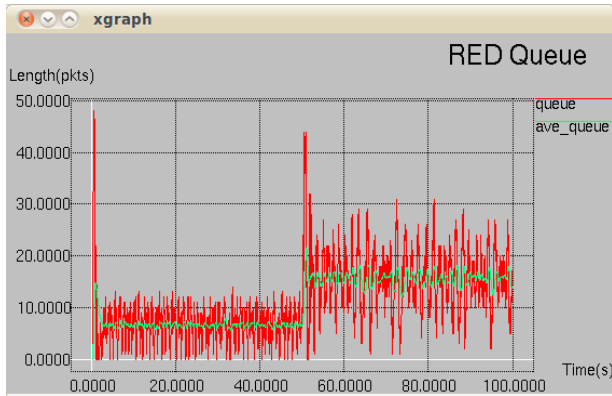
Bảng 4.3. Kết quả thống kê của mô phỏng 2 so sánh DropTail/RED

4.1.5. So sánh RED với Tail-Drop

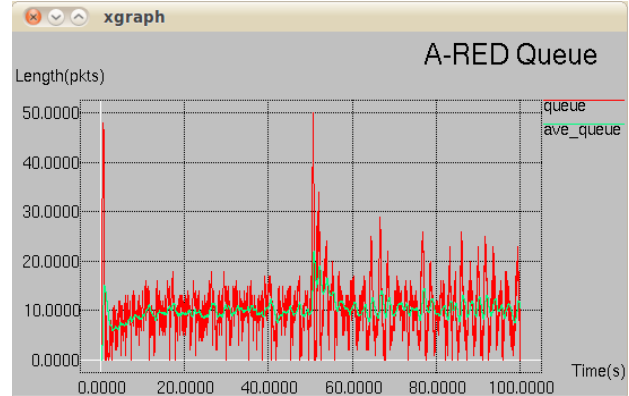
4.2. Đánh giá hiệu suất của chiến lược quản lý hàng đợi A-RED

4.2.1. Kịch bản mô phỏng 1: Tăng cường độ tắc nghẽn với các luồng lưu lượng

Kịch bản được thiết lập như sau: đầu tiên hai kết nối tcp0 và tcp1 được đưa vào mạng (ở 0.1s và 0.2s), đến nửa thời gian mô phỏng (giây thứ 50), 18 luồng mới (tcp2-tcp19) được đưa vào mạng, mỗi luồng cách nhau 0.1 giây. Mục đích của việc đưa các luồng tcp2 –tcp19 vào mạng để lưu lượng mạng được làm tăng đột ngột khi đó băng thông của các luồng truyền sẽ lớn hơn băng thông tại hàng đợi sẽ gây hiện tượng tắc nghẽn. Chúng ta sẽ theo dõi kết quả của các chiến lược trên các hình dưới đây.



Hình 4.8. RED với sự tăng cường độ tắc nghẽn



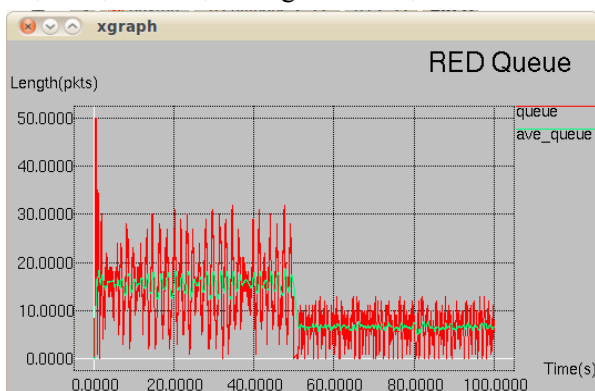
Hình 4.9. ARED với sự tăng cường độ tắc nghẽn

❖ Nhận xét:

