



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0041066

(51)^{2021.01} G06F 16/951; G06Q 50/10

(13) B

(21) 1-2022-04688

(22) 08/01/2021

(86) PCT/KR2021/000258 08/01/2021

(87) WO 2021/141441 15/07/2021

(30) 10-2020-0003063 09/01/2020 KR

(45) 25/09/2024 438

(43) 25/10/2022 415

(73) KAKAOBANK CORP. (KR)

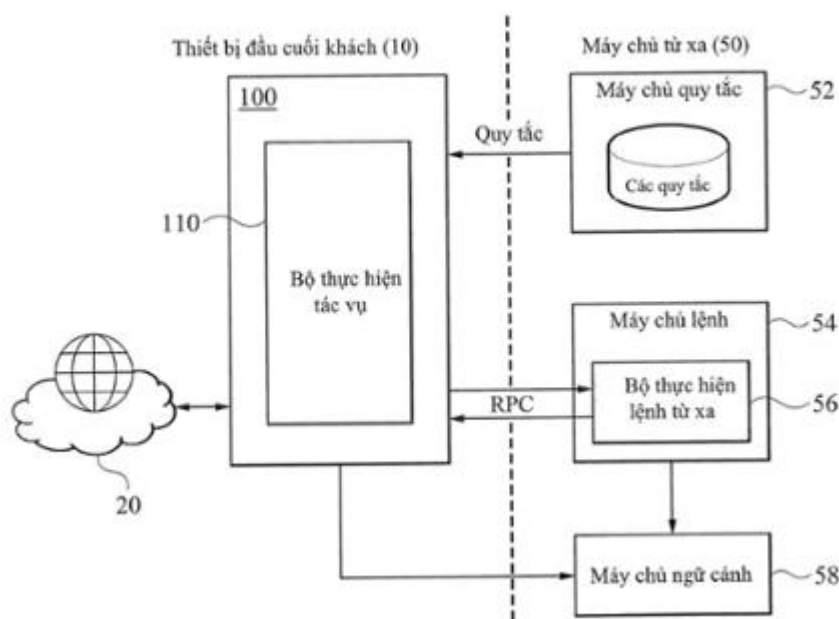
5F, 231, Pangyoyeok-ro, Bundang-gu, Seongnam-si, Gyeonggi-do 13494, Republic of Korea

(72) KIM, Dohyoung Keedi (KR); KIM, Hyeon Seung (KR).

(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)

(54) HỆ THỐNG CUNG CẤP DỊCH VỤ DỰA TRÊN CÀO DỮ LIỆU VÀ MÁY CHỦ TỪ XA

(57) Sáng chế đề cập đến hệ thống cung cấp dịch vụ dựa trên cao dữ liệu. Phương pháp cung cấp dịch vụ dựa trên cao dữ liệu bao gồm các bước: yêu cầu và nhận ít nhất một trong số thông tin quy tắc cao dữ liệu xác định một loạt tác vụ để thu dữ liệu bên ngoài; thu thập dữ liệu bên ngoài bằng cách thực hiện các tác vụ theo thông tin quy tắc đã nhận; và cung cấp thông tin dịch vụ cho người dùng bằng cách sử dụng dữ liệu bên ngoài đã thu thập. Theo sáng chế, nhờ cung cấp dịch vụ cao dữ liệu thông qua thiết bị đầu cuối khách và máy chủ từ xa theo cách phân phối, có thể tăng hiệu quả của tác vụ cao dữ liệu và đối phó tích cực với các thay đổi trong đích cao dữ liệu.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp cung cấp dịch vụ dựa trên cào dữ liệu (scraping).

Tình trạng kỹ thuật của sáng chế

Cào dữ liệu là công nghệ trích xuất dữ liệu cần thiết để cung cấp thông tin trong số dữ liệu được phân phối và cung cấp dữ liệu được trích xuất cho người dùng, và cho phép thông tin được phân phối ở các địa điểm khác nhau như các công ty tài chính, tổ chức công cộng và trang web chính phủ cần được tích hợp và được tham chiếu đến và được quản lý bằng cách sử dụng thông tin xác thực của khách hàng.

Cào dữ liệu màn hình, là một trong số các phương pháp cào, thu thập và cung cấp dữ liệu được hiển thị trên màn hình thông qua thiết bị hiển thị, và còn được gọi là cào dữ liệu web vì nó thu thập và cung cấp thông tin thường được hiển thị trên các trang web hoặc dạng tương tự.

Các công nghệ cào dữ liệu đã được sử dụng chủ yếu trong lĩnh vực quản lý tài sản và chi tiêu của các cá nhân được phân phối trong các tổ chức tài chính tương ứng, và đã được sử dụng để tích hợp và quản lý phí dịch vụ hoặc tích hợp và tham chiếu đến các thư điện tử và cung cấp các thư điện tử này cho người dùng.

Nghĩa là, khi cào dữ liệu được sử dụng, thì có ưu điểm là thông tin được phân phối đến các thực thể quản lý thông tin tương ứng có thể được thu thập và được quản lý.

Thực thể cung cấp dịch vụ thông qua việc cào dữ liệu cần phải đối phó theo cách riêng lẻ và liên tục với sự thay đổi của phương pháp xác thực hoặc các yêu cầu hệ thống trong trang web trong đó thông tin đích cào dữ liệu được quản lý.

Ứng dụng cung cấp dịch vụ thông qua việc cào dữ liệu thông thường đã được vận hành theo cách tạo cấu hình mô đun cào dữ liệu riêng biệt cho mỗi trang web đích để đối

phó trực tiếp với các thay đổi trong các trang web đích, giám sát liên tục các thay đổi và cập nhật mô đun cào dữ liệu này theo sự thay đổi trong trang web đích.

Sự thay đổi này của mô đun cào dữ liệu gây ra vấn đề là khi nhiều trang web đích cào dữ liệu tồn tại để cung cấp một dịch vụ, ứng dụng cần được cải biến và phân phối lại mỗi khi có sự thay đổi của mỗi trang web.

Một vấn đề là ứng dụng được cài đặt cần được cài đặt lại mỗi lần để sử dụng dịch vụ xảy ra, và do đó, người dùng cảm thấy bất tiện trong khi sử dụng dịch vụ.

Theo đó, để giải quyết vấn đề này, cần đưa ra phương pháp cung cấp dịch vụ dựa trên cào dữ liệu bằng cách đối phó tích cực với tình huống thay đổi của đích cào dữ liệu và giảm thiểu việc cập nhật ứng dụng.

Bản chất kỹ thuật của sáng chế

Vấn đề kỹ thuật

Mục đích của sáng chế là đề xuất phương pháp cung cấp dịch vụ cào dữ liệu một cách hiệu quả.

Chi tiết hơn là, mục đích của sáng chế là đề xuất phương pháp đối phó tích cực với sự thay đổi của thực thể đích cào dữ liệu và thực hiện tác vụ cào dữ liệu một cách hiệu quả hơn.

Ngoài ra, mục đích của sáng chế là đề xuất phương pháp tạo cấu hình thư viện dùng cho dịch vụ cào dữ liệu, và theo đó, làm giảm việc cập nhật hoặc phân phối lại ứng dụng riêng biệt.

Giải pháp kỹ thuật

Theo khía cạnh của sáng chế, phương pháp cung cấp dịch vụ dựa trên cào dữ liệu bao gồm các bước: yêu cầu và nhận ít nhất một trong số thông tin quy tắc cào dữ liệu xác định một loạt tác vụ để thu dữ liệu bên ngoài; thu thập dữ liệu bên ngoài bằng cách thực hiện các tác vụ theo thông tin quy tắc nhận được; và cung cấp thông tin dịch vụ cho người dùng bằng cách sử dụng dữ liệu bên ngoài đã thu thập.

Ở bước thu thập dữ liệu bên ngoài, dữ liệu bên ngoài có thể được thu thập bằng cách thực hiện các lệnh cấu thành các tác vụ được xác định trong thông tin quy tắc trên thiết bị đầu cuối khách hoặc máy chủ từ xa.

Tác vụ có thể bao gồm tổ hợp của các lệnh cục bộ được thực hiện trên thiết bị đầu cuối khách và được tạo ra trước.

Tác vụ có thể còn bao gồm lệnh từ xa tùy chỉnh được thực hiện trên máy chủ từ xa.

Thiết bị đầu cuối khách có thể truyền cuộc gọi thủ tục từ xa đến máy chủ từ xa để thực hiện lệnh từ xa, và máy chủ từ xa có thể trả về kết quả thực hiện của tác vụ được thực hiện theo cuộc gọi thủ tục từ xa cho thiết bị đầu cuối khách.

Ở bước thu thập dữ liệu bên ngoài, kết quả thực hiện của tác vụ được thực hiện có thể được xác minh để thực hiện mỗi tác vụ, và việc sửa đổi thông tin quy tắc hoặc cuộc gọi thủ tục từ xa đối với lệnh có thể được yêu cầu khi có sự bất thường khi thực hiện tác vụ.

Khi các lệnh từ xa là liên lục với số lần định trước hoặc nhiều hơn theo lịch trình thực hiện của tác vụ, thì thiết bị đầu cuối khách có thể cung cấp quyền điều khiển để điều khiển việc thực hiện lệnh cho máy chủ từ xa, và nhận việc trả về kết quả thực hiện của các lệnh từ xa được thực hiện liên tục và quyền điều khiển từ máy chủ từ xa mà có quyền điều khiển đã nhận.

Các tác vụ có thể được thực hiện theo các đơn vị khối, sao cho các kết quả thực hiện của các lệnh trong các tác vụ được lưu trữ trong các khối, và các kết quả thực hiện của các tác vụ có thể được tham chiếu bằng cách sử dụng các tên tác vụ và các biến liên quan ở thời điểm tham chiếu các tác vụ.

Lệnh cục bộ và lệnh từ xa có thể được thực hiện theo cách phân phối thông qua thiết bị đầu cuối khách hoặc máy chủ từ xa, và tác vụ có thể được thăm dò theo cách không đồng bộ.

Theo một khía cạnh khác của sáng chế, ứng dụng để cung cấp dịch vụ dựa trên

cào dữ liệu bao gồm: bộ yêu cầu quy tắc cào dữ liệu được tạo cấu hình để yêu cầu và nhận ít nhất một trong số thông tin quy tắc cào dữ liệu xác định một loạt tác vụ để thu dữ liệu bên ngoài; bộ thực hiện tác vụ được tạo cấu hình để thu thập dữ liệu bên ngoài bằng cách thực hiện các tác vụ theo thông tin quy tắc đã nhận; và nhà cung cấp dịch vụ được tạo cấu hình để cung cấp thông tin dịch vụ cho người dùng bằng cách sử dụng dữ liệu bên ngoài đã thu thập.

Bộ thực hiện tác vụ có thể được tạo cấu hình để thu thập dữ liệu bên ngoài bằng cách thực hiện các lệnh cấu thành các tác vụ được xác định trong thông tin quy tắc trên thiết bị đầu cuối khách hoặc máy chủ từ xa.

