

BỘ TÀI NGUYÊN VÀ MÔI TRƯỜNG  
CỤC CHUYÊN ĐỔI SỐ VÀ THÔNG TIN DỮ LIỆU  
TÀI NGUYÊN MÔI TRƯỜNG

-----o0o-----

Quyển III

**THIẾT KẾ CHI TIẾT**

HẠNG MỤC: **XÂY DỰNG CƠ SỞ DỮ LIỆU ...**

**THUỘC DỰ ÁN:**

**XÂY DỰNG, HOÀN THIỆN HỆ THỐNG THÔNG TIN, CƠ  
SỞ  
DỮ LIỆU TÀI NGUYÊN VÀ MÔI TRƯỜNG (GIAI ĐOẠN  
I)**

## CÁC THUẬT NGỮ, TỪ VIẾT TẮT

TT	Thuật ngữ/ Viết tắt	Ý nghĩa
1	CNTT	Công nghệ thông tin
2	HTTT	Hệ thống thông tin
3	CSDL	Cơ sở dữ liệu
4	LIS	Hệ thống thông tin đất đai
5	TN&MT	Tài nguyên và Môi trường
6	CQNN	Cơ quan nhà nước
7	CPĐT	Chính phủ điện tử
8	ADM	Architecture Development Method Phương pháp phát triển kiến trúc
9	BPM	Business Process Management Quản lý quy trình nghiệp vụ
10	BPR	Business Process Reengineering Tái cấu trúc quy trình nghiệp vụ
11	COTS	Commercial Off-The-Shelf Hệ thống thương mại mở
12	DRM	Data Reference Model Mô hình tham chiếu dữ liệu
13	EA	Enterprise Architecture Kiến trúc tổng thể
14	ESB	Enterprise Service Bus Trục tích hợp
15	G2B	Government to Business Services Dịch vụ và quan hệ giữa Chính phủ và doanh nghiệp
16	G2C	Government to Citizen Services Dịch vụ và quan hệ giữa Chính phủ và người dân
17	G2E	Government to Employee Services Dịch vụ và quan hệ giữa Chính phủ và cán bộ, công chức viên chức
18	G2G	Government to Government Services Dịch vụ và quan hệ giữa Chính phủ với Chính phủ

19	ISO	International Standards Organization
		Tổ chức các tiêu chuẩn quốc tế
20	IT	Information Technology
		Công nghệ thông tin
21	ITIL	The Information Technology Infrastructure Library
		Thư viện cơ sở hạ tầng công nghệ thông tin
22	KPI	Key Performance Indicator
		Chỉ số đo lường kết quả hoạt động
23	LAN	Local Area Network
		Mạng cục bộ
24	SOA	Service-Oriented Architecture
		Kiến trúc hướng dịch vụ
25	TOGAF	The Open Group Architecture Framework
		Khung kiến trúc nhóm mở
26	TRM	Technical Reference Model
		Mô hình tham chiếu công nghệ
27	WAN	Wide Area Network
		Mạng diện rộng
28	XML	Extensible Markup Language
		Ngôn ngữ đánh dấu mở rộng
29	NGSP	National Government Service Platform
		Hệ thống kết nối, liên thông các hệ thống thông tin ở Trung ương và địa phương
30	LGSP	Local Service Platform
		Nền tảng tích hợp, chia sẻ dữ liệu
31	ICT	Information and Communication Technologies: Công nghệ thông tin và Truyền thông

