

Building database of WEBGIS for the exchange of marine data between Vietnam and ASEAN countries

Do Huy Cuong*, Nguyen The Luan, Pham Hong Cuong, Nguyen Xuan Tung,
Bui Thi Bao Anh, Nguyen Thi Nhan

Institute of Marine Geology and Geophysics, VAST, Vietnam

*E-mail: dhcuong@imgg.vast.vn

Received: 25 July 2019; Accepted: 6 October 2019

©2019 Vietnam Academy of Science and Technology (VAST)

Abstract

The system of oceanic database management and exchange is built with the purpose of exchanging the oceanic data between Vietnam and other ASEAN countries. The system can meet the demand of sharing and exchanging oceanic data through internet connection between Vietnam and ASEAN member countries. Besides, the system of oceanic database management and exchange can meet the demands of researches, managements and share of data domestically and internationally. In this paper, we focused on the details of the system of oceanic database management and exchange, such as hardware and software, data storage, data format and data structure, data management and integration, and other issues of interface, security, standards. The WEBGIS oceanic thematic database is properly built and managed for exchanging data domestically and with other ASEAN countries.

Keywords: WebGIS, ASEAN, marine data.

Xây dựng cơ sở dữ liệu WEBGIS phục vụ trao đổi dữ liệu biển giữa Việt Nam với các nước ASEAN

Đỗ Huy Cường*, Nguyễn Thế Luân, Phạm Hồng Cường, Nguyễn Xuân Tùng, Bùi Thị Bảo Anh, Nguyễn Thị Nhân

Viện Địa chất và Địa vật lý biển, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, Việt Nam

*E-mail: dhcuong@imgg.vast.vn

Nhận bài: 25-7-2019; Chấp nhận đăng: 6-10-2019

Tóm tắt

Hệ thống quản lý và trao đổi dữ liệu biển được xây dựng với mục đích chia sẻ dữ liệu biển giữa Việt Nam và các nước ASEAN. Hệ thống đáp ứng yêu cầu chia sẻ, trao đổi dữ liệu thông qua kết nối mạng giữa Việt Nam và các nước ASEAN. Bên cạnh đó, hệ thống quản lý và trao đổi dữ liệu biển đáp ứng yêu cầu về công tác nghiên cứu, quản lý và chia sẻ dữ liệu trong nước và quốc tế. Trong bài báo này chúng tôi tập trung trình bày các chi tiết của hệ thống quản lý và trao đổi dữ liệu biển, như phần cứng và phần mềm, bộ lưu trữ số liệu, định dạng và cấu trúc dữ liệu, quản lý và tích hợp dữ liệu, cũng như các nội dung khác như giao diện, bảo mật và các dạng dữ liệu tiêu chuẩn. Cơ sở dữ liệu WEBGIS với các số liệu chuyên đề về hải dương học được thành lập và quản lý phù hợp với việc trao đổi số liệu trong nước và với các nước ASEAN.

Từ khóa: WebGIS, ASEAN, dữ liệu biển.

MỞ ĐẦU

Các mục tiêu đặt ra của hệ thống trao đổi và chia sẻ dữ liệu là xây dựng cơ sở dữ liệu biển của Việt Nam phục vụ công tác trao đổi, chia sẻ thông tin, dữ liệu biển với các cơ quan ban ngành trong nước cũng như chia sẻ dữ liệu với cơ sở dữ liệu chung của các nước ASEAN. Việc chia sẻ dữ liệu phải đảm bảo tính bảo mật và hoạt động độc lập của hệ thống tại Việt Nam, phù hợp với quy định về trao đổi dữ liệu của Việt Nam và quốc tế. Hệ thống đáp ứng yêu cầu nhiệm vụ, đồng thời đáp ứng yêu cầu về kỹ thuật trên các phương diện như tổ chức lưu trữ, quản lý dung lượng lớn, đáp ứng tiêu chuẩn trao đổi dữ liệu trong và ngoài nước. Hệ thống đảm bảo đáp ứng yêu cầu truy xuất tốc độ cao, bảo mật, phân quyền khai thác dữ liệu biển.

Đã có một số công trình nghiên cứu xây dựng cơ sở dữ liệu về dữ liệu biển, như cơ sở dữ liệu thông tin cơ bản về biển, về tài nguyên,

địa chất, môi trường và dữ liệu phục vụ các lĩnh vực và nghề khác nhau của quốc gia [1]. Tuy nhiên, sản phẩm các công trình nghiên cứu trên thực hiện theo từng giai đoạn, có công nghệ và kỹ thuật cũng như dữ liệu phù hợp với giai đoạn đó. Đặc biệt, sản phẩm không là công (portal) đại diện và phù hợp với chuẩn cho việc trao đổi và nối mạng theo chương trình trao đổi giữa các nước ASEAN và các quốc gia lân cận. Viện Địa chất và Địa vật lý biển thuộc Viện Hàn lâm Khoa học và Công nghệ Việt Nam chịu trách nhiệm xây dựng hệ thống trao đổi thông tin dữ liệu biển và nối mạng tại Việt Nam và lấy tên viết tắt là VNODC.

TỔNG QUAN VỀ QUẢN LÝ CƠ SỞ DỮ LIỆU BIỂN TRÊN THẾ GIỚI VÀ VIỆT NAM

Tổng quan về quản lý cơ sở dữ liệu biển trên thế giới

Trên thế giới, công tác nghiên cứu xây dựng cơ sở dữ liệu về tài nguyên môi trường biển đã có từ lâu và rất phát triển. Có nhiều trung tâm cơ sở dữ liệu quốc gia và quốc tế được thành lập nhằm cung cấp các dữ liệu cho phát triển kinh tế và quốc phòng. Trung tâm dữ liệu thế giới WDC (World Data Center) [2] ngày đầu được thành lập năm 1958 nhằm chia sẻ dữ liệu quan trắc vật lý địa cầu quốc tế và được đặt tại Hoa Kỳ, Châu Âu, Liên Xô và Nhật Bản. Sau này, do yêu cầu thực tiễn và sự phát triển của khoa học cũng như nhu cầu về chia sẻ dữ liệu địa cầu nói chung và dữ liệu biển nói riêng. Trung tâm dữ liệu quốc tế được phát triển mạnh theo nhiều hướng, trong đó có Trung tâm dữ liệu biển.

