



(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)



1-0037793

(51)^{2020.01} G06F 9/00

(13) B

(21) 1-2021-06858

(22) 28/10/2021

(45) 25/12/2023 429

(43) 27/12/2021 405

(73) TẬP ĐOÀN CÔNG NGHIỆP - VIỆN THÔNG QUÂN ĐỘI (VN)

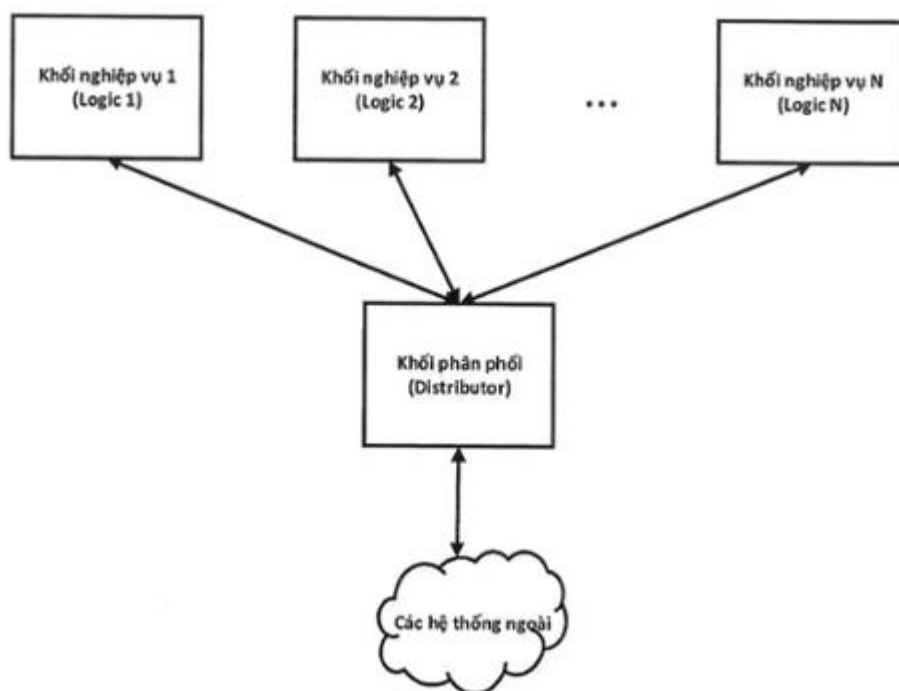
Lô D26 Khu đô thị mới Cầu Giấy, phường Yên Hoà, quận Cầu Giấy, thành phố Hà Nội

(72) Phạm Quang Điệp (VN); Lê Hồng Quân (VN); Lê Minh Tú (VN); Lê Văn Hiên (VN).

(74) Công ty TNHH NACILAW (NACILAW)

(54) PHƯƠNG PHÁP CÂN BẰNG TẢI XỬ LÝ DỰA TRÊN TRẠNG THÁI TÀI NGUYÊN VÀ TRÍCH CHỌN NGHIỆP VỤ XỬ LÝ

(57) Phương pháp cân bằng tải xử lý dựa trên trạng thái tài nguyên và trích chọn nghiệp vụ xử lý, bao gồm: bước 1: lập danh sách các khối nghiệp vụ có thể phân phối và tổ hợp danh sách; bước 2: thực hiện phân phối nghiệp vụ tới các khối nghiệp vụ; bước 3: cập nhật dữ liệu theo định kỳ để phục vụ bước 2. Phương pháp giúp chuẩn bị trước các dữ liệu và thực hiện phân phối nghiệp vụ đảm bảo cân bằng tải giữa các khối nghiệp vụ tham gia xử lý.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp cân bằng tải xử lý dựa trên trạng thái tài nguyên và trích chọn nghiệp vụ xử lý. Cụ thể, phương pháp được đề xuất áp dụng trong việc xây dựng các ứng dụng phần mềm yêu cầu dung lượng lớn và cho phép mở rộng linh hoạt.