# Contents

<b>CÁC THUẬT NGỮ, TỪ VIẾT TẮT .....</b>	<b>2</b>
<b>I. Yêu cầu các tiêu chuẩn kỹ thuật .....</b>	<b>6</b>
<b>I.1. Yêu cầu chung đối với hệ thống .....</b>	<b>6</b>
<b>I.2. Yêu cầu về cơ sở toán học với dữ liệu không gian .....</b>	<b>7</b>
<b>I.3. Yêu cầu kỹ thuật.....</b>	<b>7</b>
<b>I.4. Yêu cầu an ninh bảo mật .....</b>	<b>8</b>
<b>II. THIẾT KẾ CHI TIẾT .....</b>	<b>9</b>
<b>II.1. Xây dựng danh mục dữ liệu chủ và thiết kế cấu trúc dữ liệu chi tiết cho cơ sở dữ liệu khí tượng thủy văn quốc gia .....</b>	<b>9</b>
<b>II.1.1. Tổng quan nguồn dữ liệu .....</b>	<b>9</b>
<b>II.1.2. Xây dựng danh mục dữ liệu chủ cho cơ sở dữ liệu khí tượng thủy văn quốc gia .....</b>	<b>9</b>
<b>II.1.3. Thiết kế cấu trúc dữ liệu chi tiết cho cơ sở dữ liệu khí tượng thủy văn quốc gia .....</b>	<b>13</b>
<b>II.1.4. Dịch vụ chia sẻ dữ liệu KTTV Quốc gia .....</b>	<b>59</b>
<b>II.1.5. Sản phẩm giao nộp .....</b>	<b>71</b>
<b>II.1.6. Tổ chức thực hiện .....</b>	<b>72</b>
<b>II.2. Xây dựng, triển khai mô hình cơ sở dữ liệu ứng dụng công nghệ Big Data cho Trung tâm dữ liệu tại Tổng cục KTTV.....</b>	<b>72</b>
<b>II.2.1. Các chỉ tiêu kỹ thuật áp dụng trong triển khai lắp đặt, cài đặt, kiểm tra và hiệu chỉnh thiết bị.....</b>	<b>72</b>
<b>II.2.2. Danh mục thiết bị lắp đặt, cài đặt và các thông số kỹ thuật của thiết bị.....</b>	<b>83</b>
<b>II.2.3. Thống kê khối lượng các thiết bị, dịch vụ .....</b>	<b>91</b>
<b>II.2.4. Chỉ dẫn biện pháp triển khai.....</b>	<b>102</b>
<b>II.2.5. Sơ đồ và thuyết minh tổng thể hệ thống .....</b>	<b>107</b>

<b>II.2.6. Yêu cầu chi tiết về việc đào tạo hướng dẫn sử dụng; triển khai, hỗ trợ, quản trị, vận hành sản phẩm .....</b>	<b>140</b>
<b>II.2.7. Kế hoạch triển khai .....</b>	<b>144</b>

## **I. Yêu cầu các tiêu chuẩn kỹ thuật**

### **I.1. Yêu cầu chung đối với hệ thống**

Danh mục dữ liệu chủ và cơ sở dữ liệu phải đáp ứng được nhu cầu quản lý và lưu trữ các thông tin, dữ liệu thuộc lĩnh vực khí tượng thủy văn.

Danh mục dữ liệu chủ và cơ sở dữ liệu đáp ứng được khả năng chia sẻ dữ liệu cho các hệ thống khác khi có nhu cầu.

Cơ sở dữ liệu đáp ứng được khả năng lưu trữ dữ liệu lớn với nhiều loại dữ liệu khác nhau.

Hệ thống phải đáp ứng được nhu cầu quản lý và khai thác các thông tin, dữ liệu thuộc các lĩnh vực phục vụ trong công tác quản lý thông tin của các cấp/ngành trên địa bàn tỉnh.

Hệ thống đáp ứng được khả năng chia sẻ dữ liệu giữa các đơn vị với nhau, tận dụng, kế thừa và sử dụng lại các nguồn dữ liệu đã được xây dựng trên hệ thống.

Hệ thống đáp ứng được khả năng lưu trữ dữ liệu lớn với nhiều loại dữ liệu khác nhau.

Hệ thống cung cấp khả năng truy cập mọi lúc, mọi nơi trên trình duyệt Web, thiết bị di động.

Hệ thống cung cấp cơ chế phân quyền ứng dụng – người dùng đảm bảo khả năng bảo mật thông tin trên hệ thống.