Trung tâm dữ liệu biển thế giới WOD [3, 4] là nơi trao đổi thông tin, dữ liệu hải dương học quốc tế, dữ liệu từ các trạm quan trắc quốc tế, dữ liệu công bố của các nghiên cứu, quốc gia, dự án liên quan đến Hải dương học. Cơ sở dữ liệu WOD hàng năm được cập nhật mới, đối tượng tham gia ngày càng nhiều, điều đó cho thấy sức hút, sự hấp dẫn của mô hình và nhu cầu trao đổi thông tin về hải dương học ngày càng cao.

Thống nhất quản lý dữ liệu hải dương trong từng quốc gia, xây dựng các trung tâm dữ liệu hải dương, quản lý vốn dữ liệu chung để nâng cao hiệu quả sử dụng. Chủ trương này được các quốc gia và các tổ chức quốc tế tiếp nhận và thực hiện ngày càng mạnh, với sự thành lập các trung tâm dữ liệu hải dương quốc gia, khu vực, toàn cầu. Hiện nay trên thế giới đã có trên 60 nước thành lập các trung tâm dữ liệu hải dương quốc gia (NODC), nhiều trung tâm dữ liệu hải dương khu vực (RNODC - Thái Bình Dương, Địa Trung Hải, Bắc Đại Tây Dương...) và 4 trung tâm dữ liệu hải dương thế giới (WDC): WDC-A ở Hoa Kỳ (Washington), WDC-B ở Nga (Moskva), WDC-C ở Australia, WDC-D ở Trung Quốc (Thiên Tân),... và còn đang hình thành các WDC mới. Ngoài ra còn có các trung tâm dữ liệu hải dương chuyên đề quốc tế được hình thành như các trung tâm dữ liệu về độ mặn và nhiệt độ (GTSP), về mực nước biển (từ các chương trình GLOSS, PSMSL) [5].

Đối với các nước lân cận, có quan hệ với nước ta cũng đã hình thành các trung tâm dữ liệu hải dương quốc gia như Trung Quốc

(CNODC), Nhật Bản (JNODC), Pháp (SISMER), Ấn Độ (INODC)... và đã đi vào hoạt động từ nhiều năm nay. Trung tâm dữ liệu hải dương quốc gia Hoa Kỳ do NOAA quản lý thường xuyên có trang Web đưa dữ liệu hải dương phổ thông lên mạng Internet để sử dụng trong nước và nước ngoài.

Tăng cường trao đổi dữ liệu hải dương giữa các quốc gia và tổ chức quốc tế để thúc đẩy hoạt động này, tổ chức IOC đã có một chương trình trao đổi dữ liệu thông qua mạng NODC ở các nước. Chương trình IODE/IOC cũng đã thành lập các WDC, tổ chức các chương trình đào tạo chuyên viên quản lý dữ liệu cho các nước (do Nhật Bản tài trợ hàng năm) xây dựng các phần mềm chuyên dụng (OCEAN-PC) và gần đây chuẩn bị xây dựng chính sách trao đổi dữ liệu hải dương quốc tế, nhằm thúc đẩy hơn nữa các hoạt động này trong các nước thành viên. Có thể thấy rằng các xu thế trên đây về hiện đại hóa, đổi mới việc quản lý và mở rộng giao lưu sử dụng dữ liệu đang là xu thế tất yếu của sự phát triển khoa học về tài nguyên môi trường thế giới, nhằm thúc đẩy hợp tác giữa các quốc gia cùng khai thác sử dụng vốn dữ liệu chung để giải quyết có hiệu quả hơn các vấn đề về biển và đại dương toàn cầu đang đặt ra cho toàn khu vực và trên thế giới. Để tham gia vào chương trình hợp tác ASEAN về Quản lý xung đột tiềm tàng ở Biển Đông và thỏa thuận xây dựng đề án “Trao đổi thông tin dữ liệu biển và nối mạng tại khu vực Biển Đông”, các nước trong khu vực đã tiến hành xây dựng Trung tâm dữ liệu riêng phục vụ mục tiêu này (tại địa chỉ website <http://www.coi.gov.cn/scs>).

Các nước ASEAN và Việt Nam đều rất quan tâm đến việc tham gia vào đề xuất này. Vì vậy, từ hội nghị lần thứ 22 năm 2012 cho đến hội nghị lần thứ 28 năm 2018 các nước đã khẳng định quyết tâm cao trong việc thúc đẩy việc tham gia vào xây dựng đề án “Trao đổi thông tin dữ liệu biển và nối mạng tại khu vực Biển Đông” của từng nước.

Tình hình quản lý cơ sở dữ liệu biển ở Việt Nam

Thông qua các khảo sát về tình hình sử dụng phần mềm trong lưu trữ và quản lý cơ sở dữ liệu biển có thể nhận thấy tại Việt Nam đã xây dựng một số phần mềm quản lý dữ liệu chuyên ngành, nhưng giao diện và công năng

của chúng còn hạn chế chưa đáp ứng cho nhu cầu trao đổi với các trung tâm dữ liệu trên thế giới. Kho tài liệu biển là kết quả nghiên cứu khoa học của các chương trình nghiên cứu cấp Quốc gia trong vòng gần 30 năm chưa được đầu tư để xây dựng một cơ sở dữ liệu hoàn chỉnh phục vụ khai thác một cách hiệu quả. Các dữ liệu chỉ mới có riêng rẽ từng mảng, chưa đồng nhất về mặt không gian cho toàn vùng và cũng chưa đồng nhất về mức độ theo một định dạng thống nhất. Tình hình xây dựng, quản lý và khai thác cơ sở dữ liệu về tài nguyên môi trường biển trong thời gian qua có thể thấy một số điểm sau đây [1–3, 6]:

Dữ liệu điều tra khảo sát biển đang được lưu giữ, quản lý ở các cơ quan nhìn chung còn chưa chính quy, hiện đại. Kỹ thuật máy tính chỉ mới được ứng dụng vào quản lý dữ liệu từ sau năm 1990. Các quy định về giao nộp, sử dụng dữ liệu ở nhiều cơ quan còn chưa chặt chẽ, chất lượng dữ liệu còn chưa kiểm soát được. Nhiều cơ quan còn chưa hình thành được bộ phận chuyên trách quản lý dữ liệu, mà thường kết hợp với công tác thư viện. Nhìn chung công tác dữ liệu còn chưa được đầu tư đúng mức cho việc thu nhận, lưu giữ, kiểm soát chất lượng cũng như nâng cao hiệu quả sử dụng. Tình hình này có nhiều nguyên nhân, nhưng chủ yếu là do từng cơ quan đều chưa có những luật lệ, quy chế rõ ràng, chặt chẽ về việc giao nộp, quản lý, sử dụng dữ liệu nói chung và dữ liệu khảo sát nói riêng. Tình trạng sở hữu cục bộ, coi dữ liệu khảo sát là tài sản riêng của cơ quan, thậm chí của cá nhân là rất nặng nề, điều này dẫn đến những tác động tiêu cực đối với việc sử dụng dữ liệu, như thương mại hóa dữ liệu, hạn chế việc sử dụng, gây khó khăn cho việc nghiên cứu.

