



## PHÂN TÍCH HIỆU NĂNG CỦA KIẾN TRÚC INTERNET WEB CACHING NHỎ SỬ DỤNG MÔ HÌNH MẠNG PETRI CÓ MÀU VÀ THỜI GIAN NGẪU NHIÊN (SCPN)

Hồ Khánh Lâm, Nguyễn Xuân Trường, Nguyễn Thị

Chung Trường Đại học Sư phạm Kỹ thuật Hưng Yên

Ngày nhận: 01/10/2016

Ngày sửa chữa: 31/10/2016

Ngày xét duyệt: 15/11/2016

### Tóm tắt:

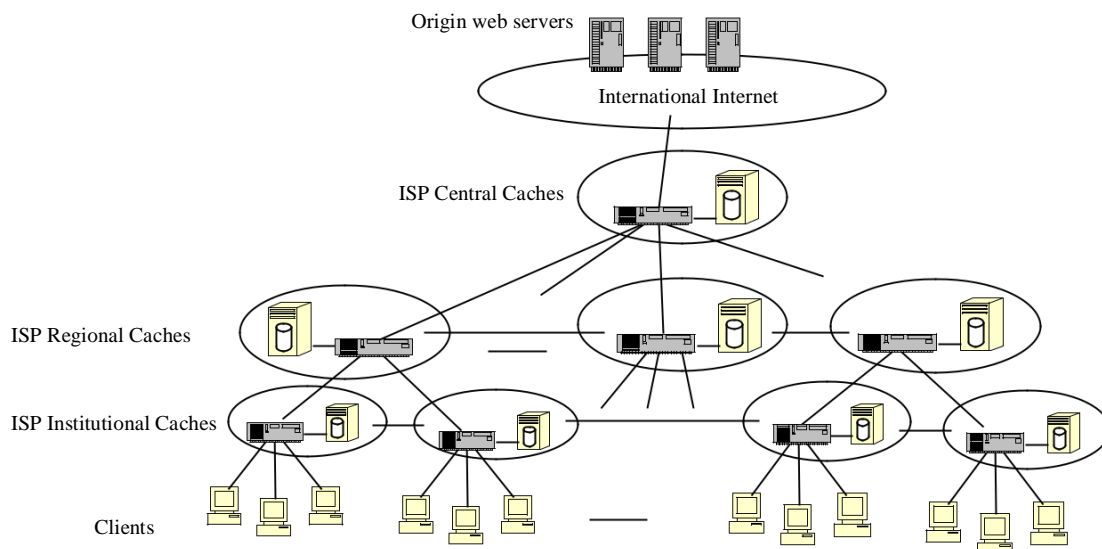
Hiện nay, các công nghệ mạng truyền thông phát triển rất nhanh cung cấp các dịch vụ đa phương tiện tốc độ cao phục vụ nhu cầu của người sử dụng các dịch vụ web. Web caching là ứng dụng ở cấp độ routing và phần lớn băng thông dùng cho web với mục tiêu làm tăng tốc độ đường truyền và tốc độ truy cập web. Hiện nay, có một số phương pháp khác nhau để đánh giá hiệu năng của kiến trúc internet web caching. Tuy nhiên, phương pháp mô hình hóa sử dụng mạng Petri có màu và thời gian ngẫu nhiên (SCPN) là một phương pháp hoàn toàn mới giúp cho việc xác định và điều chỉnh các thông số hiệu năng làm tối ưu kiến trúc internet web caching dựa trên 2 thông số: trễ đáp ứng và chi phí băng thông kênh truyền dẫn. Bài báo này đưa ra phương pháp phân tích hiệu năng của kiến trúc Internet web caching theo hướng tiếp cận dựa trên mạng Petri nhằm mục tiêu tối ưu hóa kiến trúc internet web caching để cải thiện tốc độ truy cập web và sử dụng các dịch vụ đa phương tiện trên mạng internet.

**Từ khóa:** web caching, mạng petri (PN), mô hình mạng Petri có màu và thời gian ngẫu nhiên (SCPN).

### 1. Đặt vấn đề

Ngày nay, các công nghệ truyền thông đều có kết nối truy nhập Internet để cung cấp các dịch vụ đa phương tiện tốc độ cao, thời gian thực đến tận người dùng đầu cuối sử dụng các thiết bị máy tính, điện thoại thông minh. Không thể tính được có bao nhiêu dịch vụ trên Internet đã và đang phát triển đáp ứng nhu cầu ngày càng cao của hàng tỷ người dùng trên toàn cầu. Hầu như tất cả các dịch vụ đa phương tiện tốc độ cao đều được truy nhập sử dụng qua các dịch vụ web, do đó nhu cầu băng thông của Internet tăng nhanh và cấp thiết. Internet cần phải đảm bảo cung cấp các dịch vụ với trễ nhỏ nhưng trong khi vẫn phải tiết kiệm băng thông. Đó là yêu cầu về hiệu năng của Internet. Cho đến nay có 3 giải pháp kiến trúc web caching mà đa số các ISP áp dụng nhằm đáp ứng được yêu cầu hiệu năng, đó là: kiến trúc phân cấp (hierarchical web caching architecture), kiến trúc web caching phân tán (distributed web caching) và kiến trúc web caching kết hợp (hybrid web caching). Các nghiên