Bộ thực hiện tác vụ có thể còn bao gồm: bộ phát cuộc gọi thủ tục từ xa (remote procedure call, RPC) được tạo cấu hình để truyền cuộc gọi thủ tục từ xa đến máy chủ từ xa để thực hiện lệnh từ xa; và bộ thu kết quả lệnh được tạo cấu hình để nhận việc trả về kết quả thực hiện của tác vụ được thực hiện theo cuộc gọi thủ tục từ xa từ máy chủ từ xa.

Theo một khía cạnh khác nữa của sáng chế, phương pháp cung cấp từ xa dịch vụ dựa trên cào dữ liệu bao gồm các bước: quản lý thông tin quy tắc cào dữ liệu xác định một loạt tác vụ để thu dữ liệu bên ngoài và truyền thông tin quy tắc theo yêu cầu từ thiết bị đầu cuối khách; thực hiện lệnh từ xa mà cần được thực hiện từ xa trong khi thiết bị đầu cuối khách thực hiện các tác vụ được xác định trong thông tin quy tắc đã truyền, theo cuộc gọi thủ tục từ xa; và trả về kết quả thực hiện của lệnh từ xa cho thiết bị đầu cuối khách.

Ở bước thực hiện lệnh từ xa, khi các lệnh từ xa là liên tục với số lần định trước hoặc nhiều hơn theo lịch trình thực hiện của tác vụ, thì quyền điều khiển để điều khiển việc thực hiện lệnh có thể nhận được từ thiết bị đầu cuối khách, và các lệnh từ xa có thể được thực hiện liên tục.

Ở bước trả về kết quả thực hiện, kết quả thực hiện của các lệnh từ xa được thực hiện liên tục trong máy chủ từ xa mà đã nhận được quyền điều khiển và quyền điều khiển này có thể được trả về.

Theo một khía cạnh khác nữa của sáng chế, máy chủ từ xa để cung cấp dịch vụ

dựa trên cào dữ liệu bao gồm: máy chủ quy tắc được tạo cấu hình để quản lý thông tin quy tắc cào dữ liệu xác định một loạt tác vụ để thu dữ liệu bên ngoài và truyền thông tin quy tắc theo yêu cầu từ thiết bị đầu cuối khách; và máy chủ lệnh được tạo cấu hình để thực hiện lệnh mà cần được thực hiện từ xa trong khi thiết bị đầu cuối khách thực hiện các tác vụ được xác định trong thông tin quy tắc đã truyền, theo cuộc gọi thủ tục từ xa, và trả về kết quả thực hiện của lệnh.

Máy chủ lệnh có thể nhận quyền điều khiển để điều khiển việc thực hiện lệnh từ thiết bị đầu cuối khách và thực hiện các lệnh từ xa liên tục, khi các lệnh từ xa là liên tục với số lần định trước hoặc nhiều hơn theo lịch trình thực hiện của tác vụ.

Máy chủ lệnh có thể trả về kết quả thực hiện của các lệnh từ xa được thực hiện liên tục trong máy chủ từ xa mà đã nhận được quyền điều khiển và quyền điều khiển.

Hiệu quả đạt được bởi sáng chế

Theo sáng chế, bằng cách cung cấp dịch vụ cào dữ liệu thông qua thiết bị đầu cuối khách và máy chủ từ xa theo cách phân phối, có thể tăng hiệu quả của tác vụ cào dữ liệu và đối phó tích cực với các thay đổi trong đích cào dữ liệu.

Ngoài ra, có thể tăng mức độ thuận tiện cho người dùng khi sử dụng bằng cách giảm thiểu sự cập nhật hoặc phân phối lại ứng dụng đáp lại sự thay đổi của đích cào dữ liệu.

Hơn nữa, có thể tăng độ bảo mật bằng cách vận hành máy chủ từ xa thực hiện cào dữ liệu một cách độc lập từ thiết bị đầu cuối khách và ngăn ngừa sự truy cập từ bên ngoài vào và sự rò rỉ của thông tin được sử dụng để thực hiện việc cào dữ liệu.

Mô tả vắn tắt các hình vẽ

Fig.1 là sơ đồ khối minh họa cấu hình của hệ thống cung cấp dịch vụ dựa trên cào dữ liệu theo một phương án của sáng chế;

Fig.2 là sơ đồ khối minh họa hệ thống bao gồm thiết bị đầu cuối khách thực hiện việc cào dữ liệu và máy chủ từ xa cung cấp quy trình cào dữ liệu theo một phương án của sáng chế;

Fig.3 là sơ đồ trình tự minh họa quy trình yêu cầu thông tin quy tắc giữa thiết bị đầu cuối khách và máy chủ quy tắc theo một phương án của sáng chế;

Fig.4 là sơ đồ trình tự minh họa quy trình thực hiện tác vụ của thiết bị đầu cuối khách theo một phương án của sáng chế;

Fig.5 là lưu đồ minh họa phương pháp trong đó ứng dụng thực hiện dịch vụ cào dữ liệu trên thiết bị đầu cuối khách theo một phương án của sáng chế;

Fig.6 là sơ đồ trình tự minh họa quy trình thực hiện tác vụ của thiết bị đầu cuối khách theo một phương án của sáng chế;

Fig.7A và Fig.7B là các sơ đồ minh họa mối quan hệ giữa và thứ tự thực hiện của các tác vụ theo một phương án của sáng chế; và

Fig.8 là sơ đồ khối minh họa cấu hình của ứng dụng theo một phương án của sáng chế.

Mô tả chi tiết sáng chế

Phần mô tả sau chỉ lấy ví dụ nguyên lý của sáng chế. Do đó, người có hiểu biết trung bình trong lĩnh vực kỹ thuật có thể sáng chế ra các thiết bị khác nhau thực hiện nguyên lý của sáng chế và nằm trong mục đích và phạm vi của sáng chế mặc dù không được mô tả hoặc minh họa một cách rõ ràng trong bản mô tả này. Ngoài ra, cần hiểu rằng tất cả các mục điều kiện và các phương án được đề cập trong bản mô tả này rõ ràng chỉ nhằm mục đích cho phép người có hiểu biết trung bình trong lĩnh vực kỹ thuật hiểu khái niệm của sáng chế về nguyên lý, và sáng chế không bị giới hạn ở các phương án và các tình trạng được đề cập cụ thể như vậy.

Các đối tượng, dấu hiệu và ưu điểm được mô tả ở trên sẽ trở nên rõ ràng hơn nhờ phần mô tả chi tiết sau được đề xuất liên quan đến các hình vẽ kèm theo, và theo đó, người có hiểu biết trung bình trong lĩnh vực kỹ thuật mà sáng chế thuộc về có thể thực hiện dễ dàng mục đích kỹ thuật của sáng chế.

Ngoài ra, trong quá trình mô tả sáng chế, khi xác định được rằng phần mô tả chi tiết về công nghệ đã biết liên quan đến sáng chế có thể khó hiểu một cách không cần

thiết ý chính của sáng chế, phần mô tả chi tiết sẽ được lược bỏ. Sau đây, một phương án của sáng chế sẽ được mô tả một cách chi tiết dựa vào các hình vẽ kèm theo.

Fig.1 là sơ đồ khối minh họa cấu hình của hệ thống cung cấp dịch vụ dựa trên cào dữ liệu theo một phương án của sáng chế.

Theo một phương án, hệ thống có thể bao gồm thiết bị đầu cuối khách 10 cung cấp trực tiếp dịch vụ cho người dùng và thông tin dựa trên Internet cung cấp thực thể 20 như trang web, máy chủ web hoặc ứng dụng mà cung cấp trang đích cào dữ liệu.

Thiết bị đầu cuối khách 10 là phương tiện để cung cấp dịch vụ được tích hợp cho người dùng bằng cách trích xuất dữ liệu cần thiết của dữ liệu khác nhau được phân phối trên các trang web, và có thể thực hiện việc cào dữ liệu màn hình.

Người dùng cài đặt ứng dụng 1000 để cung cấp dịch vụ cho thiết bị đầu cuối khách 10, và ứng dụng 1000 có thể tạo cấu hình môđun dưới dạng thư viện 100 để cung cấp chức năng mong muốn thông qua thiết bị đầu cuối khách 10, và cung cấp dịch vụ mong muốn cho người dùng.

Theo một phương án, ứng dụng 1000 có thể tạo cấu hình tập hợp môđun thực hiện các tác vụ để cung cấp dịch vụ cụ thể dựa trên cào dữ liệu cho người dùng dưới dạng thư viện 100.

Môđun cụ thể trong thư viện 100 có thể thực hiện tuần tự các quy trình truy cập trang web 20 bên ngoài, thực hiện việc đăng nhập hoặc xác thực công khai người dùng khi người dùng sử dụng trực tiếp trang web 20 và thu dữ liệu.

Sau đây, phương pháp thực hiện tác vụ của thư viện 100 theo một phương án sẽ được mô tả một cách chi tiết hơn dựa vào Fig.2.

Fig.2 là sơ đồ khối minh họa hệ thống bao gồm thiết bị đầu cuối khách 10 thực hiện việc cào dữ liệu và máy chủ từ xa 50 cung cấp quy trình cào dữ liệu theo một phương án của sáng chế.

Thư viện 100 được tạo cấu hình trên thiết bị đầu cuối khách 10 có thể bao gồm bộ thực hiện tác vụ 110 mà thực hiện tác vụ để cào dữ liệu.

Bộ thực hiện tác vụ 110 truy cập vào trang web 20 bên ngoài thông qua Internet để truyền và nhận thông tin và cào và thu thập dữ liệu có trong trang web 20 này.

Máy chủ từ xa 50 có thể cung cấp thông tin quy tắc cào dữ liệu RULE thông qua truyền thông với thiết bị đầu cuối khách 10, và thực hiện trực tiếp các lệnh mà cần được thực hiện từ xa thông qua cuộc gọi thủ tục từ xa (RPC).

Thông tin quy tắc có thể xác định các kiểu và thứ tự thực hiện của các tác vụ. Ví dụ, khi có các tác vụ A, B, C và D, thì thông tin quy tắc có thể xác định thứ tự thực hiện của tác vụ tương ứng là A-B-C hoặc A-B-D.

Ngoài ra, các tác vụ trong thông tin quy tắc có thể được tạo cấu hình hoặc được tạo cấu hình lại khi xem xét việc phát hiện lỗi hoặc hiệu quả của việc cào dữ liệu.

Ví dụ, khi các tác vụ được phân chia và được tạo cấu hình trong các đơn vị thông điệp lỗi, nếu thông điệp lỗi nhận được trong quy trình cào dữ liệu, thì có thể dễ dàng xác định nguyên nhân mà tác vụ tạo ra lỗi.

Hơn nữa, nó cũng có thể là lớp và quản lý các tác vụ này.