Trình độ quản lý dữ liệu ở các cơ quan nhìn chung còn hạn chế. Công nghệ thông tin đã được ứng dụng nhưng chưa có những công cụ kỹ thuật thích hợp, thống nhất cho việc quản lý chung và thiếu hẳn sự đầu tư cho một cơ quan chuyên trách thường xuyên nghiên cứu sáng tạo, cải tiến, nâng cấp các công cụ kỹ thuật quản lý dữ liệu biển đáp ứng yêu cầu công tác này trong cả nước. Tình hình này đã hạn chế việc giao lưu, trao đổi dữ liệu giữa các cơ quan trong nước.

Quan hệ trao đổi dữ liệu với các nước và các tổ chức quốc tế chỉ mới bắt đầu, chưa

thường xuyên và hiệu quả chưa cao. Chúng ta dường như còn chưa nhập cuộc thực sự với hoạt động trao đổi dữ liệu quốc tế vì vậy chưa tranh thủ được nhiều sự hỗ trợ quốc tế về mặt này.

Tình hình quản lý dữ liệu biển, trình độ kỹ thuật về quản lý, khai thác sử dụng ở nước ta còn chậm so với trình độ phát triển của thế giới, cả về kỹ thuật và tổ chức. Nhà nước chưa có một chính sách rõ ràng toàn diện, thống nhất về quản lý và khai thác dữ liệu biển, chưa có được những quy chế chặt chẽ được các cơ quan các ngành tuân thủ, cộng đồng khoa học đồng tình nhất trí. Vì vậy cũng chưa thường xuyên nắm được tình hình dữ liệu biển hiện có trong từng giai đoạn để có quyết định đúng trong việc đầu tư khảo sát. Cũng chưa có một tổ chức quốc gia nào được giao nhiệm vụ thống nhất quản lý dữ liệu biển làm nòng cốt và điều hành việc thực hiện các quy định của Nhà nước, cũng như tham gia hoạt động trao đổi dữ liệu hải dương quốc tế với danh nghĩa quốc gia.

Hệ thống quản lý dữ liệu thông tin về các nhiệm vụ nghiên cứu khoa học và công nghệ biển là một hệ thống hợp nhất các dữ liệu nghiên cứu khoa học của các ngành thuộc lĩnh vực biển và tài nguyên môi trường. Hệ thống được tổ chức theo mô hình dữ liệu không gian hướng đối tượng của hệ thống thông tin địa lý (GIS), các dữ liệu được tích hợp trên nền thành phần cơ bản là hạ tầng thông tin địa lý, các chức năng cơ bản của hệ thống là cập nhật thông tin, quản lý, phân tích, trình bày và phân phối thông tin nhằm phục vụ các cơ quan quản lý nhà nước và các tổ chức ở cấp Trung ương, các Viện, Bộ, Ngành, các nhu cầu sử dụng trong nghiên cứu khoa học,...

Cơ sở dữ liệu hiện được nhiều cơ quan và địa phương lưu giữ và quản lý. Tình hình quản lý dữ liệu mang tính cục bộ, coi dữ liệu khảo sát và nghiên cứu là của riêng, dẫn đến tiêu cực như thương mại hóa dữ liệu, do đó các thành quả khảo sát và nghiên cứu không được sử dụng hợp lý, gây tổn hại cho kinh tế quốc gia.

Việc tập hợp và xử lý dữ liệu trong một định dạng thống nhất là chưa có.

Các phần mềm quản lý được xây dựng cho một số ít lĩnh vực riêng rẽ chưa có sự thống nhất chung.

Việc cập nhật và khai thác dữ liệu còn gặp nhiều khó khăn.

Việc trao đổi thông tin dữ liệu bị hạn chế do đó dẫn đến việc lãng phí tài nguyên dữ liệu.

Để đưa CSDL phục vụ cho phát triển kinh tế một cách hiệu quả thì việc đánh giá các mô hình quản lý cơ sở dữ liệu biển để đưa ra các giải pháp quản lý và khai thác CSDL hiệu quả là rất cần thiết.

Vì vậy, đề án “Xây dựng hệ thống trao đổi thông tin dữ liệu biển và nối mạng ở Việt Nam”, trong khuôn khổ dự án hợp tác giữa các nước ASEAN về quản lý xung đột tiềm tàng trên Biển Đông cần xây dựng mới và ứng dụng các công nghệ tiên tiến để đáp ứng được yêu cầu kỹ thuật về lưu trữ, chuyển đổi dữ liệu và trao đổi thông tin trong nước và quốc tế.

HỆ THỐNG QUẢN LÝ CƠ SỞ DỮ LIỆU WEBGIS

Thành phần khai thác, vận hành hệ thống

Các thành phần tham gia khai thác hệ thống được tập trung đến các đối tượng có đề cập đến khuôn khổ hợp tác và mục đích nội địa, bao gồm các thành phần sau:

Đối tác trong khuôn khổ hợp tác trao đổi dữ liệu biển ASEAN, các tài khoản thuộc khuôn khổ hợp tác giữa các bên nói trên được cấp quyền khai thác dữ liệu biển Việt Nam theo nội dung, quyền hạn và mức độ chi tiết theo thỏa thuận giữa các bên.

Đơn vị, cá nhân trong nước được quyền khai thác dữ liệu theo phương thức kết nối mạng trực tiếp, sao chép nội dung theo nhu cầu và phạm vi quyền hạn của đơn vị và cá nhân phục vụ các nhiệm vụ an ninh, quốc phòng, kinh tế, xã hội và quản lý nhà nước.