cứ đánh giá hiệu năng của hệ thống web caching ở các bài báo [1], [2], [3], [4], [5] đều cho thấy ở từng cấp mạng phải có những nghiên cứu đánh giá riêng do sự khác nhau về lưu lượng. Tuy nhiên, hiện nay các nhà mạng đều sử dụng kiến trúc kết hợp. Tại từng cấp mạng thực hiện kiến trúc web caching phân tán, song không phải ở tất cả các nút có các hệ thống web cache, vì lý do đảm bảo tiết kiệm chi phí, mà chỉ ở những nút có yêu cầu băng thông cao do có số đông dân cư sử dụng mạng Internet. Và tất cả các hệ thống web cache như vậy liên kết ngang hàng trong một tầng mạng (kết nối P2P) với nhau để tăng mức độ sử dụng của hệ thống web caching ở từng cấp mạng. Kiến trúc web cache phân tầng kết hợp ở đây đảm bảo các hệ thống web cache của các tầng liên kết với nhau, như vậy đảm bảo tỷ lệ trùng web cao khi yêu cầu của các client chuyển lên tầng mạng trên và cũng làm tiết kiệm băng thông giữa các tầng mạng. Hình 1 là sơ đồ của kiến trúc web caching kết hợp của các tầng mạng liên kết với có 4 cấp chính.



**Hình 1.** Kiến trúc Internet web caching kết hợp

Cấp cao nhất của toàn kiến trúc web caching kết hợp là hệ thống web caching trung tâm (cấp thứ nhất) của mạng trực Internet quốc gia, ISP CC (Central Caches). Tại các mạng khu vực của mạng ISP (ISP Regional Network) có các hệ thống Web cache khu vực (cấp thứ hai), RC (Regional Caches). Cấp tiếp theo (cấp thứ ba) là các hệ thống web cache của các mạng địa phương, IC (Institutional Caches). Những người sử dụng đầu cuối (clients) ở cấp thứ 4. Chúng kết nối với các mạng truy nhập địa phương. Các trạm viễn thông tỉnh, thành phố có các nút POP (Network Point of Presence) là các nút truy nhập địa phương của mạng Internet. Tại các POP đặt các hệ thống web cache, IC (Institutional Caches). Các client có thể là một máy tính PC, một điện thoại di động, trực tiếp hay thông qua mạng LAN, kết nối với Internet qua các POP bằng các mạng truy nhập như Dial-up, ADSL, mạng di động. Các POP có các liên kết truyền dẫn tốc độ cao với các mạng khu vực. Các client bằng trình duyệt web có thể trực tiếp hoặc gián tiếp qua Proxy server cục bộ của LAN gửi yêu cầu về trang web qua mạng truy nhập đến mạng địa phương. Trong trường hợp Proxy server cục bộ không có nội dung của trang web yêu cầu,