Tình trạng kỹ thuật của sáng chế

Hiện tại trong quá trình xây dựng các ứng dụng phần mềm yêu cầu dung lượng lớn và cho phép mở rộng linh hoạt thường áp dụng mô hình khối phân phối - khối nghiệp vụ (logic). Trong đó:

- Khối phân phối: làm nhiệm vụ quản lý kết nối, giao tiếp, trao đổi các bản tin nghiệp vụ giữa ứng dụng và các hệ thống có liên quan.
- Khối nghiệp vụ: gồm nhiều khối làm chức năng xử lý nghiệp vụ của ứng dụng và trao đổi thông tin với các hệ thống có liên quan thông qua khối phân phối.
- Thông thường khối phân phối sẽ có hiệu năng cao hơn rất nhiều lần so với khối nghiệp vụ, do đó hệ thống hoàn toàn có thể mở rộng dung lượng theo chiều ngang bằng cách tăng cường số lượng khối nghiệp vụ.

Bài toán đặt ra là tải của các khối xử lý cần phân phối đều và chính xác các thông tin nghiệp vụ giao tiếp từ hệ thống ngoài vào các khối nghiệp vụ. Hiện tại, có một số giải pháp để phân phối như sau:

- Phân phối xoay vòng (round robin): xoay vòng các khối nghiệp vụ trong danh sách mỗi khi có luồng nghiệp vụ mới cần xử lý.
- Phân phối đều theo số lượng khối nghiệp vụ: dựa theo tiêu chí định trước để phân phối trải đều trên các khối nghiệp vụ: theo định danh phiên (session id), theo định danh người sử dụng (user id), theo định danh dịch vụ (service id) ...

Các phương pháp phân phối này phù hợp với các ứng dụng xử lý theo hướng sự kiện hoặc phiên ngắn, dữ liệu sau đó được cập nhật chung hoặc không cần lưu trữ. Tuy nhiên, có một số nhược điểm sau:

- Phân phối xoay vòng: với các ứng dụng yêu cầu mỗi khối nghiệp vụ quản lý dữ liệu nội tại (có thể đồng bộ thêm sang một khối nghiệp vụ khác để đảm bảo an toàn) để đáp ứng xử lý nhanh, độ trễ thấp thì không hiệu quả trong thời gian dài hoạt động. Dữ liệu trên các khối nghiệp vụ có thể được thêm/xóa trong quá trình hoạt động theo hành vi người dùng hoặc dữ liệu sử dụng, do đó, phát sinh việc có các khối nghiệp vụ có nhiều dữ liệu (thêm nhiều xóa ít), có các khối nghiệp vụ có ít dữ liệu (xóa nhiều thêm ít) không đồng đều nhưng vì theo cơ chế xoay

vòng các khối nghiệp vụ vẫn nhận đều nghiệp vụ tới => mất cân bằng tài nguyên giữa các khối nghiệp vụ xử lý.

- Phân phối đều theo số lượng khối nghiệp vụ: dữ liệu được sử dụng làm tiêu chí phân phối không được phân phối đều về mặt giá trị (trong một thời điểm hoạt động nào đó, các dữ liệu tiêu chí này tập trung vào một nhóm giá trị nhất định) dẫn tới có các khối nghiệp vụ không cân bằng tải.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất phương pháp phân phối nghiệp vụ dựa vào số lượng và tải của các khối nghiệp vụ cùng với độ ưu tiên theo yêu cầu. Phương pháp theo sáng chế bao gồm:

Bước 1: lập danh sách các khối nghiệp vụ có thể phân phối và tổ hợp danh sách; bước này được thực hiện để dữ liệu sử dụng ngay khi khởi động ứng dụng.

Bước 2: thực hiện phân phối nghiệp vụ tới các khối nghiệp vụ; các khối nghiệp vụ là các khối trong danh sách được tạo ở bước 1 (và định kỳ được cập nhật ở bước 3). Với việc phân phối như vậy đảm bảo tải của các khối nghiệp vụ xấp xỉ nhau.

Bước 3: cập nhật dữ liệu theo định kỳ để phục vụ bước 2. Kết quả của bước 3 để thường xuyên tổ hợp lại danh sách khối nghiệp vụ theo tải thực tế ở mỗi thời điểm cập nhật.