Đơn vị quản lý là Viện Địa chất và Địa vật lý biển thuộc Viện Hàn lâm Khoa học và Công nghệ Việt Nam thực hiện quản lý toàn bộ hệ thống, chịu trách nhiệm trước Chủ tịch Viện Hàn lâm Khoa học và Công nghệ Việt Nam về nhiệm vụ của mình, tham mưu đề xuất nghiên cứu phát triển nâng cao hiệu quả của hệ thống.

Quản trị hệ thống thuộc Viện Địa chất và Địa vật lý biển duy trì vận hành hệ thống bảo đảm ổn định, bảo mật và cập nhật dữ liệu kết quả nghiên cứu liên quan.

Thành phần kỹ thuật của hệ thống

Hệ thống xây dựng được cấu thành từ nhiều thành phần kỹ thuật khác nhau để hình thành 3

khối chức năng lớn là: 1- Lưu trữ dữ liệu; 2- Khai thác dữ liệu và 3- Quản trị dữ liệu. Để các khối chức năng trên hoạt động cần có các thành phần phần cứng và phần mềm hoạt động đồng bộ, cụ thể như sau [7-9]:

Kho dữ liệu sơ cấp lưu trữ dữ liệu về biển bao gồm dữ liệu không gian (bản đồ, hình ảnh), dữ liệu phi không gian (dữ liệu quản lý không gian, mô tả không gian, dữ liệu ghi trực tiếp kết quả khảo sát, kết quả nghiên cứu, thông tư, văn bản liên quan...).

Kho dữ liệu thứ cấp lưu trữ dữ liệu đã được xử lý, phân mức khai thác phù hợp chuẩn kết nối, trao đổi đáp ứng yêu cầu hợp tác giữa các bên, yêu cầu bảo mật thông tin, khẳng định chủ quyền lãnh thổ của Việt Nam.

Hệ thống máy chủ gồm 2 hệ thống máy chủ tương ứng điều hành kho dữ liệu sơ cấp và kho dữ liệu thứ cấp), cấu hình hệ thống mạnh, đáp ứng nhiều người dùng khai thác đồng thời.

Thiết bị lưu trữ, sao lưu dữ liệu NAS Server lưu trữ kho dữ liệu lớn lên đến trên 100 TB và có khả năng mở rộng trong quá trình vận hành theo thời gian.

Thiết bị kết nối mạng bao gồm các thiết bị chuyển mạch (switch layer 3), cáp mạng và thiết bị kết nối đầu cuối.

Tường lửa - Firewall sử dụng thiết lập hàng rào kỹ thuật kiểm soát dữ liệu vào/ra và kết nối bên ngoài với hệ thống.

Hệ phần mềm quản lý bản đồ ArcGIS phiên bản for Server và phiên bản for Desktop sử dụng tương ứng cho lưu trữ, quản lý và xử lý tập trung và kết nối trên máy trạm. Dữ liệu được chuyển đổi từ các kho dữ liệu bản đồ khác nhau quản lý tập trung và phân tán, trong đó việc chuyển đổi tiêu chuẩn được thực hiện tự động kết hợp nhân công. Bên cạnh hệ quản lý bản đồ ArcGIS, hệ phần mềm quản lý PostGIS được sử dụng phối hợp trong các module lập trình WebGIS cho các mức chia sẻ dữ liệu khác nhau theo phân mức quyền hạn khai thác.

Hệ quản trị cơ sở dữ liệu SQL Server, MySQL có chức năng lưu trữ và quản lý kho dữ liệu phi không gian bao gồm dữ liệu quản trị, dữ liệu ghi trực tiếp, dữ liệu xử lý... Dữ liệu quản trị theo mô hình phức hợp giữa tập trung - phân tán, thực thể - liên kết và hướng đối tượng, đáp ứng yêu cầu kết nối các kho dữ liệu khác nhau, đa dạng về định dạng dữ liệu [7-9].

Hệ điều hành máy tính bao gồm hệ điều hành máy chủ Windows Server và hệ điều hành máy trạm Windows Pro được cấu hình domain quản trị, điều hành theo tiêu chuẩn bảo mật và phân phối tài nguyên.

Các module chức năng phần mềm là khối nội dung nghiên cứu phát triển đáp ứng yêu cầu đặc thù riêng của hệ thống, gồm các khối sau:

Front-end: Là phần tương tác với người dùng khai thác hệ thống, bao gồm:

Web portal, giao diện tương tác người dùng qua đó có thể khai thác, kết nối và tra cứu dữ liệu theo phân quyền tài khoản được cung cấp.

Webservice, cung cấp công cụ cho phép hệ thống phần mềm kết nối khai thác dữ liệu.

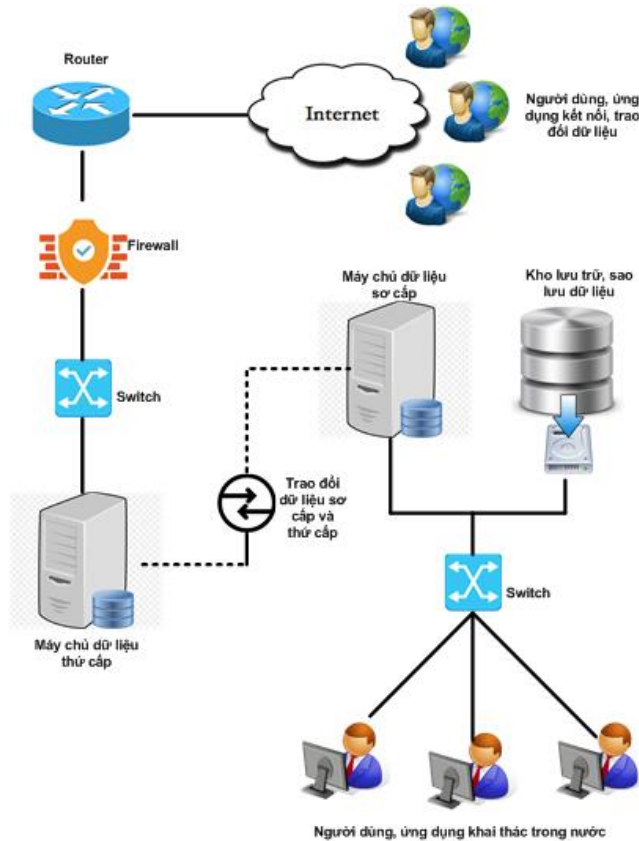
API function, cung cấp hàm kết nối đến hệ thống, tạo tính đa dạng trong kết nối phục vụ cho các hệ thống.