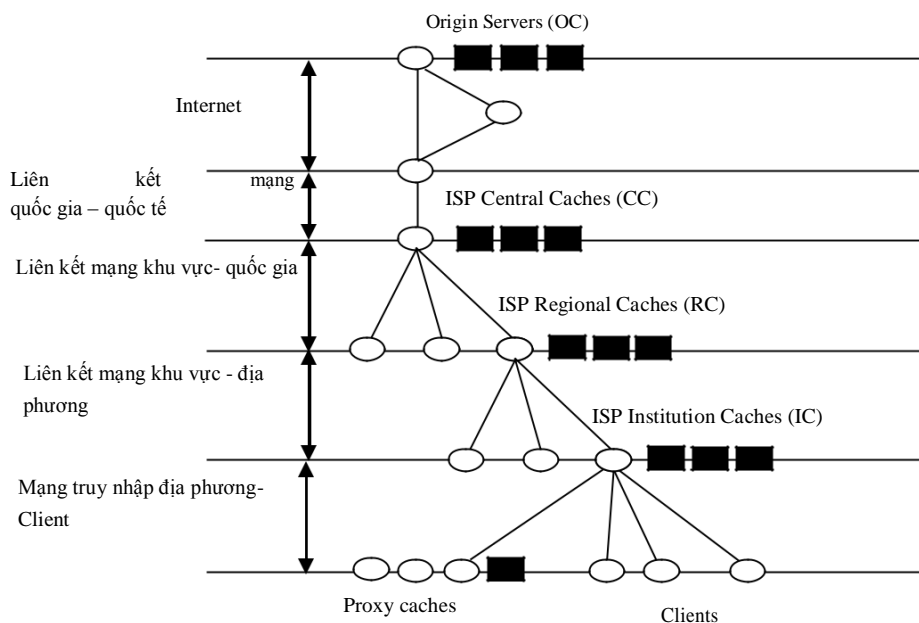
thì Proxy server chuyển yêu cầu của client đến hệ thống IC ở các POP địa phương. Nếu hệ thống IC có nội dung mà client yêu cầu (IC hit) thì IC chuyển nội dung yêu cầu về cho client (đồng thời Proxy server cục bộ cũng lưu nội dung trang web này). Khi trượt IC (IC miss) nghĩa là nội dung trang web mà client yêu cầu không có tại hệ thống IC, thì từ hệ thống IC yêu cầu được chuyển lên mạng cấp trên - mạng khu vực. Tại mạng khu vực, nếu hệ thống RC có nội dung yêu cầu (RC hit) thì nó chuyển nội dung về hệ thống IC, từ hệ thống IC chuyển tiếp về Proxy server và client. Nếu trượt RC (RC miss), yêu cầu của client được chuyển lên hệ thống CC của mạng quốc gia. Nếu hệ thống CC có nội dung thì nội dung (CC hit) được chuyển về hệ thống RC, rồi hệ thống IC, và đến Proxy server, đến client. Nếu trượt CC (CC miss), yêu cầu từ client được chuyển đến Internet quốc tế, đến web server gốc.

Đã có nhiều nghiên cứu tập trung vào cải thiện hiệu năng của Web dựa trên các thuật toán thay thế các trang web và các giải pháp web caching. Các yêu cầu nội dung web từ người dùng phải được đáp ứng với trễ nhỏ nhất ở các tầng mạng. Đáp ứng nhanh nhất đối với các yêu cầu

truy cập web phụ thuộc vào các xác suất trúng cache (hit rate). Nhưng khi yêu cầu trang web không có ở tầng mạng thì các yêu cầu phải được chuyển lên tầng mạng trên, như vậy trễ đáp ứng còn phụ thuộc cả vào các liên kết truyền thông giữa các tầng mạng. Do đó để có một kiến trúc

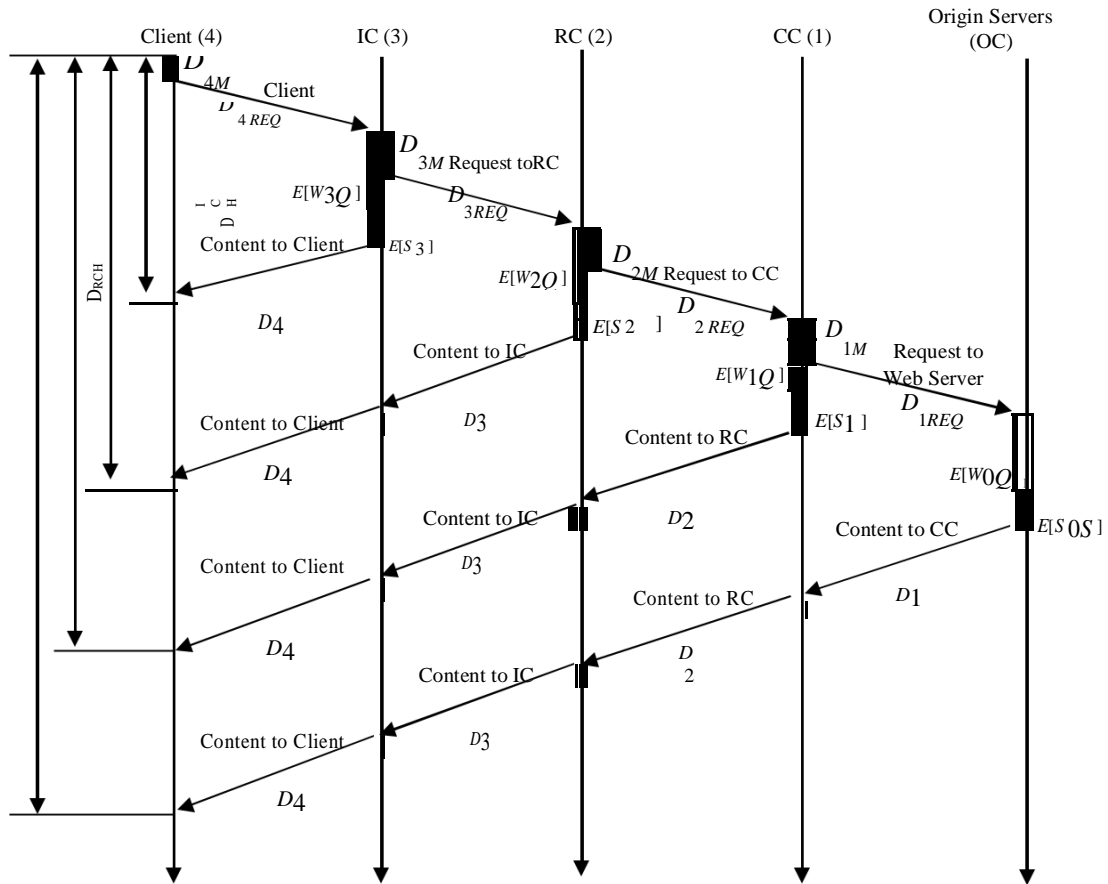
Internet web caching tối ưu về hiệu năng dựa trên hai thông số: trễ đáp ứng và chi phí băng thông kênh truyền dẫn, bài báo này đưa ra phương pháp phân tích hiệu năng của kiến trúc Internet web caching dựa trên mạng Petri [6].

