

ỨNG DỤNG CƠ SỞ DỮ LIỆU PHÂN TÁN VÀO QUẢN LÝ VÀ KIỂM SOÁT HỆ THỐNG ĐIỆN QUỐC GIA

NGUYỄN NAM HẢI & PHÍ MẠNH LỢI

Trường Đại học Bách khoa Hà nội, Viện Công nghệ thông tin

Summary. The paper presents an application of Distributed database into controlling and managing national electrical system. It allows to overcome some shortcomings in old system. The new system was constructed in opening point of view. This point of view permits the system being easy in developing and servicing for different applications. The performance of using system is higher than the old system. The cost in developing system is cheaper than traditional approach.

I. ĐẶT VẤN ĐỀ

Trong một số năm gần đây, sự phát triển mạnh mẽ về lý thuyết cũng như ứng dụng của lĩnh vực cơ sở dữ liệu phân tán đã đem lại hiệu quả rất lớn cho các ngành kinh tế, quân sự... Sự ra đời và phát triển của lý thuyết và ứng dụng CSDL phân tán đã khắc phục được khá nhiều những điểm hạn chế của CSDL tập trung, nó mang lại một sức sống mới cho CSDL. Trong bài báo này chúng tôi trình bày một ứng dụng CSDL phân tán vào điều hành và quản lý hệ thống điện quốc gia. Hệ thống được cài đặt không phải là một hệ thống độc lập hoàn toàn mà nó hoạt động song song với một hệ thống chuyên dụng. Hệ thống mới có thêm một số đặc điểm sau:

- Có khả năng quan sát hệ thống liên tục trong một chu kỳ thời gian dài hơn.
- Dễ dàng mở rộng hệ thống để đáp ứng cho những nhu cầu ứng dụng khác nhau.
- Tăng được tính độc lập cho việc xây dựng hệ thống theo dõi, quản lý và khai thác CSDL điện. (Ví dụ: tăng người khai thác dữ liệu, tăng thêm loại hình ứng dụng mới...).
- Độ tin cậy cũng như tính sẵn sàng của hệ thống được tăng thêm.

Hệ thống mới cùng với những đặc tính bổ sung kết hợp với hệ thống đang tồn tại giúp cho quá trình kiểm soát và điều hành lưới điện quốc gia có hiệu quả hơn.

II. MÔ HÌNH HỆ THỐNG HIỆN TẠI

2.1. Kiến trúc hệ thống

Để thực hiện được chức năng quản lý, theo dõi và điều hành hệ thống điện, hệ thống điện hiện nay gồm một tập các RTU (Remot Terminal Unit) và một máy chuyên dụng thực hiện chức năng quản lý và điều hành tạm gọi là Host System Unit (HSU). Giữa các RTU và HSU được nối với nhau bởi kênh liên lạc hai chiều. Các RTU có nhiệm vụ đều đặn xác định trạng thái tại điểm nó quản lý gửi về cho HSU và thực hiện các yêu cầu được gửi từ HSU. HSU nhận các thông tin dữ liệu gửi về từ các RTU. Dữ liệu này liên tục được phân tích, xử lý. Dựa trên các kết quả phân tích, HSU sẽ đưa ra những quyết định yêu cầu các RTU thực hiện. Mọi quyết định yêu cầu các RTU thực hiện. Mọi quyết định đều phải được thực hiện từ bàn điều khiển của HSU.

2.2. Phân tích hệ thống

2./2.1. Ưu điểm

Hệ thống cũ là một hệ thống chuyên dụng, được xây dựng và phát triển từ lâu nên nó có những mặt mạnh mà các hệ thống đa dụng không có được. Trong những ưu điểm thì những ưu điểm rõ nhất là :

- Có khả năng phản ứng nhanh với những trí thức mà HSU đã hiểu biết.
- Có khả năng xử lý thời gian thực.
- Không cần đến sự can thiệp của thao tác viên khi HSU đã nắm được vấn đề.