Các thành phần trên được cung cấp khai thác riêng biệt theo các kho dữ liệu sơ cấp và thứ cấp.

Back-end: Các hoạt động của phần mềm xử các yêu cầu, các hoạt động kết nối dữ liệu và các hoạt động liên quan đến phục vụ front - end người dùng, các chức năng quản trị, cập nhật dữ liệu hệ thống duy trì vận hành hệ thống đồng bộ và ổn định.

Sơ đồ kết nối các thành phần của hệ thống

Sơ đồ mô tả kết nối giữa các thành phần cơ bản tham gia khai thác hệ thống và các thành phần kỹ thuật [11–12]. Trong đó, dữ liệu thứ cấp được xử lý, chuyển đổi từ dữ liệu sơ cấp qua module phần mềm và các quản trị viên đảm bảo đáp ứng đúng chuẩn dữ liệu, chuẩn kết nối và đúng nội dung, đúng thông tin và đúng tài khoản tham gia chia sẻ. Hệ thống thứ cấp được offline với hệ thống dữ liệu sơ cấp, kết nối thông qua sao chép thủ công hoặc module kết nối có kiểm soát dữ liệu bảo đảm an toàn tuyệt đối việc xâm nhập từ hệ thống thứ cấp sang hệ thống sơ cấp của giao dịch không mong muốn và không được chào đón.



Hình 1. Sơ đồ kết nối các thành phần trao đổi thông tin dữ liệu biên

Người dùng hoặc ứng dụng các bên tham gia hợp tác ASEAN truy xuất dữ liệu thuộc kho dữ liệu thứ cấp được kiểm soát nội dung, thông tin dữ liệu qua đường internet. Các tài khoản khai thác sử dụng khai thác trực tiếp trên web portal hoặc kết nối thông qua Web Service, hàm API.

Đối với người dùng trong nước thuộc cơ quan, đơn vị, cá nhân được phép khai thác phục vụ nhiệm vụ an ninh, quốc phòng, kinh tế, xã hội và quản lý nhà nước kết nối khai thác thông qua đường kết nối nội bộ hoặc mạng riêng ảo (VPN), nội dung được chia sẻ, kiểm soát theo quy định nhiệm vụ và hợp tác thực hiện.

Đối với người dùng, tài khoản thuộc Viện Hàn lâm Khoa học và Công nghệ Việt Nam được kết nối nội bộ thông qua đường kết nối

nội bộ, kết nối đường mạng riêng ảo (VPN) theo quy định.

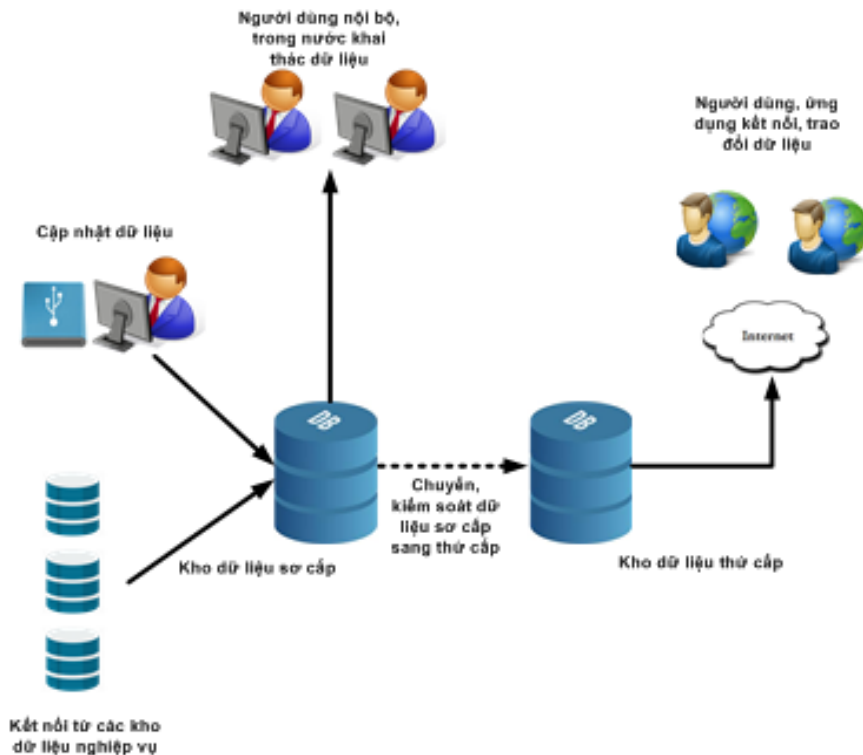
Sơ đồ khái quát kết nối luồng dữ liệu

1- Dữ liệu được cập nhật bởi cán bộ nghiệp vụ Trung tâm dữ liệu biển-Viện Địa chất và Địa vật lý biển hoặc kết nối từ các kho dữ liệu nghiệp vụ.

2- Dữ liệu được phân quyền khai thác cho cán bộ nội bộ, cơ quan đơn vị trong nước theo yêu cầu và nhiệm vụ cụ thể phục vụ các nhiệm vụ quốc gia.

3- Dữ liệu được xử lý, chọn lọc và chuyển từ kho dữ liệu sơ cấp sang kho dữ liệu thứ cấp.

4- Dữ liệu được khai thác thông qua người dùng khai thác trực tiếp trên portal (front-end) hoặc kết nối thông qua Webservice, API function.



Hình 2. Sơ đồ khái quát luồng dữ liệu

KỸ THUẬT, CÔNG NGHỆ XÂY DỰNG HỆ THỐNG VÀ CƠ SỞ DỮ LIỆU WEBGIS

Hệ thống chia sẻ dữ liệu biển, trong đó dữ liệu không gian và dữ liệu phi không gian được đồng bộ, lưu trữ và xử lý trên các lớp [11–14]:

Lớp Database Server: Quản trị kho dữ liệu không gian, phi không gian.

ArcGIS Server, PostGIS: Quản lý dữ liệu bản đồ.