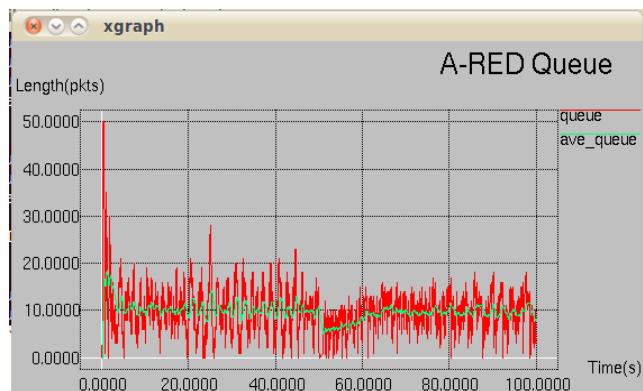
Trên hình 4.8 và 4.9 là đồ thị của kích thước hàng đợi hiện thời (màu đỏ) và kích thước hàng đợi trung bình (màu xanh) ứng với thuật toán RED và A-RED. Nhìn vào đồ thị chúng ta thấy rằng khi tắc nghẽn được tăng cường ở nửa thời gian mô phỏng (ở giây thứ 50), cả RED và A-RED đều làm kích thước hàng đợi tăng lên tối đa; dẫn tới kích thước hàng đợi trung bình tăng lên; với RED là từ 7 lên khoảng 16 gói tin, với A-RED là từ 9 đến 18 gói tin; tuy nhiên sau khoảng 10s (từ giây 60 trở đi), A-RED đã kéo kích thước hàng đợi trở về khoảng mục tiêu và dao động ở mức 10 ± 3 gói tin, trong khi RED vẫn giữ kích thước trung bình ở mức cao (16 ± 3 gói tin).

4.2.2. Kịch bản mô phỏng 2: Giảm cường độ tắc nghẽn với các luồng lưu lượng

Kịch bản được thiết lập như sau: đầu tiên tất cả các kết nối từ tcp0 đến tcp19 được đưa vào mạng (bắt đầu từ 0.1s, mỗi luồng cách nhau 0.1 giây), đến nửa thời gian mô phỏng (giây thứ 50), 18 luồng (từ tcp2-tcp19) ngừng hoạt động. Như vậy lưu lượng mạng được làm giảm đột ngột, phản ứng của từng chiến lược được thể hiện trong các đồ thị hình 4.10 và hình 4.11.



Hình 4.10 RED với sự giảm cường độ tắc nghẽn



Hình 4.11. ARED với sự giảm cường độ tắc nghẽn

❖ Nhận xét:

Nhìn vào đồ đồ thị trong trường hợp giảm cường độ tắc nghẽn ta thấy: với RED, tại thời điểm