2.2.2. Nhược điểm

Bên cạnh những mặt mạnh mà hệ thống có được, những nhược điểm cũng dễ nhận thấy được từ nhiều góc độ khác nhau. Đó là:

- Hệ thống không mềm dẻo, linh hoạt,
- Việc mở rộng khó khăn tốn kém.
- Không có khả năng xây dựng, phát triển các ứng dụng mới trên hệ thống hiện tại.
- Việc áp dụng những tiến bộ kỹ thuật tin học vào rất khó khăn, phức tạp vì ngành điện không có hệ phát triển cho hệ thống hiện tại. Đây là một trong những điểm làm cho hệ thống nhanh chóng bị lạc hậu so với những phát triển nhanh chóng của kỹ thuật tin học trong thời đại hiện nay.

2.3. Tiếp cận đề xuất

Từ những phân tích trong mục 2.2 cho thấy để có thể đáp ứng một cách hiệu quả nhu cầu thực tế, hệ thống hiện tại cần có sự phát triển thích hợp. Chúng tôi tìm cách phát triển hệ thống này sao cho vẫn giữ được những ưu điểm của hệ thống cũ. Nhưng đồng thời lại khắc phục được những nhược điểm của hệ thống đó.

Vấn đề được giải quyết theo một cách: một mặt chúng tôi vẫn tiếp tục duy trì hoạt động của hệ thống cũ, mặt khác xây dựng một hệ thống mới hoạt động song song với hệ thống cũ cho phép giải quyết những vấn đề mà hệ thống cũ không có khả năng hoặc nếu giải quyết được thì phải chi phí với một mức độ tốn kém hơn nhiều. Hệ mới được thiết kế hoàn toàn mở, dễ dàng phát triển, tận dụng triệt để những thành tựu mới của kỹ thuật tin học áp dụng vào quản lý và điều hành hệ thống điện. Quan điểm tiếp cận này cho phép hệ thống luân đổi mới, phát triển dễ dàng. So với cách mở rộng truyền thống thì cách tiếp cận này cho hiệu quả kinh tế hơn. Mô tả chi tiết hệ thống sẽ được trình bày trong mục III.

III. MÔ HÌNH HỆ THỐNG MỚI

Việc can thiệp vào hệ thống cũ để cho có thể thoả mãn được những mục tiêu đặt ra là một công việc hết sức khó khăn và mạo hiểm. Do vậy, việc xây dựng hệ thống sẽ được thực hiện theo tiếp cận đã được đề xuất trong mục 2.3. Hệ thống mới của chúng tôi được phát triển theo nhiều giai đoạn. Ở giai đoạn đầu chúng tôi thực hiện:

- Xây dựng CSDL phân tán điện.
- Xử lý dữ liệu từ RTU gửi về HSU.

Mọi dữ liệu từ các RTU gửi về HSU cũng sẽ được chuyển sang phần hệ thống mới. Những dữ liệu này sẽ được phân tích, xử lý chuyển đổi thành những dạng dữ liệu thích hợp để cập nhật vào CSDL. Kết quả ở giai đoạn này chủ yếu giúp cho việc khai thác thông tin hệ thống điện và hỗ trợ cho việc điều hành được hiệu quả hơn. Trong các giai đoạn tiếp theo, hệ thống mới sẽ được phát triển để có thể trực tiếp tham gia vào quá trình điều hành hệ thống. Trong khuôn khổ bài báo này, chúng tôi chủ yếu trình bày những phát triển hệ thống trong giai đoạn 1.

3.1. Tổ chức CSDL

Trong phần này chúng tôi trình bày một số vấn đề mấu chốt khi xây dựng CSDL phân tán thông tin lưới điện quốc gia. Phần CSDL sẽ được trình bày dưới dạng tổng quát chứ không đi sâu vào mô tả chi tiết.