Mô tả vắn tắt các hình vẽ

Hình 1 là hình vẽ mô tả mô hình khối phân phối – khối nghiệp vụ;

Hình 2 là hình vẽ mô tả các bước khởi tạo dữ liệu;

Hình 3 là hình vẽ mô tả các bước thực hiện phân phối nghiệp vụ tới các khối nghiệp vụ;

Hình 4 là hình vẽ mô tả các bước thực hiện cập nhật dữ liệu.

Mô tả chi tiết sáng chế

Nội dung được mô tả dưới đây nhằm thực hiện phương pháp cân bằng tải xử lý dựa trên trạng thái tài nguyên và trích chọn nghiệp vụ xử lý, tham chiếu Hình 1, Hình 2, Hình 3 và Hình 4, các bước cụ thể như sau:

Bước 1: lập danh sách các khối nghiệp vụ có thể phân phối và tổ hợp danh sách. Thông tin mỗi khối nghiệp vụ bao gồm:

- + Định danh cho mỗi khối nghiệp vụ;
- + Thông tin tải hiện tại (khởi tạo là 1);
- + Thông tin độ ưu tiên: giá trị quy định 1, 2, 3 (tối đa ba loại độ ưu tiên tùy theo ứng dụng). Việc đặt độ ưu tiên có thể theo loại nghiệp vụ, loại khối nghiệp vụ.

Từ đó, lập ra danh sách phân phối theo tiêu chí:

- + Các khối nghiệp vụ có độ ưu tiên giống nhau được xếp chung nhóm danh sách.
- + Trong một danh sách, tính tương quan số lượng tải để quyết định số lần xuất hiện trong danh sách, tải càng cao thì số lần xuất hiện càng ít (có thể dựa theo tính tỷ lệ hoặc tính theo giá trị chênh lệch). Tuy nhiên, giữa số lần xuất hiện nhiều nhất và ít nhất không chênh quá 5. Thứ tự trong danh sách sắp xếp các khối nghiệp vụ xuất hiện nhiều nhất trước.

⇒ Danh sách khi khởi tạo thì mỗi khối nghiệp vụ xuất hiện một lần do có cùng thông tin tải là 1.

Bước 2: thực hiện phân phối nghiệp vụ tới các khối nghiệp vụ. Cụ thể bước này được thực hiện như sau:

+ B2.1: từ độ ưu tiên nghiệp vụ (mỗi nghiệp vụ được gán với độ ưu tiên hoặc mặc định là 1 với trường hợp không quy định), chọn ra nhóm khối nghiệp vụ tương ứng để phân phối.

+ B2.2: từ nhóm chọn được ở bước 1, áp dụng cách phân phối xoay vòng trong danh sách theo thuật toán phân phối xoay vòng (round robin) các phần tử trong danh sách

Bước 3: cập nhật dữ liệu theo định kỳ để phục vụ bước 2; kết quả của bước 3 mục đích để cập nhật lại danh sách khối nghiệp vụ theo tải thực tế.

+ Trường hợp có thay đổi khối nghiệp vụ: nếu có khối nghiệp vụ bị dừng hoạt động, ngay lập tức loại bỏ khối nghiệp vụ tương ứng khỏi danh sách phân phối. Nếu có khối nghiệp vụ mới được khởi động: thực hiện tính toán và cập nhật ở chu kỳ tiếp theo;

+ Cập nhật theo chu kỳ:

- Thực hiện cập nhật lại nhóm khối nghiệp vụ theo độ ưu tiên (nếu có thay đổi độ ưu tiên trong chu kỳ liền kề trước đó).
- Tính toán độ chênh lệch tải giữa các khối nghiệp vụ hiện có trong danh sách theo luật như ở bước 1, có hai cách tính:

Cách thứ nhất: tính theo độ tương quan của các khối nghiệp vụ với khối nghiệp vụ có tải cao nhất. Ví dụ ba khối nghiệp vụ có tải lần lượt là: 5, 12, 24
 => số lần xuất hiện của khối nghiệp vụ thứ nhất = $24 / 5 \sim 5$, số lần xuất hiện của khối nghiệp vụ thứ hai = $12 / 5 \sim 3$;

Cách thứ hai: tính theo giá trị chênh lệch tuyệt đối với khối nghiệp vụ có tải cao nhất (tự quy định giá trị chênh lệch), ví dụ: ba khối nghiệp vụ có tải lần lượt là: 5, 12, 24 và quy đổi chênh lệch 6 được tính là một lần xuất hiện => số lần xuất hiện của khối nghiệp vụ thứ nhất = $(24-5)/6 \sim 4$, số lần xuất hiện của khối nghiệp vụ thứ hai = $(24-12)/6 = 2$.