## **I.2. Yêu cầu về cơ sở toán học với dữ liệu không gian**

Dữ liệu không gian thuộc hệ thống cơ sở dữ liệu tài nguyên và môi trường tỉnh Gia Lai phải được tham chiếu đến hệ quy chiếu quốc gia và hệ tọa độ quốc gia VN2000 với nguyên tắc sau:

- Elipsoid WGS-84 toàn cầu được định vị phù hợp với lãnh thổ của Việt Nam;

- Kinh tuyến trục và múi chiếu của từng loại bản đồ của tỉnh Gia Lai tuân thủ đúng theo quy định hiện hành do Cục Đo đạc bản đồ Việt Nam quy định.

Các văn bản cụ thể cần phải áp dụng gồm:

- Căn cứ Quyết định số 83/2000/QĐ-TTg ngày 12-7-2000 của Thủ tướng Chính phủ về việc áp dụng Hệ quy chiếu và Hệ tọa độ quốc gia VN-2000.

- Thông tư 973/2001/TT-TCĐC ngày 20 tháng 06 năm 2001 của Tổng cục Địa chính (nay là Bộ TN&MT) về việc hướng dẫn áp dụng hệ quy chiếu và hệ tọa độ quốc gia VN-2000.

- Quyết định số 05/QĐ-BTNMT ngày 27 tháng 02 năm 2007 của Bộ Tài nguyên và Môi trường về việc sử dụng hệ thống tham số tính chuyển giữa hệ tọa độ quốc tế WGS84 với hệ tọa độ quốc gia VN2000.

## **I.3. Yêu cầu kỹ thuật**

Việc xây dựng danh mục dữ liệu chủ và thiết kế cấu trúc dữ liệu chi tiết cho CSDL KTTV quốc gia phải đáp ứng các yêu cầu kỹ thuật như sau:

### **Yêu cầu chung:**

Xây dựng, hoàn thiện danh mục dữ liệu chủ và thiết kế cấu trúc dữ liệu chi tiết cho CSDL KTTV quốc gia được tổ chức theo mô hình dữ liệu phi không gian hướng đối tượng.

Thiết kế dữ liệu có thể áp dụng mô hình tập trung hoặc mô hình phân tán.

Hệ thống cơ sở dữ liệu được thiết kế theo tiêu chí chạy trên mạng LAN, WAN, Internet để phục vụ các yêu cầu về nghiệp vụ của cán bộ tài nguyên môi trường; các dịch vụ chia sẻ (API/Web Service), cung cấp thông tin, dữ liệu được xây dựng trên nền tảng công nghệ Internet để có thể hoạt động được trên mạng WAN, LAN, INTERNET.

Dữ liệu phải được chia sẻ thông tin thuận tiện, hiệu quả với các cơ sở dữ liệu khác.

Sử dụng mã dựng sẵn trong bộ mã chuẩn tiếng Việt TCVN 6909:2001.

Hệ thống được thiết kế, xây dựng và triển khai phải tuân thủ các tiêu chuẩn kỹ thuật về kết nối hệ thống, tích hợp dữ liệu, đặc tả dữ liệu và truy xuất thông tin quy định tại Thông tư số 39/2017/TT-BTTTT của Bộ Thông tin và Truyền thông: Ban hành Danh mục tiêu chuẩn kỹ thuật về ứng dụng công nghệ thông tin trong cơ quan nhà nước.

#### **Yêu cầu về cài đặt, triển khai:**

Danh mục dữ liệu chủ của CSDL KTTV quốc gia được thiết kế, xây dựng và triển khai trên hạ tầng kỹ thuật và nền tảng Big Data của Tổng Cục khí tượng thủy văn.

Sẵn sàng cung cấp các dịch vụ chia sẻ (API/Web Service) dữ liệu cho các hệ thống khác khi có nhu cầu.

#### **I.4. Yêu cầu an ninh bảo mật**

Hệ thống phải đáp ứng khả năng an toàn, bảo mật theo nhiều mức: mức mạng, mức xác thực người sử dụng, mức CSDL.

Hỗ trợ người sử dụng trao đổi thông tin, dữ liệu với trên mạng Internet theo các chuẩn về an toàn thông tin như S/MIME v3.0, SSL v3.0, HTTPS, ...