## 2. Mô hình mạng của Internet web caching kết hợp



Hình 2. Mô hình cây của kiến trúc web caching phân tầng

Mô hình cây của kiến trúc Internet web caching kết hợp cho ở hình 2 và hình 3 là đồ thị diễn giải các trễ mà yêu cầu HTTP từ các client đến các web servers [7].



Hình 3. Đồ thị thời gian trễ của giao dịch HTTP của client trên mạng Internet với kiến trúc Web caching kết hợp

Tổng quát, nếu kiến trúc web caching của một mạng ISP có  $n$  cấp mạng, yêu cầu HTTP của client trượt web ở cấp mạng thứ  $n$ , trượt web ở hệ thống web caching ở cấp mạng thứ  $i$ , mà  $n > i$ ,

thì đáp ứng của hệ thống web caching ở cấp mạng thứ  $i$  cho yêu cầu HTTP của client sẽ bằng:

$$E[R_i] (D_{nM} \ D_{nREQ}) (D_{n-1M} \ D_{n-1REQ}) \dots (D_{i+1M} \ D_{i+1REQ}) E[C_i] D_{i-1} \dots D_{n-1} \ D_n \quad (1)$$

Trong đó:  $D_{nM}$  - trễ do trượt web ở cấp

mạng thứ  $n$ ;  $D_{nREQ}$  - trễ phụ thuộc băng thông kênh truyền dẫn mà yêu cầu HTTP của client (hay từ proxy server cục bộ) chuyển từ cấp mạng

thứ  $n$  đến cấp mạng thứ  $n - 1$ ;  $D_n$  - trễ trả về nội dung web yêu cầu cho client phụ thuộc băng thông kênh truyền dẫn từ cấp mạng thứ  $n - 1$  đến cấp

mạng thứ  $n$ , và phụ thuộc kích thước của nội dung web. Với kiến trúc 4 tầng thì ta có:

$D_{4M}$  - thời gian trượt web cục bộ tại mạng

của client, thời gian này phụ thuộc vào tốc độ của LAN (trong đó có proxy server) của client. Nếu client là một máy tính đơn lẻ không qua LAN thì thời gian này có thể bỏ qua.

$D_{4REQ}$  - thời gian mà yêu cầu HTTP của client được gửi đến mạng địa phương/cơ sở (đến

POP hoặc đến Router cơ sở). Thời gian này phụ thuộc vào tốc độ đường truyền kết nối client qua mạng truy nhập (dial-up, ADSL, vô tuyến, di động, đường truyền trực tiếp) tới các nút mạng địa phương/cơ sở, và phụ thuộc kích thước gói yêu cầu, trễ qua các nút mạng truy nhập trung gian của từng khu vực như quận, huyện đến các POP.

$D_4$  - trễ trả về client khi trúng web ở hệ thống IC (trúng web cấp 3), phụ thuộc vào băng thông kênh truyền dẫn và kích thước nội dung web trả về từ cấp mạng địa phương đến client ở mạng cấp 4.

$D_{ICH}$  - thời gian đáp ứng trung bình của hệ thống web caching IC khi trúng web ở IC (IC hit). Thời gian này bao gồm: thời gian tổng thời gian ( $D_{4M} + D_{4REQ}$ ), thời gian đáp ứng trung

bình của hệ thống web caching ở cấp mạng 3,  $E(C_3)$  và thời gian trả nội dung web  $D_4$ .

$$\left. \begin{aligned} D_{ICH} &= (D_{4M} + D_{4REQ}) E[C_3] + D_4 \\ &+ (D_{4M} + D_{4REQ}) E[W_{3Q}] + E[S_3] + D_4 \\ &+ (D_{4M} + D_{4REQ}) \frac{E[N_{3Q}]}{1} + \frac{1}{3} D_4 \end{aligned} \right\} (2)$$