Khi các tác vụ được phân chia trong các đơn vị thông điệp lỗi như được mô tả ở trên, thì có thể hiệu quả về mặt phát hiện lỗi, nhưng có thể dẫn đến không hiệu quả về mặt truyền thông. Ví dụ, việc truyền dẫn và tiếp nhận dữ liệu đến và từ máy chủ bên ngoài được tạo ra trong quá trình thực hiện tác vụ do các đặc điểm của việc cào dữ liệu, nhưng có thể không hiệu quả nếu nhận sự trả về của mỗi kết quả truyền dẫn và tiếp nhận dưới dạng thông điệp.

Nghĩa là, vì việc phân loại chi tiết các tác vụ có thể làm ảnh hưởng đến hiệu quả của toàn bộ việc cào dữ liệu, nên một số tác vụ có thể được tạo nhóm thành và được quản lý dưới dạng một tác vụ cao hơn. Ví dụ, các tác vụ liên tiếp có thể được tạo nhóm thành số lượng định trước và được quản lý dưới dạng một tác vụ cao hơn, và việc truyền dẫn và tiếp nhận dữ liệu và các kết quả tương ứng có thể được trả về trong các đơn vị tác vụ cao hơn này.

Trong trường hợp này, khi lỗi xảy ra trong khi thực hiện một trong số các tác vụ

thấp hơn thuộc về tác vụ cao hơn, thông điệp lỗi có thể được tạo ra trong các đơn vị tác vụ cao hơn. Theo đó, người quản lý có thể xác định xem có hay không có lỗi cụ thể chỉ tồn tại đối với các tác vụ thấp hơn thuộc về tác vụ cao hơn để thực hiện việc thực hiện và quản lý hiệu quả các tác vụ nhất có thể.

Máy chủ từ xa 50 có thể bao gồm máy chủ quy tắc 52, máy chủ lệnh 54 và máy chủ ngữ cảnh 58.

Máy chủ quy tắc 52 là thiết bị mà xác định các tác vụ để cào dữ liệu và trình tự của các tác vụ dưới dạng thông tin quy tắc trước và lưu trữ và quản lý thông tin quy tắc đã xác định, và có thể được truy cập từ xa. Theo đó, người quản lý có thể truy cập máy chủ quy tắc 52 để sửa đổi trực tiếp thông tin quy tắc đã xác định, và có thể phản ánh trực tiếp tác vụ mà thay đổi theo sự thay đổi của đích cào dữ liệu trong thông tin quy tắc.

Máy chủ lệnh 54 có thể thực hiện lệnh cụ thể từ xa bằng cách vận hành kết hợp với thiết bị đầu cuối khách 10. Máy chủ lệnh 54 này có thể thực hiện việc cào dữ liệu theo thông tin quy tắc được cung cấp bởi máy chủ quy tắc 52, nhưng có thể nhận yêu cầu để thực hiện trực tiếp lệnh mà cần được thực hiện từ xa và trả về kết quả thực hiện cho thiết bị đầu cuối khách 10.

Nghĩa là, các lệnh là các lệnh có trong tác vụ, và có thể được phân chia thành lệnh cục bộ được lưu trữ trong thiết bị đầu cuối khách 10 và lệnh từ xa được lưu trữ trong máy chủ lệnh 54 theo vị trí lưu trữ.

Lệnh cục bộ được lưu trữ trong thiết bị đầu cuối khách 10 được bao gồm và được cài đặt trong ứng dụng 1000, và có thể được tạo cấu hình trong thư viện 100. Theo đó, nói chung, lệnh cục bộ có thể có dạng tĩnh trong đó nó có thể không được sửa đổi trong thiết bị đầu cuối khách 10.

Mặt khác, lệnh từ xa được lưu trữ trong máy chủ lệnh 54 có thể được sửa đổi tự do bởi người quản lý truy cập vào máy chủ lệnh 54, và có thể có dạng động trong đó nó có thể được tùy chỉnh hoặc sửa đổi theo mục đích của việc cào dữ liệu.

Ngoài ra, một số trong số các lệnh có thể được tạo cấu hình dưới dạng các lệnh

thử nghiệm.

Lệnh thử nghiệm có thể được thực hiện định kỳ kể cả khi không có yêu cầu của người dùng, và có thể có các nội dung giống với các nội dung ở thời điểm thực hiện tác vụ cào dữ liệu theo yêu cầu thực tế.

Theo đó, lệnh thử nghiệm được thực hiện định kỳ để phát hiện định kỳ sự thay đổi trong đích cào dữ liệu, và có thể xác định sự xuất hiện của lỗi theo đó.

Máy chủ quy tắc 52 và máy chủ lệnh 54 có thể được quản lý từ xa và được tạo cấu hình một cách độc lập với nhau. Quy trình được thực hiện để cào dữ liệu thực trong máy chủ lệnh 54 có thể không truy cập được từ bên ngoài, và có thể chỉ được điều khiển bởi tác vụ được xác định trong máy chủ quy tắc 52.

Theo đó, sự rò rỉ thông tin được sử dụng đối với lệnh được thực hiện từ xa có thể được ngăn ngừa, và ví dụ, sự rò rỉ thông tin xác thực công khai của người dùng hoặc mật khẩu được yêu cầu để đăng nhập có thể được ngăn ngừa.

Tuy nhiên, theo một phương án, máy chủ quy tắc 52 và máy chủ lệnh 54 có thể được tạo cấu hình và được quản lý cùng nhau trong máy chủ từ xa 50 ngoại trừ việc chúng được tách biệt về mặt vật lý và được vận hành dưới dạng các thực thể riêng biệt.

Một ví dụ là, máy chủ quy tắc 52 và máy chủ lệnh 54 có thể được tạo cấu hình dưới dạng các mô đun bên trong thực hiện các tác vụ khác nhau dưới dạng các thiết bị trong máy chủ từ xa 50. Máy chủ quy tắc 52 và máy chủ lệnh 54 có thể được tạo cấu hình và được vận hành lần lượt dưới dạng mô đun quy tắc và mô đun lệnh, trong máy chủ từ xa 50, nhưng mô đun lệnh có thể được thực hiện trong máy chủ từ xa sao cho việc điều khiển của hoặc truy cập vào mô đun lệnh từ bên ngoài là không thể.

Máy chủ ngữ cảnh 58 có thể lưu trữ giá trị cần được tham chiếu ở thời điểm thực hiện tác vụ trong máy chủ lệnh 54.

Ví dụ, máy chủ ngữ cảnh 58 tạo cấu hình dữ liệu ngữ cảnh A.Block, B.Title và C.Desc bao gồm các tên tác vụ A, B và C, và trong trường hợp này, dữ liệu ngữ cảnh có thể bao gồm các nội dung chi tiết như giá trị dữ liệu cần được tham chiếu ở thời điểm

thực hiện tác vụ thực.

Máy chủ ngữ cảnh 58 lưu trữ lại kết quả thực hiện lệnh dưới dạng dữ liệu ngữ cảnh, và trong trường hợp này, máy chủ liên hệ 58 có thể lưu trữ kết quả thực hiện dưới dạng dữ liệu ngữ cảnh ở trạng thái trong đó tên tác vụ bao gồm lệnh tương ứng được đưa ra cho dữ liệu ngữ cảnh. Theo đó, khi kết quả thực hiện thực sự được lưu trữ, kết quả thực hiện có thể được lưu trữ dưới dạng dữ liệu ngữ cảnh được phân chia cho mỗi phiên cào dữ liệu.

Dữ liệu ngữ cảnh có thể được lưu trữ dưới dạng “task name.variable name A.Block” để giảm khả năng xung đột giữa các biến được xác định khác nhau và đồng thời, cho phép mối quan hệ phụ thuộc giữa các tác vụ cần được nắm bắt.

Ví dụ, người quản lý có thể dễ dàng nắm bắt từ dữ liệu ngữ cảnh được tham chiếu A.Block mà dữ liệu ngữ cảnh là kết quả của tác vụ A, và có thể nắm bắt rằng một số lệnh của tác vụ B tham chiếu đến A.Block dưới dạng mối quan hệ phụ thuộc của tác vụ B trên tác vụ A.

Ngoài ra, cũng có thể dự đoán xem có hay không có vòng lặp vô hạn xảy ra do sự tham chiếu chéo của các lệnh trong tác vụ thông qua mối quan hệ phụ thuộc và khả năng xảy ra lỗi của tác vụ do vòng lặp vô hạn này.

Ngoài ra, máy chủ ngữ cảnh 58 có thể thực hiện việc đồng bộ hóa với thiết bị đầu cuối khách 10 trong khi phiên để cào dữ liệu cụ thể được duy trì, và nắm bắt một loạt hoạt động.

Trong khi phiên được duy trì, máy chủ ngữ cảnh 58 có thể nhận các giá trị kết quả thực hiện tác vụ được kết xuất từ thiết bị đầu cuối khách 10 và ghi lại các giá trị kết quả thực hiện tác vụ này dưới dạng nhật ký. Máy chủ ngữ cảnh 58 lưu trữ các kết quả thực hiện dưới dạng nhật ký để cho phép dễ dàng nắm bắt được nguyên nhân của lỗi khi lỗi xảy ra trong quy trình cào dữ liệu.

Trong quy trình cào dữ liệu thông thường, khi lỗi xảy ra trong thiết bị đầu cuối khách 10, chỉ giá trị kết quả của lỗi là nhận được, và do đó khó, rất khó để xác nhận lỗi

này một cách chi tiết. Tuy nhiên, máy chủ ngữ cảnh 50 lưu trữ các kết quả thực hiện trong một phiên dưới dạng nhật ký để cho phép xác nhận một cách dễ dàng lỗi xảy ra ở bước nào của các quy trình liên tiếp.

Ngoài ra, máy chủ ngữ cảnh 58 có thể ngăn ngừa sự rò rỉ thông tin cá nhân thu được bằng cách cào dữ liệu bằng cách quy định rằng thông tin cá nhân được loại bỏ sau khoảng thời gian lưu trữ định trước khi thông tin cá nhân có trong thông tin được lưu trữ dưới dạng nhật ký.

Bộ thực hiện tác vụ 110 có thể nhận thông tin quy tắc từ máy chủ quy tắc 52 và thực hiện tác vụ theo thông tin quy tắc này.

Bộ thực hiện tác vụ 110 có thể thực hiện tác vụ theo thông tin quy tắc nhận được, nhưng có thể yêu cầu máy chủ lệnh 54 thực hiện một số lệnh được xác định trong tác vụ này.

Bộ thực hiện tác vụ 110 có thể truyền cuộc gọi thủ tục từ xa đến máy chủ lệnh 54 trong trường hợp lệnh từ xa, và bộ thực hiện lệnh từ xa 56 của máy chủ lệnh 54 mà đã nhận được cuộc gọi thủ tục từ xa có thể thực hiện lệnh từ xa và trả về kết quả thực hiện của lệnh từ xa đến bộ thực hiện tác vụ 110.