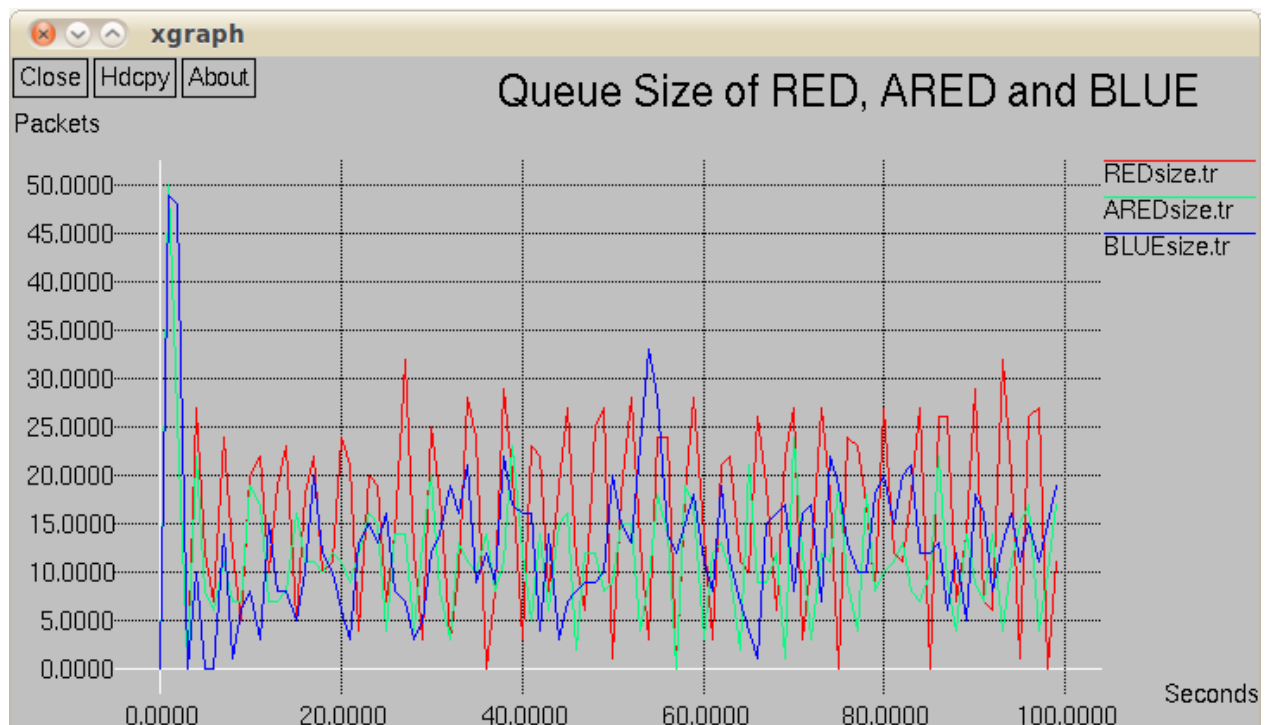
xảy ra đột biến, kích thước hàng đợi trung bình giảm xuống nhanh chóng và ổn định ở một mức mới thấp hơn (từ 16 ± 3 đến 7 ± 1 gói tin); còn với A-RED kích thước hàng đợi trung bình cũng giảm xuống, tuy nhiên mức giảm không đột ngột như RED (từ 10 ± 3 xuống 7 ± 1) và nó nhanh chóng được kéo lên và ổn định ở mức mục tiêu 10 ± 2 gói tin. Với các kết quả đã đưa ra ta thấy rằng hiệu năng của A-RED tốt hơn RED.

4.2.3. So sánh thuật toán RED và ARED

4.3. Đánh giá hiệu suất của chiến lược quản lý hàng đợi BLUE

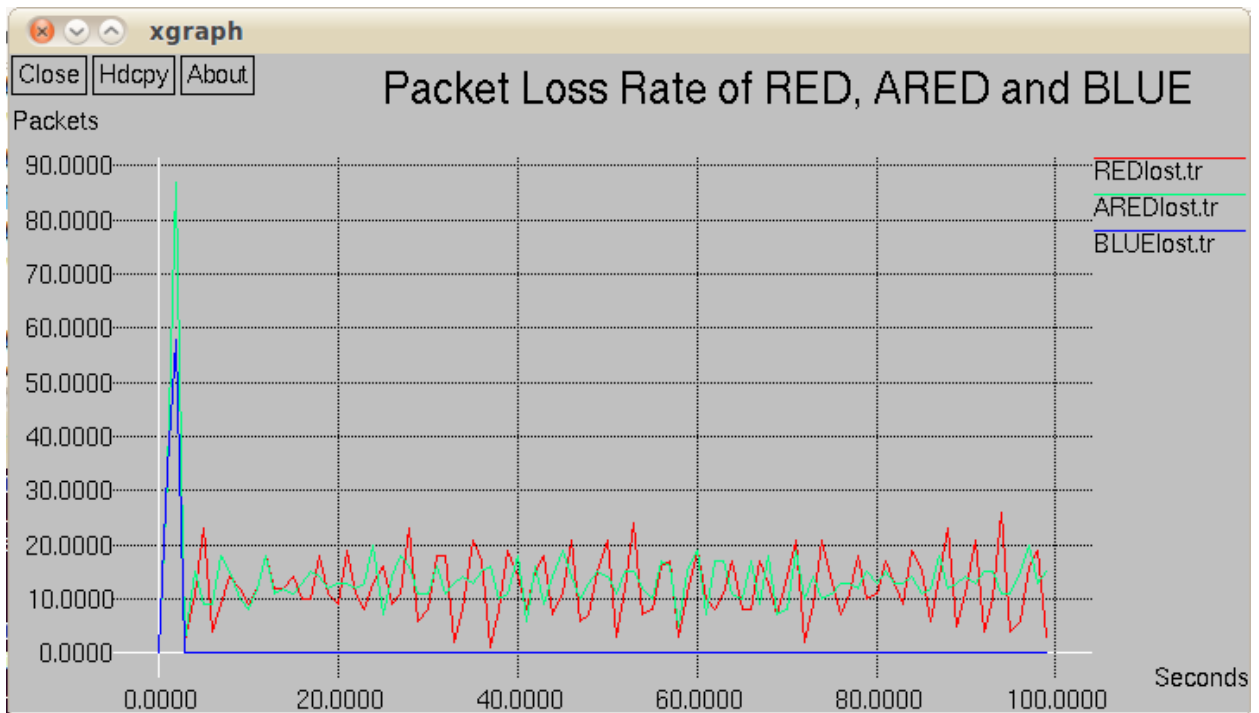
Trong phần này chúng tôi vẫn sử dụng cấu hình mạng như hình 4.6 và với việc thiết lập kịch bản mô phỏng và các tham số cho RED và ARED như mục 4.2 chúng ta sẽ xem xét các kết quả của từng giải thuật RED, A-RED và BLUE dựa trên các tham số: Kích thước hàng đợi trung bình, tỉ lệ mất gói tin và thông lượng sử dụng. Kết quả mô phỏng mạng được thể hiện ở trên các hình 4.12, 4.13 và 4.14.