$D_{RCH}$  - đáp ứng trung bình của hệ thống web caching ở mạng khu vực khi trúng RC (RC hit):

$$\left. \begin{aligned} D_{RCH} &= (D_{4M} + D_{4REQ}) (D_{3M} + D_{3REQ}) E[C_2] + D_3 + D_4 \\ &+ (D_{4REQ} + D_{4M}) (D_{3REQ} + D_{3M}) E[W_{2Q}] + E[S_2] + D_3 + D_4 \\ &+ (D_{4M} + D_{4REQ}) (D_{3M} + D_{3REQ}) \frac{E[N_{2Q}]}{2} + \frac{1}{2} D_3 + D_4 \end{aligned} \right\} (3)$$

$D_{CCH}$  - đáp ứng trung bình của hệ thống web caching ở mạng quốc gia khi trúng CC (CC hit):

$$\left. \begin{aligned} D_{CCH} &= (D_{4M} + D_{4REQ}) (D_{3M} + D_{3REQ}) (D_{2M} + D_{2REQ}) E[C_1] + D_2 + D_3 + D_4 \\ &+ (D_{4REQ} + D_{4M}) (D_{3REQ} + D_{3M}) (D_{2REQ} + D_{2M}) \\ &+ (D_{4M} + D_{4REQ}) (D_{3M} + D_{3REQ}) (D_{2M} + D_{2REQ}) E[W_{1Q}] + E[S_1] + D_2 + D_3 + D_4 \\ &+ (D_{4M} + D_{4REQ}) (D_{3M} + D_{3REQ}) (D_{2M} + D_{2REQ}) \frac{E[N_{1Q}]}{1} + \frac{1}{1} D_2 + D_3 + D_4 \end{aligned} \right\} (4)$$

$D_{WSH}$  - đáp ứng trung bình của Internet quốc tế và web server nguồn:

$$\left. \begin{aligned} D_{WSH} &= (D_{4M} + D_{4REQ}) (D_{3M} + D_{3REQ}) (D_{2M} + D_{2REQ}) (D_{1M} + D_{1REQ}) E[C_0] + D_1 + D_2 + D_3 + D_4 \\ &+ (D_{4REQ} + D_{4M}) (D_{3REQ} + D_{3M}) (D_{2REQ} + D_{2M}) \\ &+ (D_{4M} + D_{4REQ}) (D_{3M} + D_{3REQ}) (D_{2M} + D_{2REQ}) E[W_{0Q}] + E[S_0] + D_1 + D_2 + D_3 + D_4 \\ &+ (D_{4M} + D_{4REQ}) (D_{3M} + D_{3REQ}) (D_{2M} + D_{2REQ}) \\ &+ (D_{4M} + D_{4REQ}) \frac{E[N_{0Q}]}{0} + \frac{1}{0} D_1 + D_2 + D_3 + D_4 \end{aligned} \right\} (5)$$

Như vậy, trường hợp xấu nhất là không trúng web yêu cầu ở tất cả các cấp mạng của ISP trong quốc

gia và chỉ trúng web trên cấp mạng Internet quốc tế ở web server nguồn. Công thức (5) là công thức tổng quát để tính trễ truy nhập web cho trường hợp xấu nhất này. Các công thức trên đây [7] cho thấy sự phụ thuộc vào từng hệ thống web caching ở từng cấp mạng: các liên kết ngang hàng, số lượng nút hệ thống web caching, giao thức thay thế web cache, các liên kết của các hệ thống web caching của các tầng mạng với nhau, và số lượng tầng mạng.

### 3. Mô hình SCPN kiến trúc Internet web caching

Mạng Petri có màu và thời gian ngẫu nhiên chung SCPN (Stochastic Colored Petri Net) được định nghĩa bởi 9 bộ thông số [6]:

$SCP_N (P, T, A, N, C, E, G, I)$ ; trong đó: P - tập hợp các vị trí có màu (có đặc tính dữ liệu) được ký hiệu bằng các vòng tròn với các thẻ (đầu chấm đen); T - tập hợp các chuyển tiếp (có thể là hàm thời gian) hoặc không có thời gian (kích hoạt ngay khi thỏa mãn điều kiện các vị trí vào của chúng đều có thẻ), được biểu diễn là các thanh chữ nhật trắng (có thời gian) hoặc thanh đen (không có thời gian); A - tập hợp các cung thỏa mãn:

$P \rightarrow T \rightarrow P \rightarrow A \rightarrow T \rightarrow A$ ; tập hợp hữu hạn của các tập hợp màu xác định trong SCPN. Tập hợp này chứa tất cả các màu có thể thẻ, các phép toán và các hàm được sử dụng trong SCPN; N - hàm nút, được xác định từ A vào trong  $P \rightarrow T \rightarrow P$ ; C - hàm màu, được xác định từ P vào trong và ánh xạ các vị trí  $p \in P$  vào trong các màu trong tập hợp, được viết là  $C: P; G$  - hàm giám sát (guard function);