Bộ thực hiện tác vụ 110 có thể thực hiện quy trình cào dữ liệu theo thông tin quy tắc đã xác định, và có thể cung cấp dữ liệu được thu thập thông qua thiết bị đầu cuối khách 10.

Một loạt dữ liệu được trao đổi trong quá trình truyền thông giữa thiết bị đầu cuối khách 10 và máy chủ từ xa 50 được mô tả ở trên có thể được nén, được mã hóa và sau đó được truyền/nhận giữa thiết bị đầu cuối khách 10 và máy chủ từ xa 50.

Dữ liệu thu được trong quy trình cào dữ liệu có thể bao gồm thông tin cá nhân, chủ yếu sử dụng thông tin của trang web 20, và do đó có thể được tạo cấu hình dưới dạng tài liệu ngôn ngữ đánh dấu siêu văn bản (hypertext markup language, HTML).

Theo đó, bằng cách nén và mã hóa chuỗi ký tự lặp trong tài liệu HTML trước tiên, hiệu quả nén có thể được tăng lên và đồng thời, kích thước của dữ liệu được mã hóa có

thể được giảm xuống để giảm bớt thời gian mã hóa.

Hơn nữa, bằng cách mã hóa thông tin cá nhân trong đó, thiệt hại do sự rò rỉ thông tin cá nhân có thể được ngăn ngừa.

Sau đây, các hoạt động giữa thiết bị đầu cuối khách 10 và máy chủ quy tắc 52 sẽ được mô tả một cách chi tiết hơn dựa vào Fig.3.

Fig.3 là sơ đồ trình tự minh họa quy trình yêu cầu thông tin quy tắc giữa thiết bị đầu cuối khách 10 và máy chủ quy tắc 52.

Trước tiên, ứng dụng 1000 được cài đặt trong thiết bị đầu cuối khách 10 có thể yêu cầu cào dữ liệu khi cần thiết để thu dữ liệu từ trang web bên ngoài trong quá trình hoạt động.

Trong trường hợp này, thiết bị đầu cuối khách 10 cung cấp giá trị ban đầu để cào dữ liệu cho bộ thực hiện tác vụ 110 (32).

Giá trị ban đầu để cào dữ liệu là thông tin nhận dạng để cào dữ liệu, và có thể bao gồm thông tin khác nhau được yêu cầu để cào dữ liệu, như thông tin về bộ định vị tài nguyên thống nhất (uniform resource locator, URL) của trang web đích cào dữ liệu và thông tin xác thực được yêu cầu để thu dữ liệu.

Ngoài ra, giá trị ban đầu để cào dữ liệu có thể còn bao gồm thông tin xác định định dạng hoặc kiểu dữ liệu cần thu được.

Bộ thực hiện tác vụ 110 yêu cầu máy chủ quy tắc truyền quy tắc cào dữ liệu để cào dữ liệu, sử dụng giá trị ban đầu để cào dữ liệu (34).

Máy chủ quy tắc 52 truyền quy tắc cào dữ liệu phù hợp của các quy tắc được lưu trữ theo giá trị ban đầu nhận được làm thông tin quy tắc (36).

Bộ thực hiện tác vụ 110 mà đã nhận được thông tin quy tắc thực hiện một loạt tác vụ được xác định trong thông tin quy tắc. Bộ thực hiện tác vụ 110 có thể chọn tác vụ ưu tiên nhất trong thông tin quy tắc và thực hiện tác vụ ưu tiên nhất (38).

Trong trường hợp này, mỗi tác vụ có thể được thực hiện dưới dạng khối bao gồm

các lệnh theo mục đích chi tiết để cào dữ liệu, và bộ thực hiện tác vụ 110 có thể đánh giá kết quả thực hiện mỗi khi việc thực hiện mỗi tác vụ được hoàn thành (40).

Bộ thực hiện tác vụ 110 thực hiện tác vụ tiếp theo trong thông tin quy tắc khi không có sự bất thường trong kết quả thực hiện.

Các quy trình được mô tả ở trên có thể được thực hiện lặp lại, tất cả các tác vụ được xác định trong thông tin quy tắc có thể được thực hiện, và dữ liệu thu được có thể được cung cấp cho thiết bị đầu cuối khách dưới dạng kết quả cào dữ liệu (44).

Sau đây, quy trình thực hiện tác vụ chi tiết của bộ thực hiện tác vụ 110 sẽ được mô tả dựa vào Fig.4. Theo một phương án, tác vụ có thể bao gồm các lệnh.

Trong trường hợp này, các lệnh có thể được phân chia thành lệnh cục bộ mà có thể được thực hiện trực tiếp bởi thiết bị đầu cuối khách và lệnh từ xa mà cần được thực hiện thông qua máy chủ từ xa bên ngoài.

Theo đó, bộ thực hiện tác vụ 110 có thể nhận yêu cầu thực hiện tác vụ (62), và có thể chọn và thực hiện lệnh mà cần được thực hiện ưu tiên.

Khi lệnh mà cần thực hiện là lệnh cục bộ, thì lệnh cục bộ có thể được thực hiện trên thiết bị đầu cuối khách (64).

Khi việc thực hiện lệnh cục bộ được hoàn thành, thì có thể xác định được xem lệnh mà cần thực hiện tiếp theo là lệnh cục bộ hay lệnh từ xa.

Khi lệnh tiếp theo là lệnh từ xa (66), thì bộ thực hiện tác vụ 110 có thể truyền cuộc gọi thủ tục từ xa đến máy chủ lệnh (68).

Cuộc gọi thủ tục từ xa có thể bao gồm thông tin xác định các nội dung của lệnh được yêu cầu để thực hiện từ xa, và máy chủ lệnh 54 có thể thực hiện lệnh từ xa thông qua cuộc gọi thủ tục từ xa.

Máy chủ lệnh 54 trả về kết quả thực hiện cho thiết bị đầu cuối khách sau khi thực hiện lệnh từ xa (72).

Các quy trình được mô tả ở trên có thể được lặp lại cho đến khi tất cả các lệnh của

tác vụ được thực hiện, và khi việc thực hiện lệnh được hoàn thành, thì kết quả thực hiện tác vụ có thể được cung cấp cho thiết bị đầu cuối khách (74).

Theo một phương án, tác vụ có thể bao gồm tổ hợp của lệnh cục bộ và lệnh từ xa, và thiết bị đầu cuối khách 10 có thể thực hiện lặp lại quy trình truyền cuộc gọi thủ tục từ xa đến máy chủ lệnh 54 và nhận việc trả về kết quả theo các đặc điểm của lệnh.

Sau đây, các quy trình của phương pháp cào dữ liệu theo một phương án sẽ được mô tả một cách chi tiết hơn dựa vào Fig.5.

Fig.5 là lưu đồ minh họa phương pháp trong đó ứng dụng thực hiện dịch vụ cào dữ liệu trên thiết bị đầu cuối khách theo một phương án của sáng chế.

Trước tiên, khi việc cào dữ liệu đối với trang web bên ngoài được yêu cầu, thì ứng dụng có thể yêu cầu thông tin quy tắc cào dữ liệu (S10).

Ứng dụng 1000 có thể yêu cầu thông tin quy tắc bằng cách truyền giá trị ban đầu để cào dữ liệu đến máy chủ quy tắc 52 thông qua môđun truyền thông của thiết bị đầu cuối khách 10.

Khi thông tin quy tắc được yêu cầu, thì máy chủ quy tắc 52 có thể tìm kiếm thông tin quy tắc xác định quy trình cào dữ liệu phù hợp và truyền thông tin quy tắc đến thiết bị đầu cuối khách 10.

Trong trường hợp này, thông tin quy tắc có thể được chọn từ danh sách định trước, hoặc khi thông tin quy tắc phù hợp không tồn tại, thì máy chủ quy tắc 52 có thể còn tạo cấu hình tác vụ thông qua tổ hợp của các lệnh cơ bản và tạo ra tác vụ dưới dạng thông tin quy tắc.

Ngoài ra, thông tin quy tắc cũng có thể được tạo ra và được cung cấp bằng cách phân chia lệnh cục bộ thực hiện được bởi thiết bị đầu cuối khách 10 và lệnh từ xa thực hiện được bởi máy chủ lệnh 54 theo thực thể thực hiện.

Hơn nữa, thông tin quy tắc có thể được sửa đổi hoặc được tạo mới thông qua sự truy cập bên ngoài của người quản lý được ủy quyền, ngoài việc được tạo cấu hình trước bởi tổ hợp của các lệnh cơ bản.

Nghĩa là, khi có sự thay đổi trong trang web cụ thể, là đích cào dữ liệu, thông tin quy tắc có thể được sửa đổi tương ứng với sự thay đổi này.

Theo đó, so với phương pháp mã hóa cứng trong đó quy tắc cào dữ liệu được xác định dưới dạng thư viện trong ứng dụng, người quản lý có thể đối phó tích cực với đích cào dữ liệu, và người dùng có thể sử dụng các dịch vụ hiện có mà không cập nhật ứng dụng này.

Thiết bị đầu cuối khách 10 nhận thông tin quy tắc cào dữ liệu từ máy chủ quy tắc 52 (S20).

Bộ thực hiện tác vụ 110 trong thiết bị đầu cuối khách 10 thực hiện các tác vụ được xác định trong thông tin quy tắc nhận được như được mô tả ở trên (S30).

Các tác vụ có thể được thực hiện tuần tự, và có thể được thực hiện dựa vào các kết quả của các tác vụ được thực hiện trước đó với sự tương quan giữa chúng.

Tác vụ được thực hiện hiện tại có thể tạo ra kết quả dựa vào các kết quả thực hiện của các tác vụ trước đó bằng cách sử dụng tên tác vụ và các biến liên quan được xác định trong thông tin quy tắc.

Tác vụ bao gồm các lệnh, và kết quả tác vụ được tạo ra thông qua quy trình thực hiện các lệnh có thể được lưu trữ cho mỗi tác vụ.

Ngoài ra, các lệnh có thể được phân chia thành lệnh cục bộ và lệnh từ xa, và có thể được xử lý theo cách phân phối lần lượt trong thiết bị đầu cuối khách và máy chủ lệnh.

Các lệnh tương ứng có thể được xử lý theo cách thăm dò không đồng bộ.

Theo đó, thiết bị đầu cuối khách 10 có thể thực hiện lệnh trong giao thức định trước, nhưng có thể yêu cầu máy chủ lệnh thực hiện lệnh từ xa bởi cuộc gọi thủ tục từ xa khi lệnh là lệnh từ xa. Máy chủ lệnh 54 có thể thực hiện lệnh từ xa khi nó nhận cuộc gọi thủ tục từ xa.

Ngoài ra, phương pháp cào dữ liệu theo một phương án là phương pháp không đồng bộ, và độ trễ có thể xảy ra khi thực hiện tác vụ theo một số lệnh từ xa được tạo cầu

hình trong tác vụ này.