Hình 4.12 cho ta thấy kích thước hàng đợi của RED dao động mạnh (từ 5packet – 25 packet), kích thước hàng đợi của ARED cũng dao động nhưng mức độ ít hơn so với RED, còn kích thước hàng đợi của BLUE thì mức độ dao động nhỏ hơn. Điều này chứng tỏ rằng độ trễ trung bình và độ lệch chuẩn của độ trễ của BLUE sẽ ít hơn so với RED và ARED.



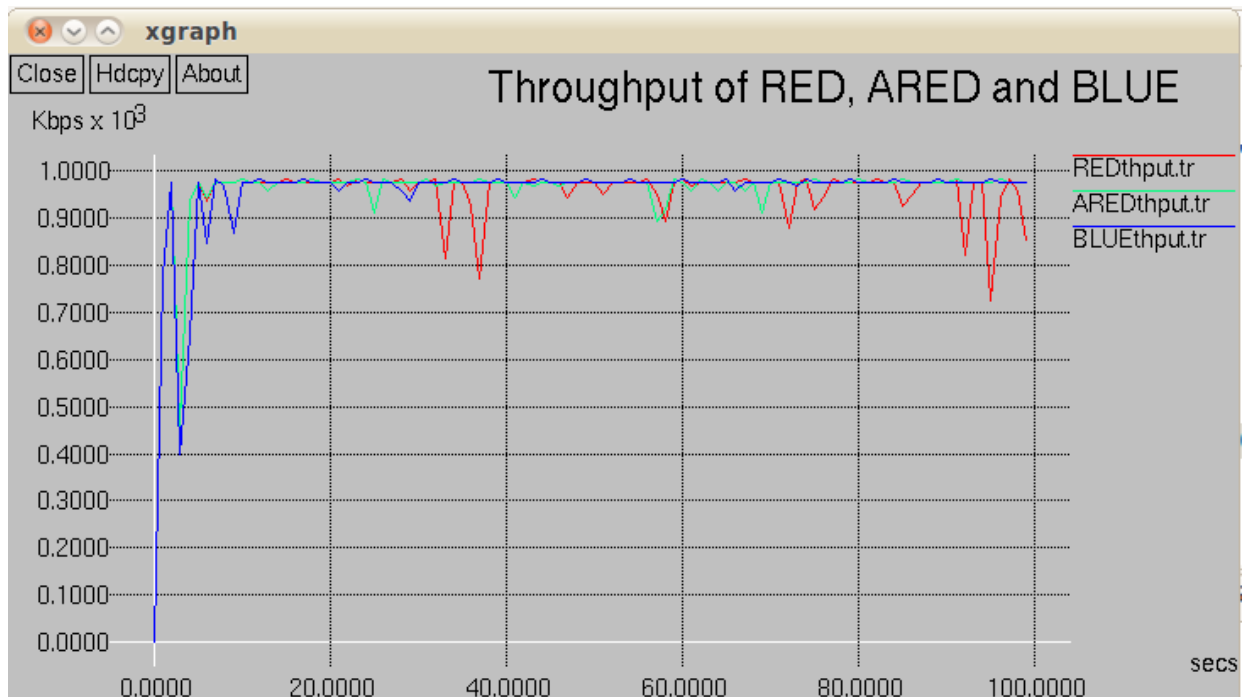
Hình 4.12. Kích thước hàng đợi của RED, A-RED và BLUE