$G$  ánh xạ từng chuyển tiếp  $t \in T$  vào biểu thức giám sát. Biểu thức giám sát phải có giá trị boolean true hay false, và được viết là:

$$t: T: Type(G(t)) \text{ BooleanType}(Var(G(t)))$$

Trong đó  $Type(Vars)$  chỉ tập hợp các kiểu  $Type(v) | v \in Vars$ ,  $Vars$  là tập hợp các biến,  $Vars(G(t))$  chỉ các biến sử dụng trong  $g(t)$ .

$E$  - hàm biểu thức của cung.  $E$  ánh xạ từng cung  $a \in A$  vào trong biểu thức  $E$ . Các kiểu của

vào và ra của các biểu thức cung phải tương ứng với kiểu của các nút được nối bởi cung:

$$a \quad A : Type(E(a)) \quad C(p(a))_{MS} \quad Type(var(E(a)))$$

Trong đó,  $p(a)$  là vị trí trong  $A(a)$  và  $C(p)_{MS}$  là tập hợp hữu hạn các multiset.

$I$ -hàm khởi tạo.  $I$  ánh xạ từng vị trí  $p$  vào trong các biểu thức khởi tạo (không có các biến) như rằng:

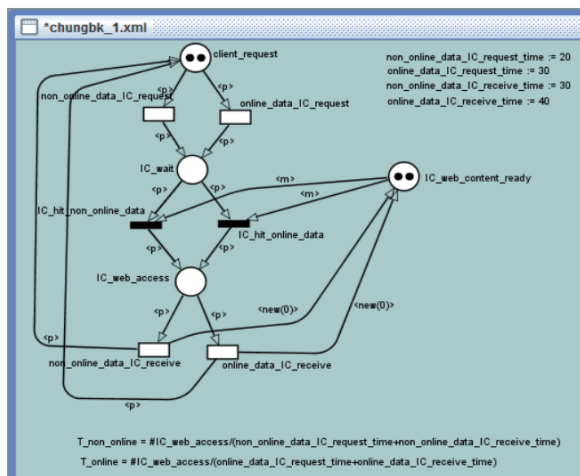
$$p \quad P : Type(I(p)) \quad C(p)_{MS} \quad Var(I(p))$$

Đối với chuyển tiếp có thời gian, khi được phép (nếu các vị trí vào của nó đều có thẻ), nó bắt đầu kích hoạt, lấy các thẻ ra khỏi các vị trí vào. Tuy nhiên các thẻ chỉ được đưa vào các vị trí ra của chuyển tiếp ngay khi thẻ kết thúc. Các thẻ bị “giữ trong chuyển tiếp” trong khoảng trễ kích hoạt.

Kết quả nghiên cứu trong [7], [8] có đề xuất sử dụng mạng hàng đợi để phân tích hiệu năng của kiến trúc Internet web

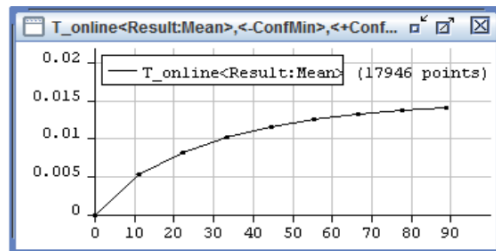
này đề xuất sử dụng SCPN để mô hình hệ thống web caching và phân tích hiệu năng của hệ thống web caching; TimeNET là phần mềm được sử dụng để mô phỏng mô hình đề xuất trên.

Hình 2(a) là mô hình SCPN hệ thống Internet web caching ở lớp IC và kịch bản cho rằng http client yêu cầu nội dung web có ngay ở IC, nghĩa là có IC-hit. Vị trí client\_request với 2 loại thẻ: thẻ 1: thẻ hiện loại dữ liệu không online (non\_online\_data) và thẻ 2: thẻ hiện loại dữ liệu online (online\_data). Hình 2(b) và (c) là kết quả xác định các thông số  $T_{non\_online}$  và  $T_{online} = (\text{số lượng truy nhập IC})/(\text{trễ trung bình ở cấp mạng IC})$  theo  $\theta$  (tọa độ Y) và chạy trong khoảng 9 seconds với số lần kích hoạt thẻ hiện số yêu cầu web từ client (tọa độ X) đối với non\_online\_data và online\_data.