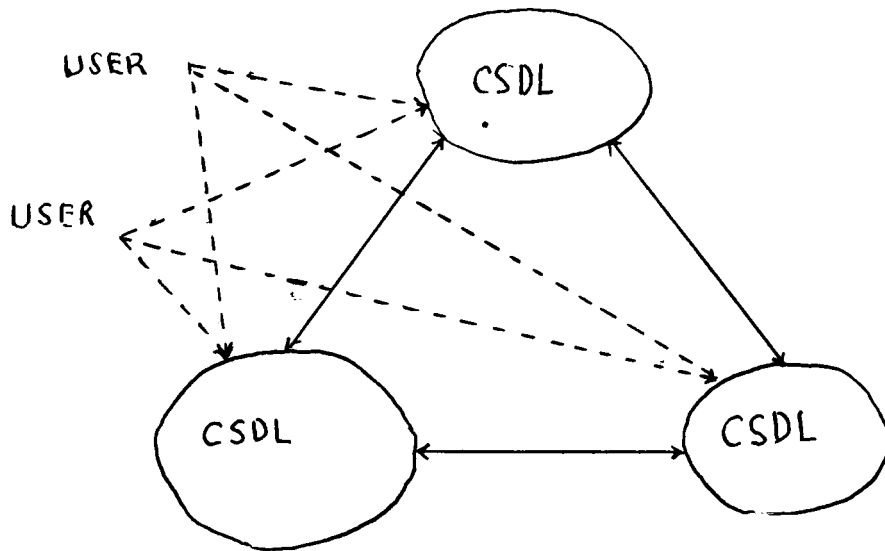
3.1.1. Tổ chức hệ thống

CSDL gồm một tập các item dữ liệu x_1, x_2, \dots, x_n nhận giá trị trong miền xác định D_1, D_2, \dots, D_n tương ứng. Các D_i không nhất thiết phải rời nhau. Một bộ (x_1, \dots, x_n) là trạng thái của CSDL. CSDL được tổ chức lập toàn phần trên 3 nút mạng. Tại mỗi trạm

được gắn một đồng hồ logic C_i ($i = 1, 2$) nhận các giá trị nguyên trong khoảng $(0, +\infty)$. Các C_i được khởi đầu giá trị 0.

Các thao tác với CSDL có 2 dạng:

- Cập nhật được gửi về từ các trạm điện và nhà máy điện.
- Truy nhập của người dùng vào CSDL qua các trạm làm việc.



Tổ chức lưu giữ CSDL

3.1.2. Xử lý cập nhật

CSDL ở đây được thiết kế lắp toàn bộ tại 3 nút mạng. Do vậy mọi cập nhật phải đảm bảo:

- Duy trì được tương thích nội tại (Internal consistency),
- Duy trì được tính tương thích tổng thể của CSDL (Mutual consistency).

Tương thích nội tại liên quan tới ngữ nghĩa của dữ liệu trong một bản sao CSDL. Tương thích tổng thể yêu cầu mọi bản sao của CSDL phải có giá trị như nhau. Do một cập nhật không thể thực hiện đồng thời tại mọi nút lưu dữ CSDL cho nên các bản sao khác nhau của CSDL không thể giống nhau hoàn toàn tại mọi thời điểm. Nhưng khi mọi cập nhật ngừng thì mọi bản sao CSDL phải có giá trị như nhau.

Để xử lý cập nhật, chúng tôi sử dụng kỹ thuật đã được trình bày trong [1]. Kỹ thuật đồng bộ ở đây được sử dụng là kỹ thuật *timestamp ordering* được trình bày trong [2], [5]. Trong sơ đồ xử lý cập nhật có tính để khả năng phân hoạch CSDL có thể xảy ra khi nút mạng hoặc kênh gặp sự cố. Vấn đề độ sẵn sàng của CSDL cũng được quan tâm đến để sao cho CSDL luôn ở trạng thái sẵn sàng phục vụ tốt. Một số kết quả về nâng cao độ sẵn sàng đã được đề cập trong [3] cũng được áp dụng trong hệ thống.