SQL Server, MySQL: Quản lý dữ liệu phi không gian, dữ liệu mô tả, dữ liệu quản lý.

Lớp Service Server: Vận hành WebService, API và các dịch vụ xử lý dữ liệu.

Lớp UI: Cung cấp giao diện người dùng (back-end, front-end), hàm giao tiếp kết nối dữ liệu với hệ thống.

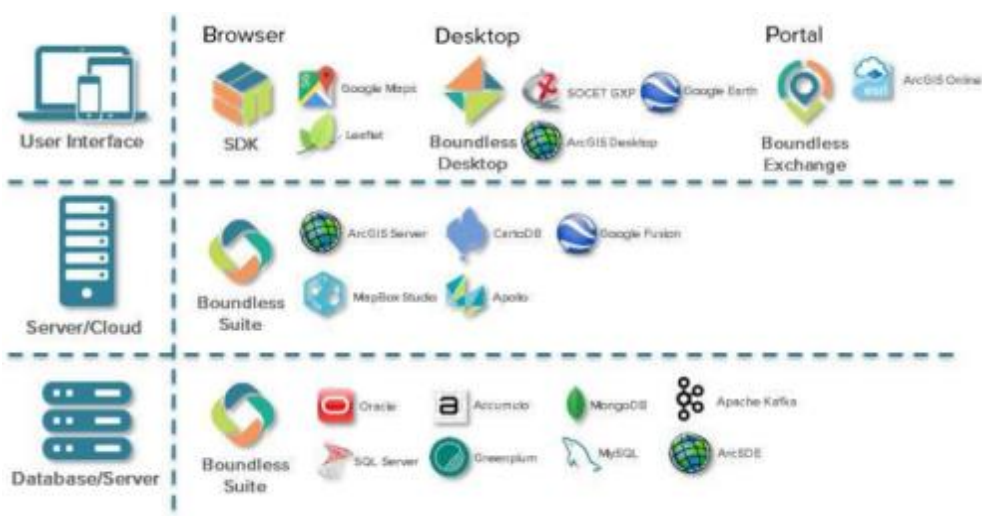
Các lớp được thiết lập hệ thống bảo mật trong (Internal FireWall - Cài đặt và lập trình trong nội tại hệ thống máy chủ lưu trữ dữ liệu), bảo mật ngoài (External FireWall) và các công

nghệ bảo mật chống tấn công, xâm nhập từ bên ngoài khác.

Nền tảng công nghệ WebGIS (hình 3) lựa chọn xây dựng hệ thống chia sẻ dữ liệu dựa trên mô hình ba lớp (hình 4), trong đó Portal, WebService và API function cung cấp khả năng kết nối dữ liệu bản đồ và dữ liệu phi không gian cho người dùng.



Hình 3. Mô hình dịch vụ WebGIS chia sẻ dữ liệu



Hình 4. Mô hình ba lớp hệ thống trao đổi dữ liệu

WebGIS xây dựng trên nền tảng quản trị dữ liệu GIS của ArcGIS Server, qua đó các ứng

dụng được phát triển dựa trên nền WebGIS khai thác dữ liệu GIS và dữ liệu phi không gian khác.



Hình 5. Nền tảng ArcGIS Server và WebGIS trong hệ thống chia sẻ dữ liệu

Để bảo đảm an toàn, bảo mật trong khai thác, chia sẻ dữ liệu của hệ thống, nhiều giải pháp được đề xuất, giải pháp mô hình bảo mật hai lớp Firewall (External firewall và internal firewall) được nhà phát triển ArcGIS đưa ra, trong đó gồm:

Lớp Internal firewall: Kiểm soát truy xuất giữa WebServer và GIS Server, DB Server, bảo đảm an toàn phiên truy vấn dữ liệu tin cậy. Lớp Internal Firewall là lớp firewall mềm kiểm cài đặt trên máy chủ kiểm soát dữ liệu và các kết nối. Phát hiện tấn công, xâm nhập bất thường đến hệ thống.

Lớp External firewall: Kiểm soát truy cập từ bên ngoài, đảm bảo an toàn dữ liệu chia sẻ qua cổng kết nối, ngăn chặn các luồng truy cập không mong muốn.

Chống tấn công từ chối dịch vụ, SQL Injection và kỹ thuật khác nhằm bảo đảm hệ thống vận hành an toàn, không bị xâm nhập trái phép.

DỮ LIỆU LƯU TRỮ, XỬ LÝ VÀ TRAO ĐỔI

Kho dữ liệu biển thiết kế là mô hình phức hợp, gồm các kho thành phần sau

Kho dữ liệu quản trị: Lưu trữ dữ liệu về quản trị hệ thống.

Kho dữ liệu quản lý dữ liệu biển phi không gian: Lưu trữ dữ liệu quản lý đối tượng, lưu trữ phục vụ các chức năng nghiệp vụ, dữ liệu kho này không chứa dữ liệu bản đồ.

Kho dữ liệu không gian: Lưu trữ dữ liệu bản đồ, hình ảnh.

Thiết kế cấu trúc kho dữ liệu: Các kho dữ liệu được thiết kế bảo đảm các yêu cầu kho dữ liệu lớn về: Dung lượng, tính hợp lệ, tính đa dạng, độ tin cậy và tốc độ xử lý. Ngoài ra, các yêu cầu về cơ sở dữ liệu quan hệ thực thể - liên kết, cơ sở dữ liệu hướng đối tượng, cơ sở dữ liệu GIS, cơ sở dữ liệu phân tán.

Định dạng dữ liệu: Dữ liệu được lưu trữ theo nhiều định dạng khác nhau: Chữ, số, thời gian, đối tượng, nhị phân, html, xml, ảnh, video, map, metadata. Mỗi định dạng dữ liệu có đặc điểm lưu trữ và xử lý khác nhau.

Dữ liệu được tổ chức lưu trữ, cấu trúc theo công nghệ dữ liệu lớn (BigData) đáp ứng các đặc tính: Dung lượng lớn, tốc độ nhanh, đa dạng, tính chính xác, giá trị.

Các số liệu được phân làm 4 mức bảo mật cung cấp dữ liệu theo các đối tượng và nhu cầu khai thác như sau:

Mức bảo mật 1: Dữ liệu công khai.

Mức bảo mật 2: Dữ liệu phục vụ chia sẻ và trao đổi với các nước ASEAN.