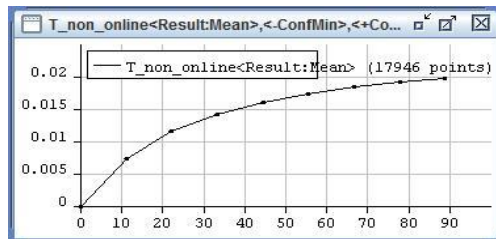


(a)

caching kết hợp. Bài báo

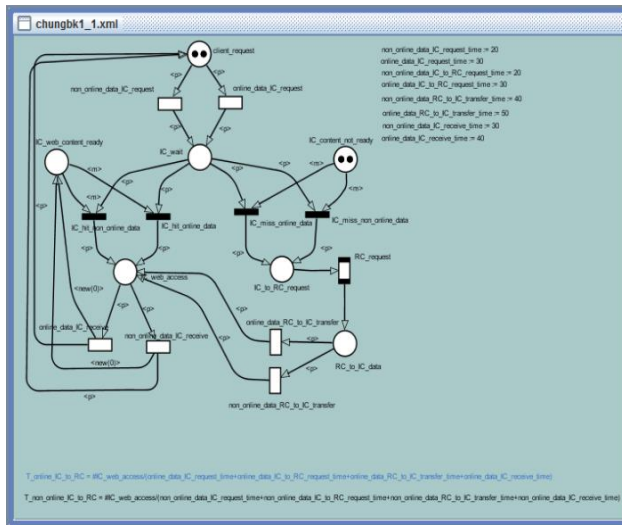


(b)

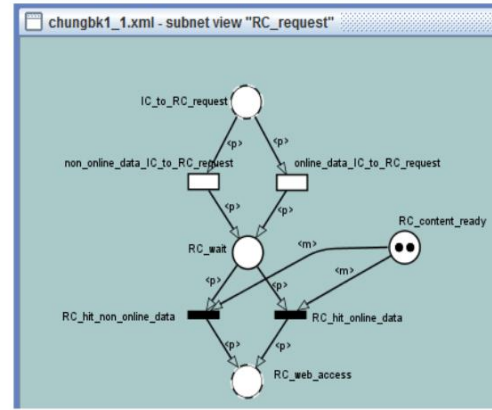


(c)

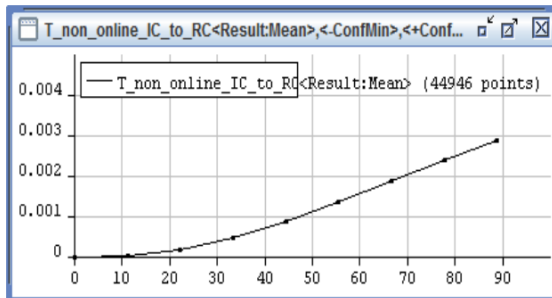
Hình 4. (a) mạng SCPN của cấp web caching cache IC khi có IC-hit  
(b) Trễ nhận nội dung web với non-online-data, (c) Trễ nhận nội dung web với online-data



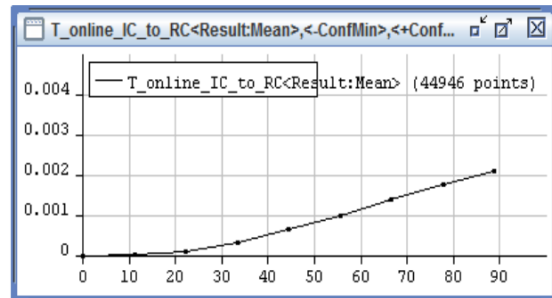
(a)



(b)



(c)



(d)

**Hình 5.** (a) mạng SCPN của hai cấp webchinh cache IC-RC khi có IC-miss và RC-hit  
(b) Trễ nhận nội dung web với non-online-data, (c) Trễ nhận nội dung web với online-data

Hình 5 (a) là mô hình SCPN hệ thống Internet hai cấp web caching IC-RC khi IC-miss và RC-hit, trong đó vị trí RC-request thể hiện mô hình SCPN con cho RC (hình 5(b)). Hình 5(c) và (d) là kết quả xác định các thông số  $T_{non\_online\_IC\_to\_RC}$  và  $T_{online\_IC\_to\_RC} = (\text{số lượng truy nhập IC}) / (\text{tổng trễ trung bình của cấp mạng IC và RC})$  theo ms (tọa độ Y) và chạy trong khoảng 64 seconds với số lần kích hoạt thể hiện số yêu cầu web từ client (tọa độ X) đối với non\_online\_data và online\_data.

Có thể bằng cách tương tự xác định được các mô hình SCPN cho hệ thống 3 cấp web

caching IC-RC-CC khi có IC-miss, RC-miss, CC-hit, và hệ thống 4 cấp web caching IC-RC-CC-OC khi IC-miss, RC-miss, CC-miss, và OC-hit.