Vì lệnh từ xa làm tăng số lượng cuộc gọi thủ tục từ xa và số lần tiếp nhận các kết quả được trả về, nên việc thực hiện tác vụ có thể bị trễ theo trạng thái truyền thông, và có thể cần chờ để thực hiện các lệnh cục bộ khác cho đến khi nhận được kết quả của lệnh từ xa.

Theo đó, trong phương pháp cào dữ liệu theo một phương án, cách truyền tạm thời quyền điều khiển của bộ thực hiện tác vụ 110 đến máy chủ lệnh 54 khi xem xét cấu hình lệnh trong tác vụ này có thể được sử dụng.

Về việc này, dựa vào Fig.6, bộ thực hiện tác vụ 110 có thể xác định xem lệnh mà cần được thực hiện được thực hiện cục bộ hay từ xa, và đồng thời, có thể xác định thêm xem có hay không có các lệnh từ xa là liên tục.

Bộ thực hiện tác vụ 110 có thể thực hiện lệnh cục bộ (82), nhưng có thể chuyển quyền điều khiển để thực hiện lệnh cho máy chủ lệnh 54 khi các lệnh từ xa là liên tục theo lịch trình thực hiện của các tác vụ.

Theo đó, máy chủ lệnh 54 đã có quyền điều khiển có thể bỏ qua quy trình lặp lại để trả về kết quả cho thiết bị đầu cuối khách 10 mỗi khi lệnh từ xa được thực hiện và nhận cuộc gọi thủ tục từ xa mới.

Nghĩa là, các lệnh từ xa liên tục có thể được thực hiện mà không nhận các cuộc gọi thủ tục từ xa riêng biệt (88 và 90). Theo đó, thời gian chờ từ lúc truyền cuộc gọi thủ tục từ xa đến khi trả về kết quả có thể được giảm xuống.

Thông qua các quy trình được mô tả ở trên, khi các lệnh từ xa liên tục được thực hiện, thì máy chủ lệnh có thể chuyển quyền điều khiển trở lại cho bộ thực hiện tác vụ cùng với việc trả về kết quả thực hiện (92).

Ngoài ra, trong phương pháp cào dữ liệu theo một phương án, việc cào dữ liệu được thực hiện theo các đơn vị tác vụ, và hiệu quả cào dữ liệu có thể được tăng lên bằng cách thực hiện các tác vụ song song khi xem xét mối quan hệ tham chiếu có thể thay đổi của các tác vụ.

Như được mô tả ở trên, có thể nắm bắt mối quan hệ phụ thuộc giữa các tác vụ theo việc xem có hay không có các lệnh trong các tác vụ được tham chiếu, và cũng có thể cho phép cải dữ liệu nhanh hơn và hiệu quả hơn bằng cách sử dụng mối quan hệ phụ thuộc trong trường hợp này.

Bộ thực hiện tác vụ 110 có thể cho phép các tác vụ độc lập trong đó các biến đầu vào và các biến đầu ra giữa các tác vụ không liên quan với nhau cần được thực hiện đồng thời song song bất kể thứ tự theo thông tin quy tắc khi xem xét quy trình thực hiện của tất cả các tác vụ.

Ví dụ, mối quan hệ phụ thuộc giữa các tác vụ được xác định trong thông tin quy tắc có thể được biểu diễn dưới dạng đồ thị phụ thuộc như được minh họa trên Fig.7A. Trong trường hợp này, các nút tương ứng của đồ thị này đề cập đến các tác vụ, và các liên kết nối các nút với nhau và các mũi tên của các liên kết có thể xác định các hướng tham chiếu của các tác vụ.

Trong trường hợp trên Fig.7A, các tác vụ B, C và D có thể tham chiếu kết quả thực hiện của tác vụ A, và tác vụ E có thể tham chiếu các kết quả thực hiện của các tác vụ B và C. Tác vụ F tham chiếu các kết quả thực hiện của tác vụ E và tác vụ D ở cùng thời điểm.

Theo đó, bộ thực hiện tác vụ 110 có thể xác định thứ tự thực hiện của các tác vụ khi xem xét mối quan hệ phụ thuộc đối với hoạt động được tối ưu hóa ở thời điểm thực hiện thực sự các tác vụ kể cả khi người quản lý tạo ra các mã để các tác vụ A, B, C, D, E và F được thực hiện tuần tự ở thời điểm tạo ra thông tin quy tắc thực.

Cụ thể, theo mối quan hệ phụ thuộc, các tác vụ B, C và D tham chiếu độc lập kết quả của tác vụ A và không tham chiếu với nhau, và do đó, bộ thực hiện tác vụ 110 có thể thực hiện đồng thời các tác vụ B, C và D để cho phép công việc cải dữ liệu nhanh hơn và hiệu quả hơn. Nghĩa là, bộ thực hiện tác vụ 110 có thể thực hiện đồng thời các tác vụ B, C và D song song ở thời điểm thực hiện thực sự các tác vụ B, C và D kể cả khi các tác vụ B, C và D đã được xác định tuần tự trong thông tin quy tắc.

Ngoài ra, máy chủ quy tắc 52 cũng có thể nắm bắt mối quan hệ phụ thuộc giữa

các tác vụ trước, xác định thứ tự thực hiện thực tế của các tác vụ, và cung cấp thứ tự thực hiện thực tế này cho thiết bị đầu cuối khách 10.

Theo đó, máy chủ quy tắc 52 có thể trích xuất đồ thị trình tự để xác định lại thứ tự thực hiện của các tác vụ thông tin quy tắc thông qua mối quan hệ phụ thuộc giữa các tác vụ được xác định như được minh hoạ trên Fig.7A dưới dạng được minh hoạ trên Fig.7B.

Dựa vào Fig.7B, đồ thị trình tự có thể xác định thứ tự thực hiện của các tác vụ mà cần được thực hiện tuần tự thông qua các liên kết theo hướng ngang, và có thể xác định mối quan hệ giữa các tác vụ mà có thể được thực hiện đồng thời song song thông qua các liên kết theo hướng dọc.

Theo đó, máy chủ quy tắc 52 có thể xác định lại thứ tự thực hiện của các tác vụ dưới dạng đồ thị trình tự và cung cấp thứ tự thực hiện đã xác định lại này cho thiết bị đầu cuối khách 10 để cho phép các tác vụ B, C và D được thực hiện đồng thời sau khi tác vụ này được thực hiện, nhờ đó tăng hiệu quả của toàn bộ việc cào dữ liệu.

Khi việc thực hiện các tác vụ cào dữ liệu được mô tả ở trên được hoàn thành, thì dữ liệu được thu thập từ trang web bên ngoài của ứng dụng dưới dạng kết quả thực hiện (S40).

Dữ liệu đã thu thập có thể được cung cấp theo dịch vụ được yêu cầu bởi người dùng thông qua ứng dụng 1000 (S50).

Sau đây, cấu hình của thư viện trong ứng dụng theo một phương án của sáng chế sẽ được mô tả dựa vào Fig.8.

Theo một phương án, ứng dụng là chương trình ứng dụng (hoặc phần mềm ứng dụng) được cài đặt và điều khiển trên thiết bị đầu cuối khách, và được thực hiện để vận hành kết hợp với máy chủ quy tắc và máy chủ lệnh.

Thư viện 100 trong ứng dụng có thể bao gồm bộ yêu cầu quy tắc cào dữ liệu 120, bộ thực hiện tác vụ 110 và nhà cung cấp dịch vụ 130.

Bộ yêu cầu quy tắc cào dữ liệu 120 có thể yêu cầu máy chủ quy tắc 52 truyền thông tin quy tắc để cào dữ liệu.

Nghĩa là, theo một phương án, ứng dụng được thực hiện theo cách lưu trữ bên ngoài và quản lý quy trình đề cào dữ liệu trong máy chủ quy tắc 52 riêng biệt thay vì cách (cách mã hóa cứng) lưu trữ bên trong quy trình đề cào dữ liệu trước đối với mỗi trang web đích.

Theo đó, có thể đối phó tích cực với các thay đổi trong đích cào dữ liệu và đối phó với các thay đổi trong đích cào dữ liệu bằng cách chỉ thay đổi thông tin quy tắc trong máy chủ quy tắc mà không cần phân phối lại ứng dụng riêng biệt.

Bộ thực hiện tác vụ 110 có thể thực hiện tác vụ bằng cách sử dụng thông tin quy tắc nhận được từ bộ yêu cầu quy tắc cào dữ liệu 120.

Cụ thể, tác vụ có thể bao gồm tổ hợp của các lệnh, và có thể bao gồm các lệnh cục bộ và lệnh từ xa.

Lệnh cục bộ là lệnh được thực hiện trực tiếp bởi thiết bị đầu cuối khách, ví dụ, lệnh được tạo ra theo cách cơ bản, và có thể được tạo ra trước trong máy chủ quy tắc.

Ví dụ, lệnh để yêu cầu yêu cầu/phản hồi giao thức truyền siêu văn bản (hypertext transfer protocol, HTTP) và đầu vào người dùng và thực hiện yêu cầu/phản hồi đăng nhập HTTP hoặc dạng tương tự có thể được tạo ra dưới dạng lệnh cục bộ.

Lệnh từ xa là lệnh được thực hiện thông qua máy chủ lệnh, và có thể được tùy chỉnh theo dữ liệu bên ngoài được yêu cầu.

Ví dụ, thông qua lệnh từ xa được tùy chỉnh, máy chủ lệnh có thể tạo ra khóa băm để đáp lại chứng nhận công khai và thực hiện tác vụ mà yêu cầu việc xử lý bổ sung như yêu cầu trang thông tin dựa vào trang tùy chỉnh.

Khi việc cập nhật trang trong trang web đích được tạo ra, thì có thể được tạo cấu hình dưới dạng trang tùy chỉnh và lệnh tương ứng có thể được thực hiện từ xa.

Bộ thực hiện tác vụ 110 có thể còn bao gồm bộ phát cuộc gọi thủ tục từ xa (RPC) 112 yêu cầu cuộc gọi thủ tục từ xa để thực hiện lệnh từ xa và bộ thu kết quả lệnh 114 nhận kết quả thực hiện của lệnh từ xa.

Nhà cung cấp dịch vụ 130 xử lý dữ liệu được thu thập bởi bộ thực hiện tác vụ 110

được mô tả ở trên và cung cấp dữ liệu đã xử lý cho người dùng.

Như được mô tả ở trên, theo sáng chế, bằng cách cung cấp dịch vụ cào dữ liệu thông qua thiết bị đầu cuối khách và máy chủ từ xa theo cách phân phối, thì có thể làm tăng hiệu quả của tác vụ cào dữ liệu và đối phó tích cực với các thay đổi trong đích cào dữ liệu.

Ngoài ra, có thể tăng mức độ thuận tiện cho người dùng khi sử dụng bằng cách giảm thiểu sự cập nhật hoặc phân phối lại của ứng dụng đáp lại sự thay đổi của đích cào dữ liệu.