⇒ Tạo ra nhóm khối nghiệp vụ và danh sách khối nghiệp vụ tương ứng mới để phục vụ cho bước 2.

Phương pháp được đề xuất trong sáng chế cần chú ý thực hiện cập nhật các thông tin:

- Thông tin khối nghiệp vụ và độ ưu tiên: không nên để quá nhiều độ ưu tiên sẽ ảnh hưởng tới hiệu năng hệ thống, giá trị đề xuất: ≤ 3 để tối ưu.
- Chu kỳ T cần tính toán dựa trên hoạt động phù hợp để có thể cập nhật danh sách tài nguyên chính xác nhất, tránh thời gian T quá lâu có thể gây lệch tải hoặc quá ngắn thì ảnh hưởng hiệu năng do cập nhật quá thường xuyên.

Ví dụ thực hiện sáng chế

Phương pháp theo sáng chế được áp dụng trong hệ thống mạng lõi chuyển mạch gói 4G (EPC) – thành phần quản lý tính di động của thuê bao (Mobility Management Entity - MME).

Dưới đây là đồ thị tài nguyên của các phần cứng xử lý nghiệp vụ trong MME.

Hiệu quả đạt được của sáng chế

Hiệu quả đạt được của sáng chế bao gồm: Tải giữa các phần cứng triển khai xử lý nghiệp vụ xấp xỉ nhau trong quá trình hoạt động. Đảm bảo phân phối đều các nghiệp vụ và thuê bao xử lý.

Yêu cầu bảo hộ

1. Phương pháp cân bằng tải xử lý dựa trên trạng thái tài nguyên và trích chọn nghiệp vụ xử lý, bao gồm:

bước 1: lập danh sách các khối nghiệp vụ có thể phân phối và tổ hợp danh sách; thông tin mỗi khối nghiệp vụ bao gồm: định danh cho mỗi khối nghiệp vụ, thông tin tải hiện tại (khởi tạo là 1), thông tin độ ưu tiên: giá trị quy định 1, 2, 3 (tối đa ba loại độ ưu tiên tùy theo ứng dụng); việc đặt độ ưu tiên có thể theo loại nghiệp vụ, loại khối nghiệp vụ, từ đó, lập ra danh sách phân phối theo tiêu chí:

các khối nghiệp vụ có độ ưu tiên giống nhau được xếp chung nhóm danh sách; trong một danh sách, tính tương quan số lượng tải để quyết định số lần xuất hiện trong danh sách, tải càng cao thì số lần xuất hiện càng ít (có thể dựa theo tính tỷ lệ hoặc tính theo giá trị chênh lệch); tuy nhiên, giữa số lần xuất hiện nhiều nhất và ít nhất không chênh quá 5; thứ tự trong danh sách sắp xếp các khối nghiệp vụ xuất hiện nhiều nhất trước; từ đó lập được danh sách khi khởi tạo thì mỗi khối nghiệp vụ xuất hiện một lần do có cùng thông tin tải là 1;

bước 2: thực hiện phân phối nghiệp vụ tới các khối nghiệp vụ; cụ thể bước này được thực hiện như sau:

từ độ ưu tiên nghiệp vụ (mỗi nghiệp vụ được gán với độ ưu tiên hoặc mặc định là 1 với trường hợp không quy định), chọn ra nhóm khối nghiệp vụ tương ứng để phân phối;

từ nhóm chọn được ở bước 1, áp dụng cách phân phối xoay vòng trong danh sách theo thuật toán phân phối xoay vòng (round robin) các phần tử trong danh sách;

bước 3: cập nhật dữ liệu theo định kỳ để phục vụ bước 2; kết quả của bước 3 là danh sách khối nghiệp vụ được cập nhật theo tải thực tế.