#### 4. Phân tích và kết luận

Với một cấp web caching, ví dụ IC (hình 4), kết quả mô phỏng phụ thuộc vào số yêu cầu web từ client và trễ trung bình ở các cấp web caching. Đối với dữ liệu online, thông thường có độ dài lớn nên chúng có chi phí trễ trung bình lớn hơn so với dữ liệu non-online (thể hiện ở các thông số đặt trước trên các hình 4(a) và 5(a)). Do đó thông số hiệu năng  $T_{online} < T_{non\_online}$  nhận được nội

dung web yêu cầu là nhỏ nhất. Với hai cấp webcaching IC-RC khi IC không sẵn sàng nội dung (vị trí IC\_content\_not\_ready có thể) kết quả mô phỏng ở hình 5(c) và (d) cho thấy ban đầu vì tỷ số IC-miss và nội dung web cần được lấy từ cấp RC với trễ truy nhập lớn nên giá trị  $T_{online\_IC\_to\_RC} < T_{online\_IC}$  và  $T_{non\_online\_IC\_to\_RC} < T_{non\_online}$ . Với sự tăng số thẻ từ RC đến vị trí IC\_web\_content\_ready, tỷ số IC\_hit cao lên, do đó trễ truy nhập giảm, khi đó  $T_{online\_IC\_to\_RC} > T_{online\_IC}$  và  $T_{non\_online\_IC\_to\_RC} >$

$T_{non\_online}$ . Như vậy, với số cấp web caching tăng ta có thể giảm trễ truy nhập Internet, trong khi không cần phải tăng dung lượng băng thông của từng cấp mạng.

Với phương pháp mô hình hóa bằng SCPN và đưa vào các đặc tính về dữ liệu cho các loại thẻ trong các vị trí như mức ưu tiên, độ dài dữ liệu (hay số lượng các gói tin), và thay đổi các thông số về trễ trung bình của các cấp mạng web caching, ta có thể xác định các thông số hiệu năng khác để đánh giá hiệu năng của kiến trúc Internet web caching.

### Tài liệu tham khảo

- [1]. Carey Williamson, Mudashiru Busari: Simulation Evaluation of Web Caching Architectures, M.Sc. Thesis, June 2000, Department of Computer science, University of Saskatchewan, <http://www.cs.usask.ca/faculty/carey/>.
- [2]. Haohuan Fu, Pui-On Au, Weijia Jia, 2004. Performance Evaluation of Replacement Algorithms In Hierarchical Web Caching. Book series Lecture notes in computer science, Publisher "Springer Berlin/Heidelberg", ISSN 0302-9743(print) 1611-3349 (online), volume 3129/2004.
- [3]. A. Rousskov, "On Performance of Caching Proxies", In ACM SIGMETRICS, Madison, USA, September 1998.
- [4]. C. Maltzahn, J.Richardson, "Performance Issues of Enterprise Level Web Proxies", 1998
- [5]. M. Deshpande, G. Karypis, " Selective Markov models for predicting Web page access". ACM Transactions on Internert technology, May 2004.
- [6]. Hồ Khánh Lâm, "Mạng Petri: lý thuyết và ứng dụng". Trường Đại học Sư phạm Kỹ thuật Hưng Yên, NXB KHKT, 2015.
- [7]. Ho Khanh Lam, Nguyen Xuan Truong, "Performance Analysis of Hybrid Web Caching Architecture". American Journal of Networks and Communications. 2015; 4(3): 37-43. Published online April 24, 2015 (<http://www.sciencepublishinggroup.com/j/ajnc>). doi: 10.11648/j.ajnc.20150403.13. ISSN: 2326-893X (Print); ISSN: 2326-8964 (Online)
- [8]. Hồ Khánh Lâm, Nguyễn Xuân Trường, "Sử dụng mạng hàng đợi phân tích ảnh hưởng của các client proxy server trong kiến trúc web caching". Tạp chí Khoa học và Công nghệ - Đại học Thái Nguyên Tập 137, số 07, 2015. ISSN: 1859 - 2171.



**ANALYSIS PERFORMANCE OF INTERNET WEB CACHING BASED ON USED PETRI  
NETWORK MODEL HAS COLORS AND RONDON TIME (SCPN)**

**Abstract:**

*At present, the communication network technology developed rapidly provide high speed multimedia services to serve the demand of web services 's users. Web caching is an application-level routing and bandwidth for most of the web with the goalsare increasing the line speed and the web access speed. Currently, there are several different methods to evaluate the performance of internet web caching architecture. However, modeling methods use colored Petri network and random time (SCPN) is a completely new method helps to identify and adjust parameters to optimize performance of internet web caching architecture based on two parameters: delay and cost to meet the transmission bandwidth channel. This paper gives a method for analyzing the performance of Internet web caching architecture oriented based approach Petri net aims to optimize web caching architecture to improve web access speed and use multimedia services on the internet.*

**Keywords:** *web caching, Petri network (PN), Petri network model has colors and random time (SCPN).*