Hơn nữa, có thể tăng độ bảo mật bằng cách vận hành máy chủ từ xa thực hiện việc cào dữ liệu một cách độc lập từ thiết bị đầu cuối khách và ngăn ngừa sự truy cập từ bên ngoài vào và sự rò rỉ thông tin được sử dụng để thực hiện việc cào dữ liệu.

Ở trên, các phương án khác nhau được mô tả trong bản mô tả này có thể được thực hiện trong vật ghi đọc được bằng máy tính lâu dài bằng cách sử dụng, ví dụ, phần mềm, phần cứng hoặc tổ hợp của các phần này.

Theo một phương án thực hiện về phần cứng, các phương án được mô tả trong bản mô tả này có thể được thực hiện bằng cách sử dụng ít nhất một trong số các mạch tích hợp cho ứng dụng cụ thể (application specific integrated circuit, ASIC), bộ xử lý tín hiệu số (digital signal processor, DSP), thiết bị xử lý tín hiệu số (digital signal processing device, DSPD), thiết bị logic lập trình được (programmable logic device, PLD), mảng cổng lập trình được dạng trường (field programmable gate array, FPGA), bộ xử lý, bộ điều khiển, bộ vi điều khiển, bộ vi xử lý và đơn vị điện tử để thực hiện các chức năng khác. Trong một số trường hợp, các phương án được mô tả trong bản mô tả này có thể được thực hiện dưới dạng chính môđun điều khiển.

Theo một phương án thực hiện về phần mềm, các phương án như các thủ tục và chức năng được mô tả trong bản mô tả này có thể được thực hiện dưới dạng các môđun phần mềm riêng biệt. Mỗi trong số các môđun phần mềm có thể thực hiện một hoặc nhiều chức năng và hoạt động được mô tả trong bản mô tả này. Mã phần mềm có thể được thực hiện dưới dạng ứng dụng phần mềm được ghi theo ngôn ngữ lập trình thích

hợp. Mã phần mềm này có thể được lưu trữ trong môđun bộ nhớ và được thực hiện bằng môđun điều khiển.

Mục đích kỹ thuật của sáng chế đã được mô tả chỉ bằng cách lấy ví dụ ở trên, và sáng chế có thể được cải biến, biến đổi và thay thế theo cách khác nhau bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật mà sáng chế thuộc về không lệch khỏi các dấu hiệu cơ bản của sáng chế.

Theo đó, các phương án được bộc lộ theo sáng chế và các hình vẽ kèm theo được đề xuất để mô tả mục đích kỹ thuật của sáng chế thay vì giới hạn mục đích kỹ thuật của sáng chế, và phạm vi của sáng chế không bị giới hạn bởi các phương án này và các hình vẽ kèm theo. Phạm vi của sáng chế cần được hiểu bởi các điểm yêu cầu bảo hộ sau, và cần được hiểu rằng tất cả các mục đích tương đương với các điểm yêu cầu bảo hộ sau đều nằm trong phạm vi của sáng chế.

YÊU CẦU BẢO HỘ

1. Hệ thống cung cấp dịch vụ dựa trên cào dữ liệu bao gồm:

thiết bị đầu cuối khách bao gồm bộ thực hiện tác vụ được tạo cấu hình để yêu cầu và nhận thông tin quy tắc để cào dữ liệu và thu dữ liệu bên ngoài bằng cách thực hiện các tác vụ trong thông tin quy tắc; và

máy chủ từ xa bao gồm máy chủ quy tắc được tạo cấu hình để quản lý thông tin quy tắc và truyền thông tin quy tắc này theo yêu cầu từ thiết bị đầu cuối khách, máy chủ lệnh được tạo cấu hình để thực hiện lệnh theo cuộc gọi thủ tục từ xa nhận được từ thiết bị đầu cuối khách và trả về kết quả thực hiện của lệnh này, và máy chủ ngữ cảnh được tạo cấu hình để ghi các giá trị kết quả tác vụ theo thông tin quy tắc dưới dạng thông tin nhật ký,

trong đó thông tin quy tắc bao gồm thông tin xác định kiểu và thứ tự thực hiện của các tác vụ,

các tác vụ được tạo cấu hình trong thông tin quy tắc theo các đơn vị thông điệp lỗi để phát hiện lỗi, và một số tác vụ độc lập được thực hiện song song theo mỗi quan hệ phụ thuộc định trước dựa vào các giá trị dữ liệu ngữ cảnh bao gồm các tên tác vụ và các tên biến ở thời điểm thực hiện các tác vụ này,

lệnh là lệnh có trong tác vụ, và bao gồm lệnh cục bộ thực hiện được bởi thiết bị đầu cuối khách, lệnh từ xa được thực hiện thông qua máy chủ từ xa, và lệnh thử nghiệm thực hiện việc xác minh định kỳ, và

lệnh từ xa có dạng động trong đó lệnh này sửa đổi được thông qua máy chủ lệnh.

2. Hệ thống cung cấp dịch vụ dựa trên cào dữ liệu theo điểm 1, trong đó bộ thực hiện tác vụ được tạo cấu hình để thu thập dữ liệu bên ngoài bằng cách thực hiện các lệnh cấu thành tác vụ được xác định trong thông tin quy tắc trên thiết bị đầu cuối khách hoặc máy chủ lệnh.

3. Hệ thống cung cấp dịch vụ dựa trên cào dữ liệu theo điểm 1, trong đó dữ liệu bên ngoài đã thu được nén và sau đó được mã hóa trong tài liệu ngôn ngữ đánh dấu siêu văn

bản (hypertext markup language, HTML).

4. Hệ thống cung cấp dịch vụ dựa trên cào dữ liệu theo điểm 1, trong đó bộ thực hiện tác vụ bao gồm:

bộ phát cuộc gọi thủ tục từ xa (remote procedure call, RPC) được tạo cấu hình để truyền cuộc gọi thủ tục từ xa đến máy chủ từ xa để thực hiện lệnh từ xa; và

bộ thu kết quả lệnh được tạo cấu hình để nhận việc trả về kết quả thực hiện của tác vụ được thực hiện theo cuộc gọi thủ tục từ xa.

5. Hệ thống cung cấp dịch vụ dựa trên cào dữ liệu theo điểm 1, trong đó bộ thực hiện tác vụ được tạo cấu hình để:

xác minh kết quả thực hiện của tác vụ được thực hiện để thực hiện mỗi tác vụ này, và

yêu cầu sửa đổi thông tin quy tắc hoặc cuộc gọi thủ tục từ xa đối với lệnh khi có sự bất thường trong khi thực hiện tác vụ.

6. Hệ thống cung cấp dịch vụ dựa trên cào dữ liệu theo điểm 1, trong đó các tác vụ được thực hiện theo các đơn vị khối, sao cho các kết quả thực hiện của các lệnh trong các tác vụ được lưu trữ trong các khối, và

các kết quả thực hiện của các tác vụ được tham chiếu bằng cách sử dụng các tên tác vụ và các biến liên quan ở thời điểm tham chiếu các tác vụ.

7. Hệ thống cung cấp dịch vụ dựa trên cào dữ liệu theo điểm 1, trong đó lệnh cục bộ và lệnh từ xa được thực hiện theo cách được phân phối thông qua thiết bị đầu cuối khách hoặc máy chủ từ xa, và

tác vụ được thăm dò theo cách không đồng bộ.

8. Hệ thống cung cấp dịch vụ dựa trên cào dữ liệu theo điểm 1, trong đó bộ thực hiện tác vụ được tạo cấu hình để truyền giá trị ban đầu để cào dữ liệu cho máy chủ quy tắc, và

máy chủ quy tắc được tạo cấu hình để truyền quy tắc cào dữ liệu phù hợp của các quy tắc được lưu trữ theo giá trị ban đầu nhận được dưới dạng thông tin quy tắc đến thiết

bị đầu cuối khách.

9. Máy chủ từ xa để cung cấp dịch vụ dựa trên cào dữ liệu bao gồm:

máy chủ quy tắc được tạo cấu hình để quản lý thông tin quy tắc để cào dữ liệu và truyền thông tin quy tắc theo yêu cầu từ thiết bị đầu cuối khách;

máy chủ lệnh được tạo cấu hình để thực hiện lệnh theo cuộc gọi thủ tục từ xa nhận được từ thiết bị đầu cuối khách và trả về kết quả thực hiện của lệnh; và

máy chủ ngữ cảnh được tạo cấu hình để ghi các giá trị kết quả tác vụ theo thông tin quy tắc dưới dạng thông tin nhật ký,

trong đó thông tin quy tắc bao gồm thông tin xác định kiểu và thứ tự thực hiện của các tác vụ,

các tác vụ được tạo cấu hình trong thông tin quy tắc theo các đơn vị thông điệp lỗi để phát hiện lỗi, và một số tác vụ độc lập được thực hiện song song theo mỗi quan hệ phụ thuộc định trước dựa vào các giá trị dữ liệu ngữ cảnh bao gồm các tên tác vụ và các tên biến ở thời điểm thực hiện các tác vụ này,

lệnh là lệnh có trong tác vụ, và bao gồm lệnh cục bộ thực hiện được bởi thiết bị đầu cuối khách, lệnh từ xa được thực hiện thông qua máy chủ từ xa, và lệnh thử nghiệm thực hiện việc xác minh định kỳ, và

lệnh từ xa có dạng động trong đó lệnh này sửa đổi được thông qua máy chủ lệnh.

10. Máy chủ từ xa để cung cấp dịch vụ dựa trên cào dữ liệu theo điểm 9, trong đó máy chủ quy tắc được tạo cấu hình để nhận giá trị ban đầu để cào dữ liệu từ thiết bị đầu cuối khách và truyền quy tắc cào dữ liệu phù hợp của các quy tắc được lưu trữ theo giá trị ban đầu nhận được dưới dạng thông tin quy tắc đến thiết bị đầu cuối khách.