trường hợp có thay đổi khối nghiệp vụ: nếu có khối nghiệp vụ bị dừng hoạt động, ngay lập tức loại bỏ khối nghiệp vụ tương ứng khỏi danh sách phân phối; nếu có khối nghiệp vụ mới được khởi động: thực hiện tính toán và cập nhật ở chu kỳ tiếp theo;

cập nhật theo chu kỳ:

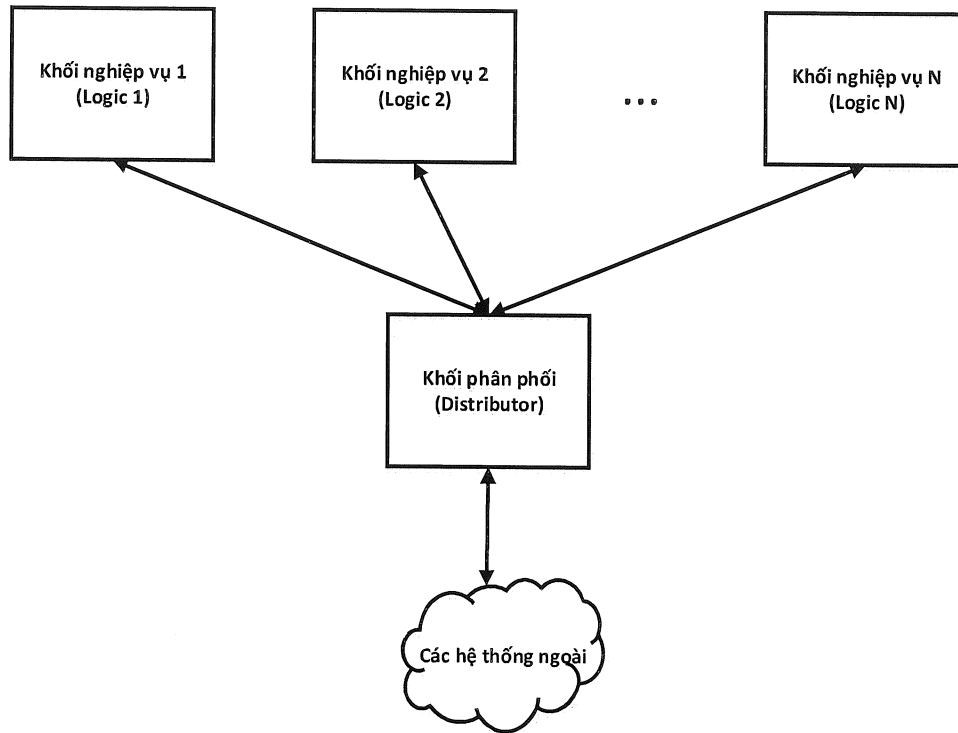
thực hiện cập nhật lại nhóm khối nghiệp vụ theo độ ưu tiên (nếu có thay đổi độ ưu tiên trong chu kỳ liền kề trước đó);

tính toán độ chênh lệch tải giữa các khối nghiệp vụ hiện có trong danh sách theo luật như ở bước 1, có hai cách tính:

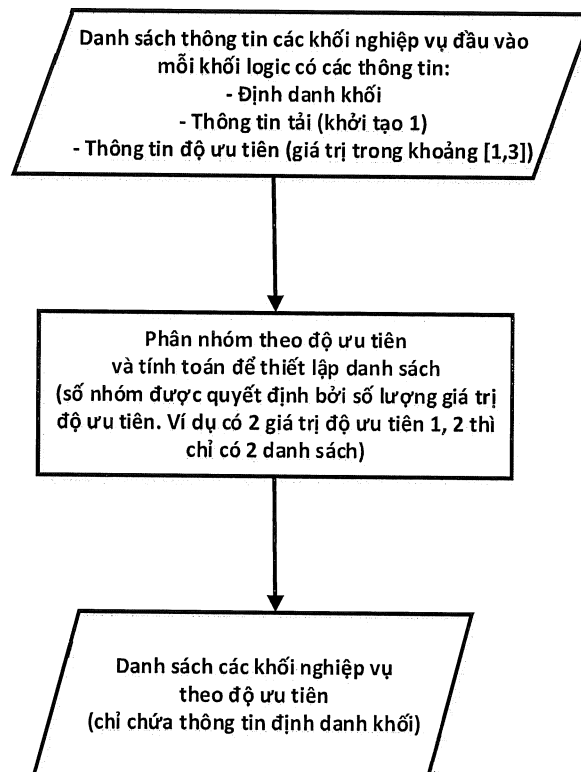
cách thứ nhất: tính theo độ tương quan của các khối nghiệp vụ với khối nghiệp vụ có tải cao nhất. ví dụ ba khối nghiệp vụ có tải lần

lượt là: 5, 12, 24 $\Rightarrow$ số lần xuất hiện của khối nghiệp vụ thứ nhất = $24 / 5 \sim 5$, số lần xuất hiện của khối nghiệp vụ thứ hai = $12 / 5 \sim 3$;
 cách thứ hai: tính theo giá trị chênh lệch tuyệt đối với khối nghiệp vụ có tải cao nhất (tự quy định giá trị chênh lệch), ví dụ: ba khối nghiệp vụ có tải lần lượt là: 5, 12, 24 và quy đổi chênh lệch 6 được tính là một lần xuất hiện $\Rightarrow$ số lần xuất hiện của khối nghiệp vụ thứ nhất = $(24-5)/6 \sim 4$, số lần xuất hiện của khối nghiệp vụ thứ hai = $(24-12)/6 = 2$;

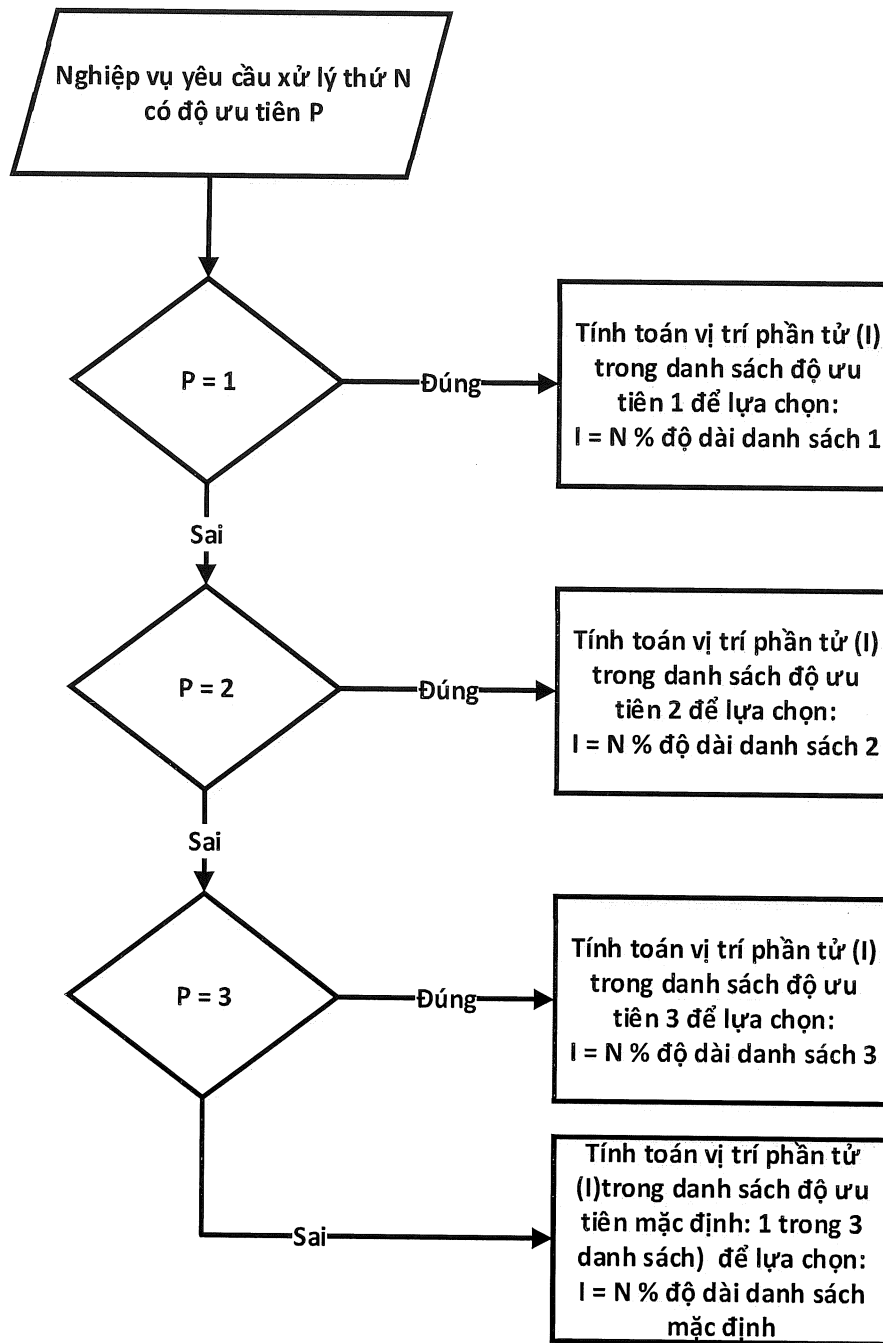
từ đó tạo ra nhóm khối nghiệp vụ và danh sách khối nghiệp vụ tương ứng mới để phục vụ cho bước 2.