Mức bảo mật 3: Dữ liệu phục vụ trao đổi, chia sẻ với các đơn vị, cá nhân trong nước theo chương trình phối hợp giữa các cơ quan.

Mức bảo mật 4: Dữ liệu nội bộ.

Công tác biên tập cơ sở dữ liệu số chuyên đề GIS

Công việc biên tập cơ sở dữ liệu số chuyên đề liên quan bao gồm các bước như sau:

Chuẩn bị tài liệu để tiến hành số hóa biên tập cần tiến hành rà soát và kiểm tra nguồn dữ liệu bản đồ đang được lưu trữ ở dưới dạng giấy và cơ sở dữ liệu ảnh.

Sử dụng các phần mềm số hóa bản đồ như NeuraMap, ArcGIS... để số hóa các bản đồ lựa chọn làm bản đồ nền.

Dựa trên các thông tin thu thập được từ các báo cáo nghiên cứu khoa học, sử dụng phần mềm ArcGIS để biên tập bản đồ vectơ.

Biên tập sửa chữa lỗi kỹ thuật tin học cho toàn bộ các bản vẽ.

Biên tập đưa về các lớp thông tin chuẩn, xây dựng cấu trúc các lớp thông tin, xây dựng các thông tin thuộc tính.

Xây dựng và nhập các dữ liệu thuộc tính.

Kiểm tra mức độ đầy đủ thông tin của các trường thuộc tính, đặc biệt đối với các trường thuộc tính bắt buộc phải nhập: Mã nhận dạng, mã đối tượng, loại đối tượng.

Kiểm tra độ chính xác thông tin của các trường thuộc tính so với thông tin từ các nguồn

dữ liệu, thuyết minh mô hình dữ liệu và quy định gán mã nhận dạng, quy định nhập thông tin thuộc tính. Sản phẩm của từng bước thực hiện trong quá trình xây dựng cơ sở dữ liệu phải được kiểm tra, đánh giá và sửa chữa triệt để, kiểm tra chất lượng. Các lỗi phát hiện trong quá trình kiểm tra, nghiệm thu các cấp phải được sửa chữa triệt để. Chất lượng sản phẩm phải đáp ứng theo yêu cầu kỹ thuật trong các tài liệu sử dụng để thi công và quy định hiện hành.

Các chức năng cơ bản của phần mềm quản lý dữ liệu VNODC

Khởi chức năng cập nhật dữ liệu bản đồ

Import dữ liệu bản đồ.

Export dữ liệu bản đồ.

Cập nhật dữ liệu bản đồ.

Kiểm tra sự chồng khớp các vùng bản đồ.

Khởi chức năng khai thác dữ liệu bản đồ

Xem bản đồ.

Lựa chọn vùng bản đồ.

Tác nghiệp với bản đồ.

Tìm kiếm trên bản đồ.

Khai thác thông tin bản đồ.

Tổng hợp dữ liệu về bản đồ.

Khởi chức năng kết xuất dữ liệu bản đồ

Nhập vùng bản đồ.

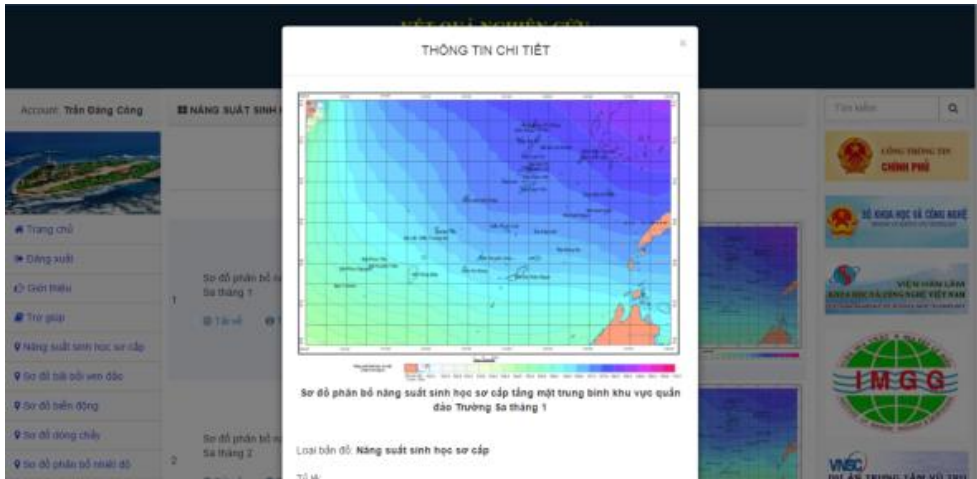
Chọn, cắt vùng bản đồ.

Nối, ghép vùng bản đồ.

Xuất vùng bản đồ ra các định dạng khác nhau.



Hình 6. Giao diện khai thác dữ liệu bản đồ



Hình 7. Giao diện thông tin chi tiết bản đồ

Khởi chức năng quản lý dữ liệu biển đảo (dữ liệu phi không gian)

Cập nhật dữ liệu các công trình nghiên cứu.

Cập nhật dữ liệu quảng bá chủ quyền biển đảo.

Tra cứu tìm kiếm.

Tổng hợp dữ liệu.

Khởi chức năng phân tích dữ liệu

Phân tích sự thay đổi dữ liệu bản đồ biển đảo.

Phân tích sự thay đổi dữ liệu liên quan đến khu vực biển đảo.

Khởi chức năng kết nối dữ liệu

Giao tiếp dữ liệu qua Webservice.

Giao tiếp (API) dữ liệu bản đồ với hệ thống khác.

Giao tiếp (API) dữ liệu phi bản đồ với các hệ thống khác.

Khởi chức năng an toàn, bảo mật dữ liệu

Phần mềm được triển khai trên mạng internet có địa chỉ <http://www.vnodc.net.vn>. Ngoài ra, phần mềm có thể triển khai trên mạng nội bộ hoặc mạng diện rộng của cơ quan, đơn vị có điều kiện triển khai. Các giao diện và thao tác chính của phần mềm có thể thấy qua một số hình ảnh giới thiệu dưới đây.

Tìm kiếm bản đồ

Nhập thông tin bản đồ cần tìm vào hộp tìm kiếm để thực hiện tìm.