Giải thuật xử lý cập nhật có thể được trình bày một cách tóm tắt như sau:

Khi một trạm điện nào đó có dữ liệu gửi về HSU, dữ liệu đó cũng đồng thời được gửi đến một nút mạng gần nhất. Tại nút mạng đó, dữ liệu được phân tích, xử lý và biến đổi về dạng dữ liệu phù hợp với cấu trúc của CSDL. Dữ liệu sau đó được chuyển cho bộ quản trị CSDL tại nút mạng đó dưới dạng một cập nhật có dạng Update (RS.SWf) trong đó RS là tập các item dữ liệu đọc bởi cập nhật, WS là tập các item dữ liệu được viết bởi cập nhật, f là hàm mô tả các tác động lên cơ sở dữ liệu với giả định rằng f luôn chuyển CSDL từ trạng thái tương thích tới trạng thái tương thích. Cập nhật này sẽ được xử lý theo giải thuật lạc quan. Một mặt cập nhật này được xếp vào hàng đợi xử lý tại nút mạng đó, mặt khác, nó được chuyển ngay tới các nút mạng có lưu giữ bản sao CSDL. Chi tiết giải thuật xem [1]. Áp dụng giải thuật này có điểm lợi là chi phí truyền thấp vì lượng dữ liệu phải truyền trên kênh nhỏ hơn nhiều so với việc tính toán cập nhật được thực hiện tại nút mạng sau đó mới gửi kết quả tính toán được cho nút mạng khác. Sự cố tại một nút mạng không làm cho hệ thống ngừng.

3.1.3. Xử lý cập nhật không còn nhất thiết phải duy trì

Xử lý cập nhật không còn nhất thiết phải duy trì là một vấn đề không kém phần quan trọng trong các hệ quản trị CSDL phân tán. Trong CSDL phân tán, tính tương thích tổng thể của các bản sao CSDL tại các nút mạng được đảm bảo bằng việc xác định một thứ tự tổng thể trên mọi cập nhật đã được phát ra. Mỗi bản sao CSDL tại thời điểm cho trước bất kỳ phải phản ánh được mọi cập nhật mà nó đã nhận được cho đến thời điểm đó. Để duy trì được tính tương thích tổng thể cho CSDL thì mọi nút mạng có lưu giữ bản sao CSDL cuối cùng phải nhận được mọi cập nhật đã phát ra các trạm làm việc. Do vậy, các bản sao cập nhật cần phải được duy trì để truyền lại cho các nút

hông hoặc kênh truyền đến nút đó gặp sự cố. Hơn nữa vì các cập nhật không đến theo thứ tự thời gian do vậy tại các nút mạng có lưu giữ bản sao CSDL cần thiết để xử lý những cập nhật đến muộn một cách chính xác. Nhưng thông tin kiểu như vậy tăng lên theo thời gian và có thể chiếm hết bộ nhớ lưu giữ tại đó nếu chúng không được bỏ đi một cách thích hợp.

Có một số thuật xử lý loại bỏ những thông tin dạng trên đã được đề xuất [4], [6], [7]. Chọn giải thuật nào cho phù hợp với mô hình dữ liệu cũng là một vấn đề cần phải xem xét. Qua tìm hiểu một loạt các giải thuật chúng tôi đã quyết định sử dụng giải thuật đã được trình bày trong [4] với một tí thay đổi cho phù hợp với mô hình CSDL thực tế. So với những tiếp cận khác thì tiếp cận này có tính hệ thống hơn. Các tiếp cận này quan tâm tới cả nhu cầu quản trị thông tin quá khứ và trạng thái của kênh truyền giữa các trạm. Cách tiếp cận này có thể trình bày một cách tóm tắt như sau.