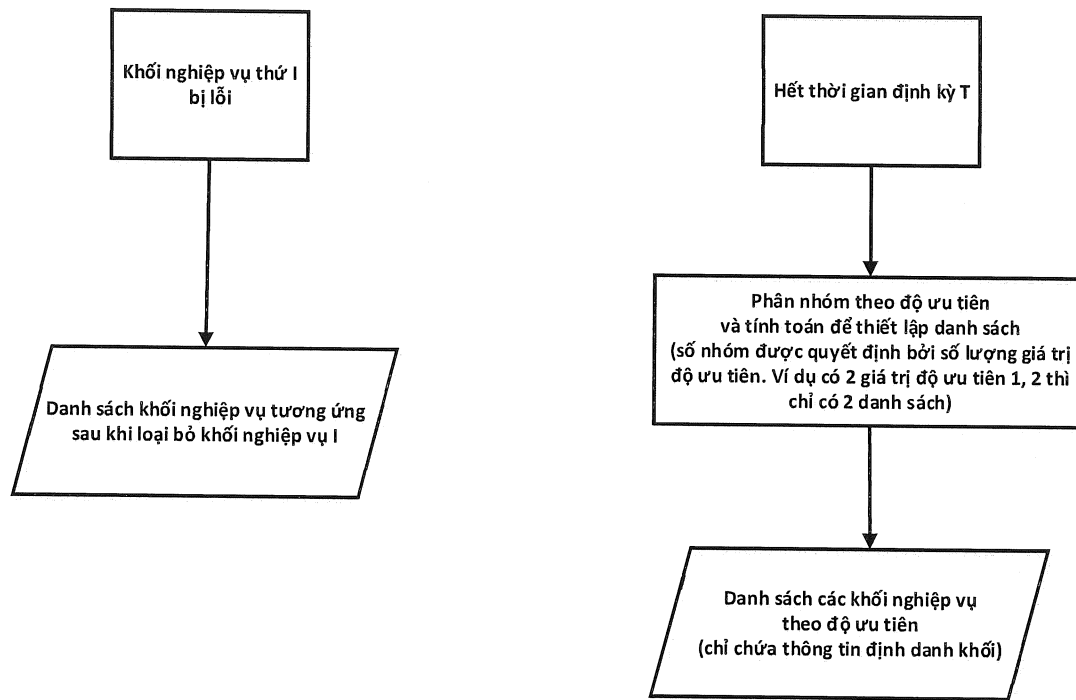
Hình 1



Hình 2



Hình 3



Hình 4