Hình 8. Giao diện tìm kiếm, tra cứu bản đồ

VẬN HÀNH VÀ KHAI THÁC HỆ THỐNG

Vận hành, Trung tâm dữ liệu biển Quốc gia, Viện Địa chất và Địa vật lý biển vận hành theo quy chế hoạt động được cấp có thẩm quyền ban hành, bảo đảm kỹ thuật cho hệ thống vận hành hiệu quả, ổn định. Các chuyên gia thuộc phòng nghiệp vụ cung cấp dữ liệu chuyên ngành chia cập nhật kho dữ liệu.

Khai thác, các bên tham gia theo chương trình hợp tác được cung cấp tài khoản xác định quyền sử dụng, khai thác dữ liệu theo đúng nội dung hợp tác. Các đơn vị, cá nhân trong nước, thuộc Viện Hàn lâm Khoa học và Công nghệ Việt Nam được cung cấp tài khoản khai thác kho dữ liệu sơ cấp theo nội dung phối hợp thuộc chương trình, nhiệm vụ cụ thể.

Nghiên cứu và phát triển, cơ quan chủ quan chủ động, phối hợp các đơn vị liên quan tham mưu, đề xuất chương trình nghiên cứu và phát triển mạnh hơn nữa hiệu quả hoạt động của hệ thống.

KẾT LUẬN

Hệ thống trao đổi thông tin dữ liệu biển được xây dựng đóng vai trò quan trọng nhằm kiểm chế xung đột trên Biển Đông, đồng thời phục vụ nhiệm vụ an ninh, quốc phòng, kinh tế xã hội và nghiên cứu trong nước.

Sự đa dạng dữ liệu về biển, dữ liệu được kết nối, tổ chức khai thác theo nhiều mức khác nhau phục vụ thiết thực từng nhóm đối tượng người dùng phát huy hiệu quả kết quả nghiên cứu của các công trình nghiên cứu trước đó, cung như nghiên cứu cập nhật mới đáp ứng tốt hơn hiệu quả khai thác sử dụng.

Hệ thống, nòng cốt là trung tâm dữ liệu về biển Việt Nam, được thiết kế, xây dựng hiện đại, bằng công nghệ mới, phù hợp với xu thế phát triển của công nghệ, bắt kịp cuộc cách mạng công nghệ 4.0, đúng theo tinh thần và sự chỉ đạo từ Trung ương đến các cấp ngành nước ta.

Lời cảm ơn: Bài báo đã được hoàn thành dưới sự trợ giúp của đề tài thuộc Chương trình Khoa học và Công nghệ cấp Quốc gia về Công nghệ vũ trụ 2016–2020, mã số đề tài: VT-UD.04/17–20.

TÀI LIỆU THAM KHẢO

- [1] Nguyễn Đình Kỳ, 2016. Xây dựng CSDL GIS sở dữ liệu và Atlas điện tử tổng hợp vùng Tây Nguyên.
- [2] World Data Center, Available at: https://en.wik-ipedia.org/wiki/World_Data_Center. (access 20 June 2019).
- [3] World Ocean Database, Available at: <http://wod.iode.org> (access 22 June 2019).
- [4] National Center Invironmental Information, Available at: <https://www.nodc.noaa.gov/> (access 25 July 2019).
- [5] Khan, M., Wu, X., Xu, X., and Dou, W., 2017. Big data challenges and opportunities in the hype of Industry 4.0. In *2017 IEEE International Conference on Communications (ICC)* (pp. 1–6). *IEEE*.
- [6] Vũ Văn Tác, Ngô Mạnh Tiến, 2012. Dữ liệu hải dương học Biển Đông trong cơ sở dữ liệu biển thế giới 2009. *Tuyển tập nghiên cứu biển*, Tập XVIII. Tr. 35–45.
- [7] Aji, A., Sun, X., Vo, H., Liu, Q., Lee, R., Zhang, X., ... and Wang, F., 2013. Demonstration of Hadoop-GIS: a spatial data warehousing system over MapReduce. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 528–531). *ACM*.
- [8] Eldawy, A., and Mokbel, M. F., 2015. Spatialhadoop: A mapreduce framework for spatial data. In *2015 IEEE 31st international conference on Data Engineering* (p. 1352–1363). *IEEE*.
- [9] Yang, C., Yu, M., Hu, F., Jiang, Y., and Li, Y., 2017. Utilizing cloud computing to address big geospatial data challenges. *Computers, Environment and Urban Systems*, 61, 120–128.
- [10] Hershey, P. C., and Silio, C. B., 2012. Procedure for detection of and response to distributed denial of service cyber attacks on complex enterprise systems. In *2012 IEEE International Systems Conference SysCon 2012* (pp. 1–6). *IEEE*.
- [11] High Efficiency, Firewalls and ArcGIS Server, Available at: <http://enterprisearcgis.com/en/>

- server/latest/administer/linux/firewall-and-arcgis-server.html. (access 20 July 2019).
- [12] Jin, S., Wang, X., Luan, C., Zhang, H., and Guo, Y., 2012. Construction of Marine Oil Spill Response Information System Based on ArcGIS Server. In *Recent Advances in Computer Science and Information Engineering* (pp. 575–581). Springer, Berlin, Heidelberg.
- [13] Yang, C., Huang, Q., Li, Z., Liu, K., and Hu, F., 2017. Big Data and cloud computing: innovation opportunities and challenges. *International Journal of Digital Earth*, 10(1), 13–53.
- [14] Grecea, C., Herban, S., and Vilceanu, C. B., 2016. WebGIS solution for urban planning strategies. *Procedia engineering*, 161, 1625–1630.
- [15] Wang, F., Aji, A., Liu, Q., and Saltz, J., 2011. Hadoop-GIS: A high performance spatial query system for analytical medical imaging with MapReduce. *Center for Comprehensive Informatics, Technical Report*. Available at: <http://www3.cs.stonybrook.edu/~fuswang/papers/CCI-TR-2011-3.pdf> (access 21 September 2015).
- [16] Fustes, D., Cantorna, D., Dafonte, C., Iglesias, A., and Arcay, B., 2012. Applications of cloud computing and gis for ocean monitoring through remote sensing. In *Smart Sensing Technology for Agriculture and Environmental Monitoring* (pp. 303–321). Springer, Berlin, Heidelberg.
- [17] Vo, H., Aji, A., and Wang, F., 2014. SATO: a spatial data partitioning framework for scalable query processing. In *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 545–548). ACM.