Một nút nào đó khi đã nhận và xử lý một cập nhật với *timestamp* T. Nút đó có thể loại bỏ không cần lưu giữ cập nhật đó nữa khi biết chắc chắn các điều kiện sau được thoả mãn:

- (Đk1). Mọi nút đã nhận được cập nhật đó.
- (Đk2). Không có cập nhật nào đang trên đường tới nút đi với *timestamp* nhỏ hơn T.
- (Đk3). Không còn nút nào phát ra thêm những cập nhật với *timestamp* nhỏ hơn T.

Điều kiện 1 để chỉ ra rằng cập nhật đó không còn cần thiết phải truyền lại tới những nút chưa nhận được nó. Đk2 . Đk3 để chỉ ra không còn cập nhật nào nữa với *timestamp* nhỏ hơn T sẽ được nhận bởi nút đang lưu trữ cập nhật đó. Như vậy, khi một cập nhật thoả mãn 3 điều kiện trên sẽ không còn cần thiết phải duy trì nữa. Những cập nhật kiểu như vậy có thể loại bỏ mà không ảnh hưởng đến việc duy trì tính tương thích của CSDL.

Để xác định T, chúng tôi sử dụng giải thuật *snapshot* phân tán đã được Chandy và Lamport trình bày trong [8].

IV. ỨNG DỤNG

Các dữ liệu về hệ thống điện là hết sức quan trọng cho các định hướng phát triển ngành điện. Do vậy, có rất nhiều những nghiên cứu, ứng dụng cần đến dữ liệu của CSDL điện. Với hệ thống cũ khả năng đáp ứng cho nhu cầu này bị hạn chế với một lượng khai

thác không nhiều. Với hệ thống mới, khả năng mở rộng người khai thác CSDL điện dễ dàng hơn. Tại mọi thời điểm CSDL có thể phục vụ cho nhiều loại người dùng khác nhau tại các vùng địa lý khác nhau. Đây là một ưu điểm mà hệ thống cũ không có được. Với cách tổ chức này, chúng ta cũng dễ tăng thêm được độ tin cậy cũng như tính sẵn sàng của CSDL. Hệ thống được thiết kế hoàn toàn mở, dễ dàng phát triển.

V. KẾT LUẬN

Bài báo này trình bày kết quả ứng dụng một số phát triển lý thuyết CSDL phân tán vào thực tiễn của ngành điện. Do tính chất đặc biệt của hệ thống điện quốc gia, việc thay thế hệ thống mới cài đặt vào hệ thống chuyên dụng cũ chưa thể tiến hành ngay được. Tuy nhiên, các kết quả thử nghiệm thu được đã chứng tỏ hệ thống mới đã đạt được những mục tiêu thiết kế đã đặt ra.

TÀI LIỆU THAM KHẢO

1. Nguyễn Nam Hải, *Nâng cao hiệu quả cập nhật CSDL phân tán*, Báo cáo tại hội nghị vô tuyến điện tử lần thứ V.
2. Nguyen Nam Hai & Nguyern Thuc Hai, *The notion of Vector time in distributed computation*, Proceedings of the NCST of Vietnam, V. 7, n. 1, 1995.
3. Nguyễn Nam Hải, *Một sơ đồ mới nhằm tăng độ sẵn sàng trong các hệ CSDL phân tán*, Tạp chí Tin học và Điều khiển học, T. 10, n. 2, 1994.
4. Sarin S.K. & Lynch N.A., *Discarding obsolete information in a replicated database system*, IEEE Trans. on Software eng. V. SE-13, n. 1 1987.
5. Lamport L., *Time, clocks and the orderings of events in a distributed systems*, Comm. ACM V. 21, 1978.
6. Birrell A.D., Levin R., Needham R.M. & Schroeder M.D., *Grapevine: An exercise in distributed comuting*, Com. ACM V. 25, n. 4, 1982.
7. Johson P.R. & Thomas R.H., *The maintenaince of duplicated database*, Arpanet request for comments (RFC)667, jan. 1975.
8. Chandy K.M. & Lamport L., *Distributed snapshots: Determing global states of distributed systems*, ACM Trans. Comput. system, V 3, n. 1, 1985.