FIG. 1

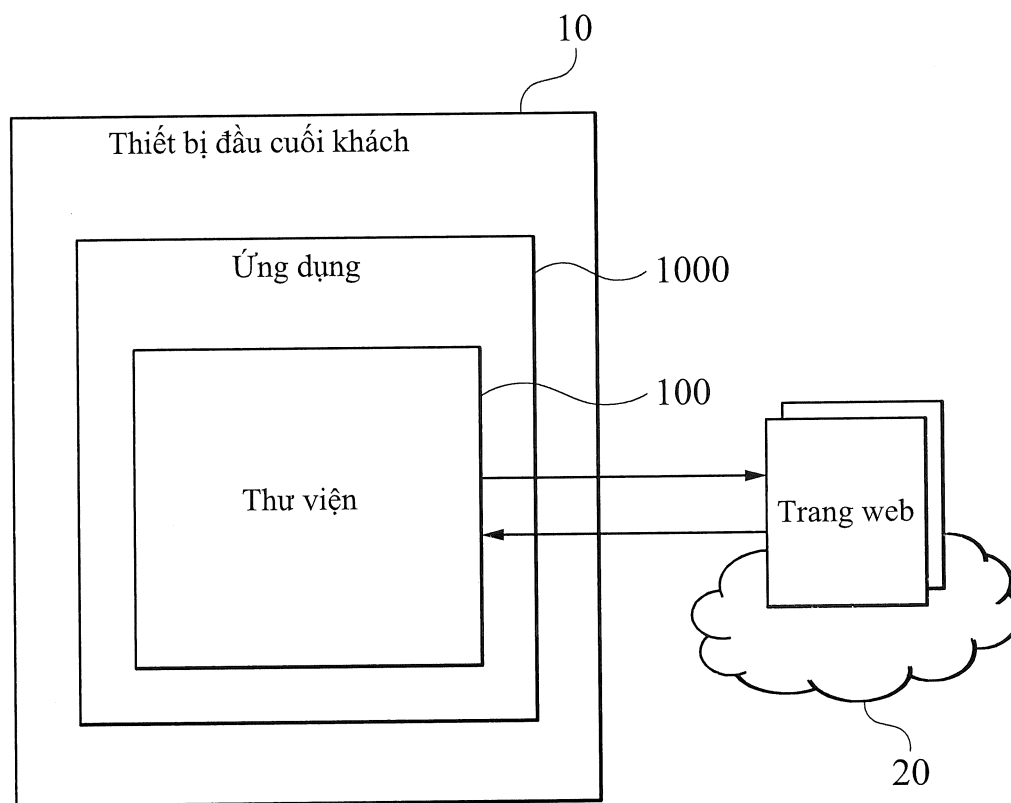


FIG. 2

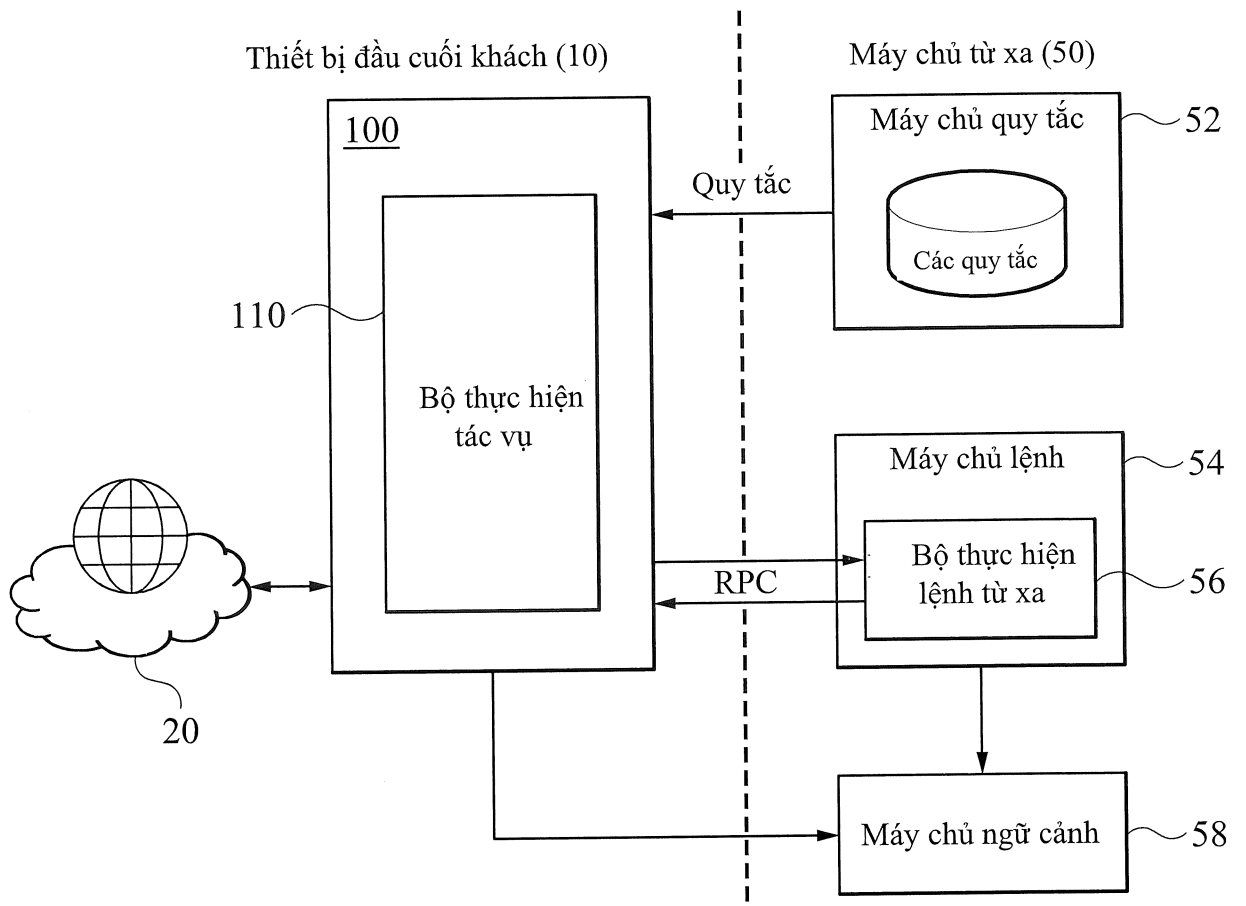


FIG. 3

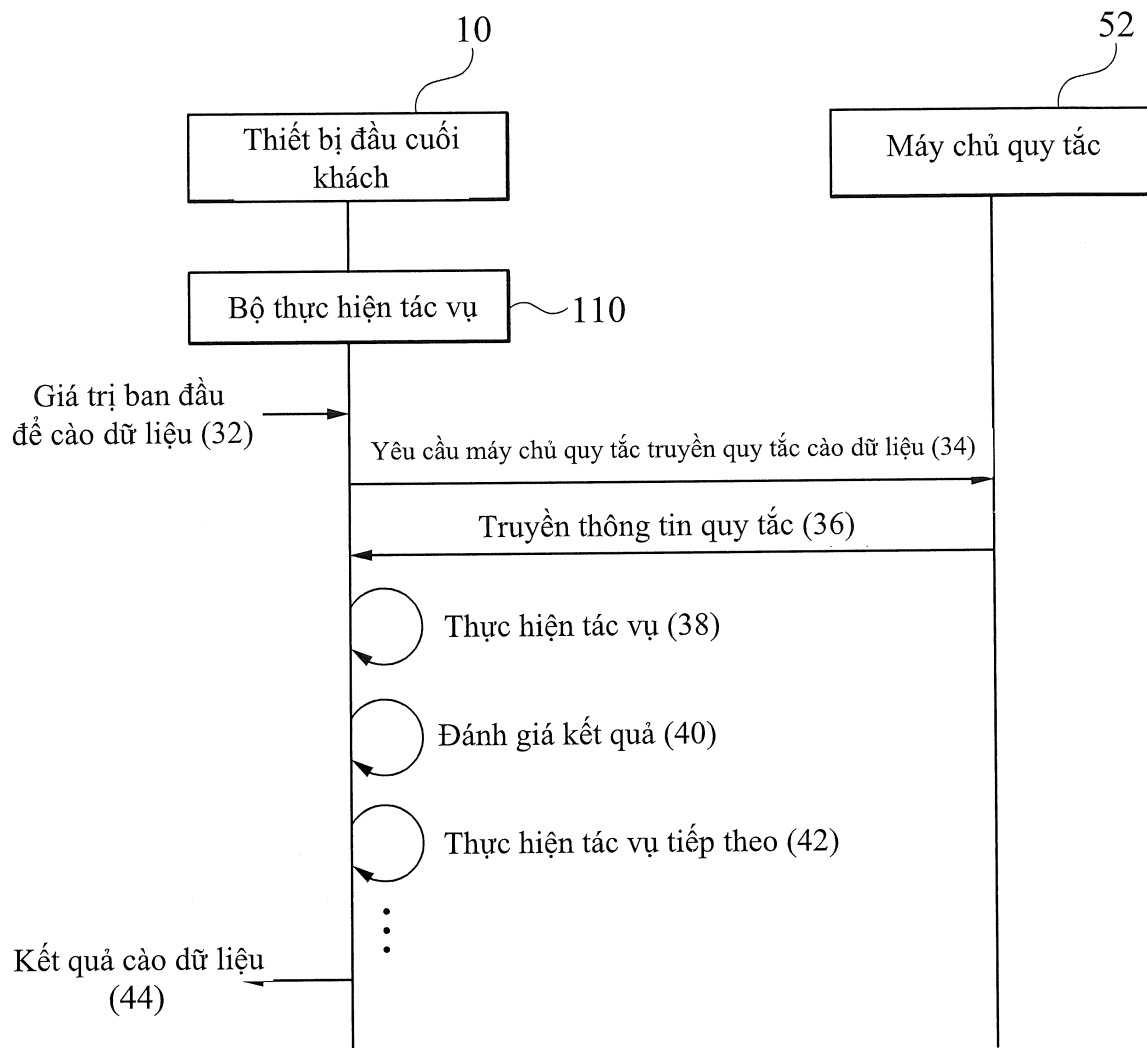


FIG. 4

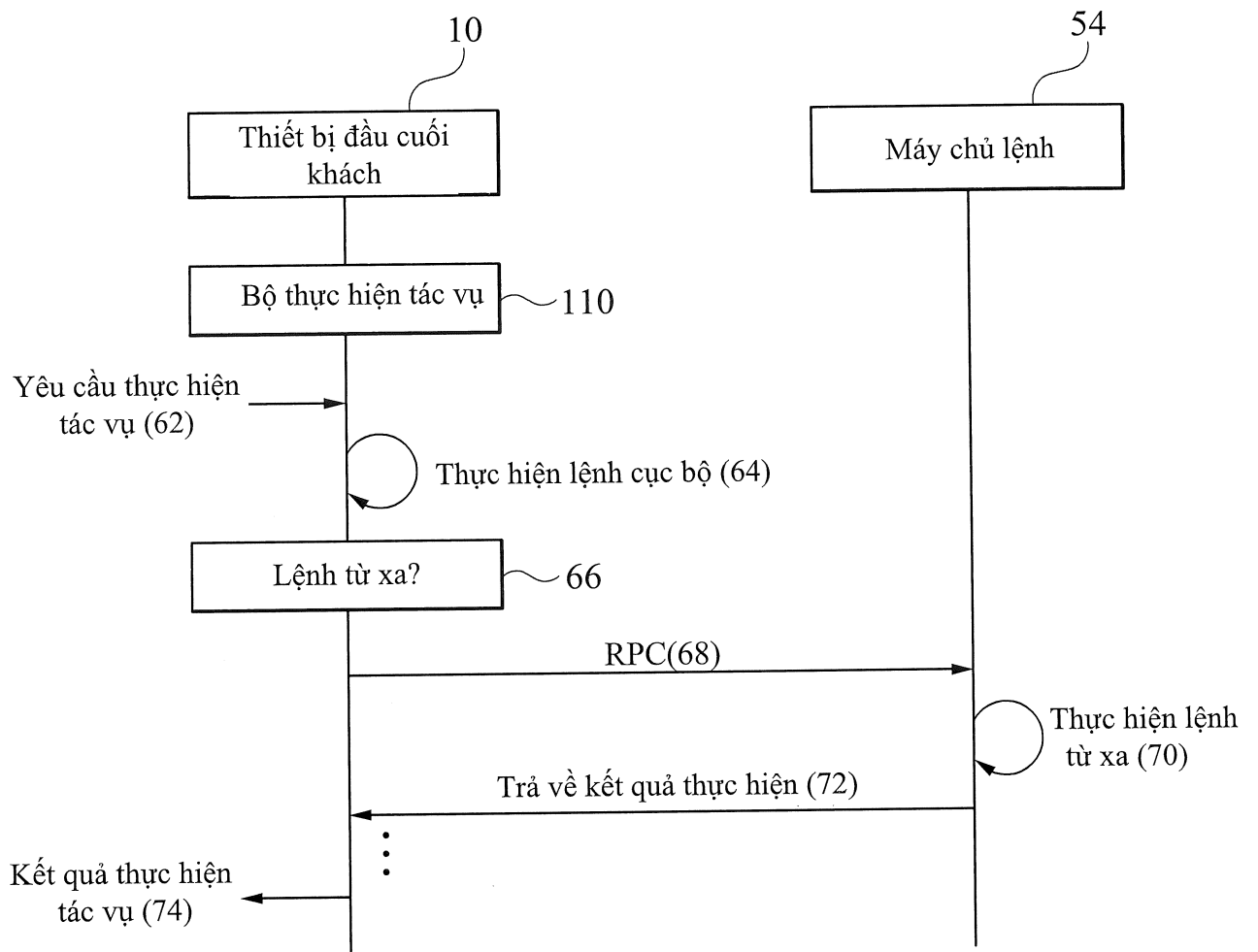


FIG. 5