Hình 4.13 cho thấy RED có tỉ lệ mất gói tin nhiều hơn so với ARED và BLUE, điều đó có nghĩa là BLUE có hiệu suất tốt hơn so với RED, ARED. Hàng đợi RED có nhiều gói tin bị mất ngay từ thời điểm bắt đầu của mô phỏng, hàng đợi ARED cũng có gói tin bị mất nhưng ít hơn so với RED, còn BLUE thì gần như không có gói tin bị loại bỏ



Hình 4.13. Tỷ lệ gói tin bị mất của RED, A-RED và BLUE

Hình 4.14 ta thấy trong cả 3 chiến lược băng thông luôn được sử dụng ở mức tối đa.



Hình 4.14. Thông lượng của RED, A-RED và BLUE

KẾT LUẬN

Các thuật toán quản trị hàng đợi tích cực AQM có nhiều ưu điểm nổi bật hơn so với các chiến lược quản lý hàng đợi và lập lịch. Mục tiêu của quản lý hàng đợi tích cực là duy trì một xác suất chủ động loại bỏ gói hợp lý nhằm hạn chế được tình trạng tắc nghẽn trong khi vẫn đảm bảo được chất lượng của các luồng lưu lượng và tính công bằng trong quan hệ giữa các luồng lưu lượng khi trạng thái động học của mạng thay đổi.

Trên cơ sở nghiên cứu các ưu khuyết điểm của các giải thuật quản lý hàng đợi tích cực RED, ARED, BLUE và các giải pháp cải tiến nhằm nâng cao hiệu năng và chất lượng dịch vụ cho truyền thông đa phương tiện. Luận văn đã đạt những kết quả như sau:

- Nghiên cứu cơ sở lý thuyết về truyền thông đa phương tiện và các yêu cầu về chất lượng dịch vụ.
- Nghiên cứu các chiến lược quản lý hàng đợi tích cực: Đối với RED: Chúng tôi đã chỉ ra được những tính năng ưu việt của RED so với DropTail, đồng thời chỉ ra những hạn chế của RED trong những điều kiện mạng cụ thể. Đó là lý do cho sự phát triển của các thuật toán sau nó. Đối với A-RED: Bằng mô phỏng chúng tôi nhận thấy A-RED có những ưu thế nổi bật hơn so với RED. Đó là khả năng tự động thiết lập các tham số, tự động hiệu chỉnh xác suất loại bỏ để duy trì kích thước hàng đợi trung bình mong muốn, trong khi vẫn đảm bảo được thông lượng cao cho mạng. Đối với BLUE: BLUE là một chiến lược quản lý hàng đợi dựa trên tải nạp, qua đó dự đoán khả năng sử dụng đường truyền liên kết, xác định tắc nghẽn và đưa ra cách xử lý. Mục đích của chiến lược là điều tiết gói tin vào nút mạng để ổn định lưu lượng gói tin đến, nhằm duy trì sự ổn định cho mạng.

Hướng phát triển tiếp theo của luận văn là mở rộng và phát triển các giải thuật cải tiến quản lý hàng đợi tích cực AQM nhằm hạn chế tối đa tắc nghẽn để mạng luôn duy trì được sự ổn định cao nhất về chất lượng. Thông qua đó có thể áp dụng các giải thuật cải tiến trên các mô hình mạng phức hợp, và mạng có tổn hao như các mạng không dây, di