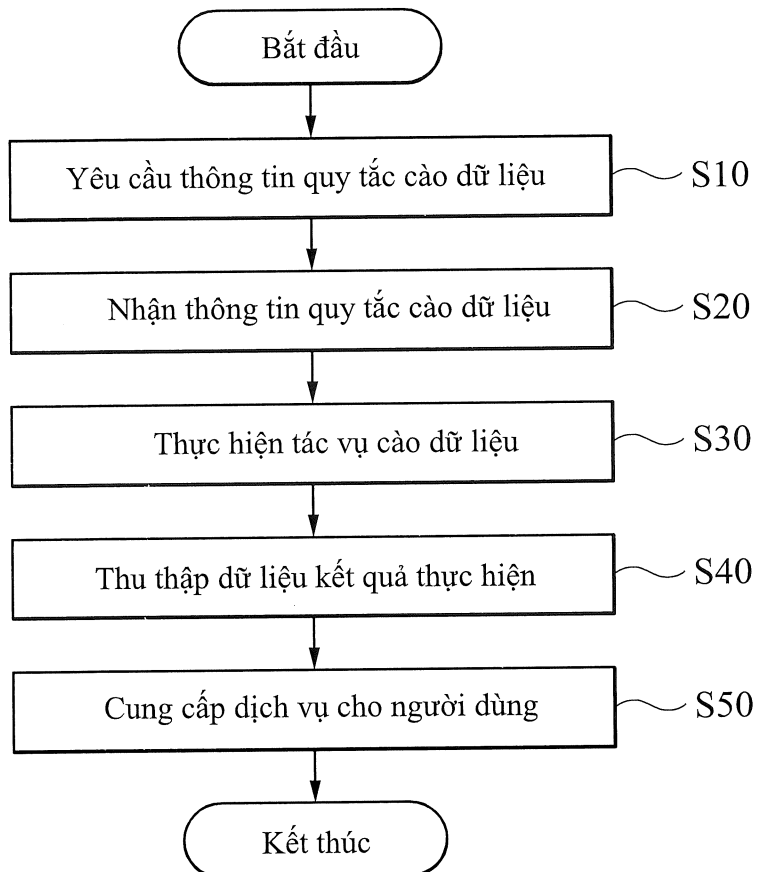


FIG. 6

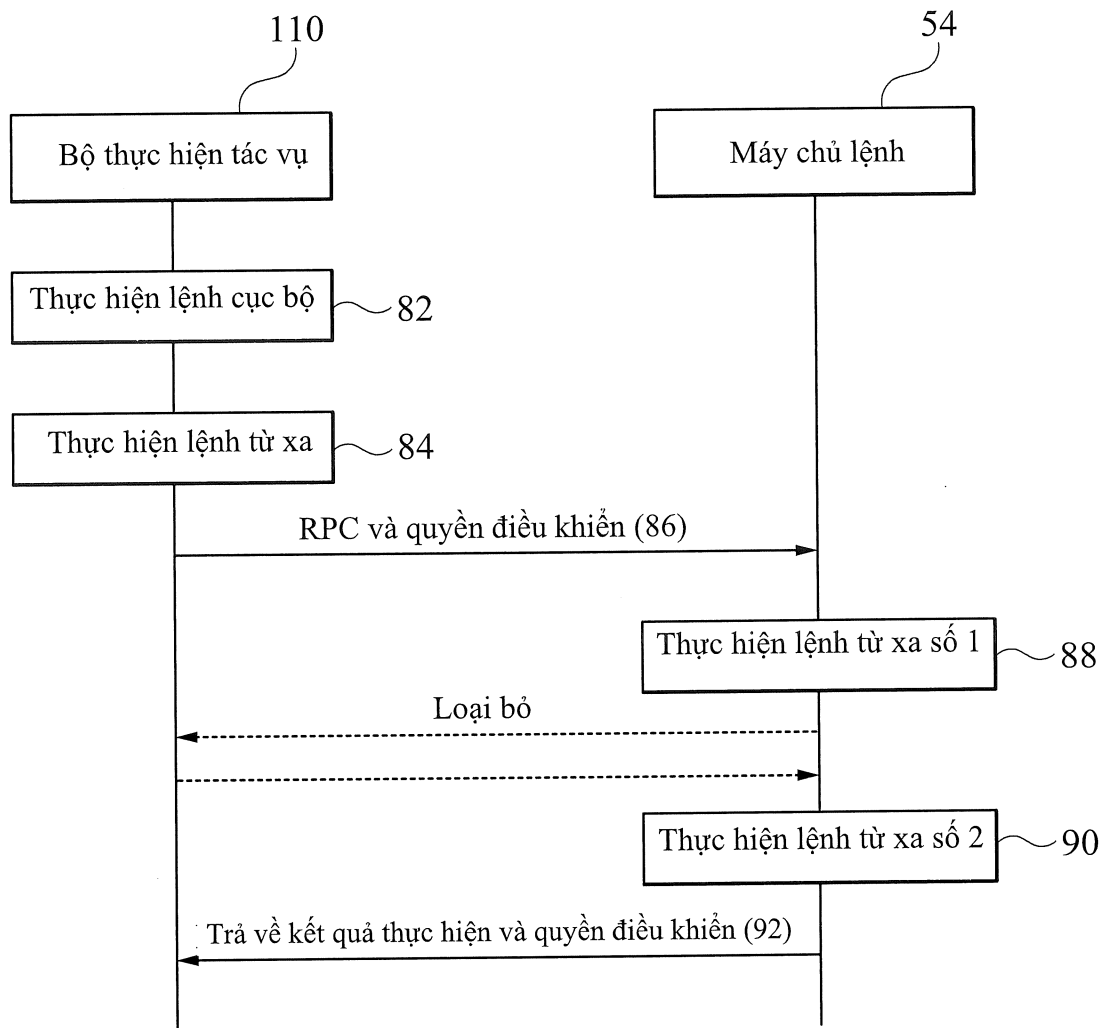


FIG. 7A

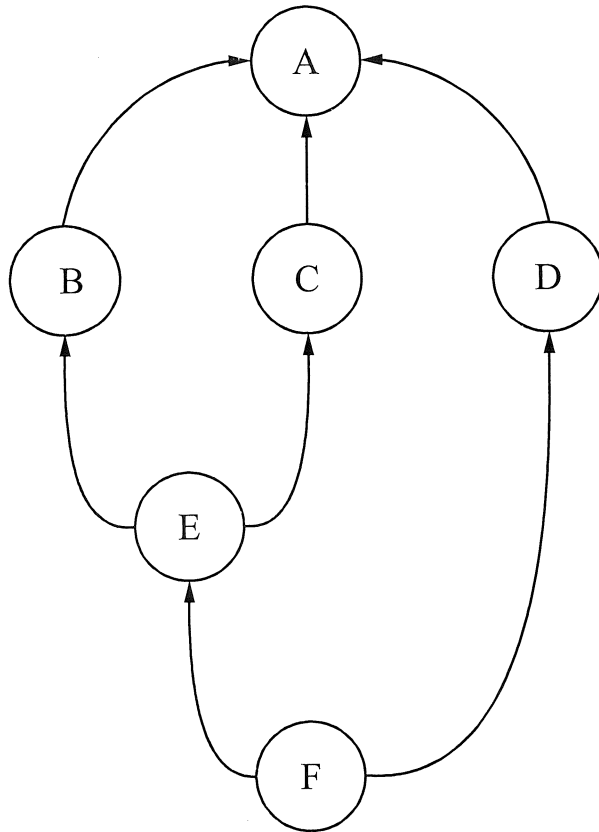


FIG. 7B

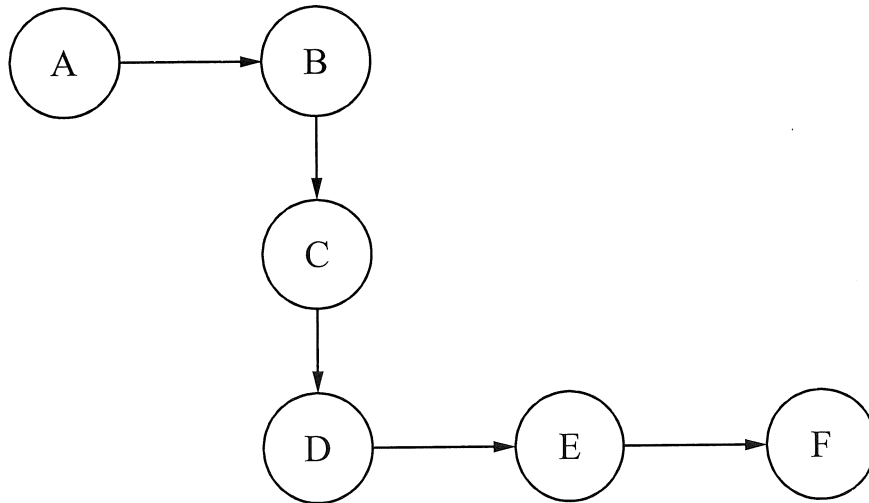


FIG. 8