Có cơ chế theo dõi và giám sát, lưu vết tất cả các hoạt động cho mỗi kênh thông tin và toàn hệ thống.

Toàn bộ các dữ liệu cần quản lý, phải được lưu trong CSDL được mã hóa và phân quyền truy cập chặt chẽ.

Phải có cơ chế sao lưu dữ liệu định kỳ, đột xuất đảm bảo nhanh chóng đưa hệ thống hoạt động trở lại trong trường hợp có sự cố xảy ra.

Hệ thống luôn phải đảm bảo tính sẵn sàng, phục vụ trực tuyến 24/7 và lưu trữ nhiều dữ liệu quan trọng, nhạy cảm về môi trường ...

Do vậy việc đảm bảo an toàn, an ninh cho hệ thống phần mềm là rất quan trọng, cần được quan tâm từ khâu thiết kế, bước lập dự án để đảm bảo sau khi đưa vào triển khai thì hệ thống đã được đảm bảo về an toàn thông tin toàn diện, cụ thể ở các vấn đề như chính sách an toàn thông tin hệ thống; mã hóa và an toàn dữ liệu; hệ thống và phần mềm, mã nguồn, các kết nối ...

Bên cạnh đó, theo quy định của Nghị định 85/2016/NĐ-CP bảo đảm an toàn hệ thống thông tin theo cấp độ và Thông tư 12/2022/TT-BTTTT hướng dẫn Nghị định 85/2016/NĐ-CP an toàn hệ thống thông tin đã quy định chi tiết về việc trách nhiệm đảm bảo an toàn hệ thống thông tin. Theo đó, cụ thể tại Điều 19 của

Nghị định 85/2016/NĐ-CP các hệ thống cần được đảm bảo an toàn thông tin từ khâu thiết kế, xây dựng và trong quá trình vận hành.

## **II. THIẾT KẾ CHI TIẾT**

### **II.1. Xây dựng danh mục dữ liệu chủ và thiết kế cấu trúc dữ liệu chi tiết cho cơ sở dữ liệu khí tượng thủy văn quốc gia**

#### ***II.1.1. Tổng quan nguồn dữ liệu***

Khí tượng thủy văn (KTTV) là một trong những lĩnh vực có nhu cầu khai thác dữ liệu rất lớn, trong thời gian vừa qua Tổng cục Khí tượng Thủy văn đã và đang sử dụng một khối lượng lớn dữ liệu phục vụ công tác chuyên môn và cập nhật định kỳ như:

- Dữ liệu quan trắc từ các trạm quan trắc thủ công, tự động;
- Dữ liệu chia sẻ qua kênh GTS;
- Dữ liệu vệ tinh;
- Dữ liệu Radar;
- Dữ liệu định vị sét;
- Sản phẩm từ các mô hình số trị, mô hình thủy văn/thủy lực và một số loại dữ liệu khác có liên quan.

#### **Các loại dữ liệu này đã được thu thập từ các nguồn dữ liệu sau:**

- Dữ liệu khí tượng thủy văn quốc gia: Được thu thập từ các trạm quan trắc KTTV do ngành KTTV đầu tư quản lý và vận hành;
- Dữ liệu khí tượng thủy văn chuyển dùng: Được thu thập từ các trạm quan trắc KTVVT do Bộ ban ngành địa phương quản lý vận hành và giao nộp dữ liệu về ngành KTTV;
- Dữ liệu khí tượng thủy văn được cung cấp bởi các tổ chức, cá nhân nước ngoài.

#### ***II.1.2. Xây dựng danh mục dữ liệu chủ cho cơ sở dữ liệu khí tượng thủy văn quốc gia***

##### ***II.1.2.1. Căn cứ pháp lý, cơ sở xác định danh mục dữ liệu chủ***

- Luật khí tượng thủy văn 2015;
- Nghị định số 38/2016/NĐ-CP ngày 15 tháng 05 năm 2016 về việc quy định chi tiết một số điều của luật khí tượng thủy văn;

- Nghị định số 48/2020/NĐ-CP ngày 15 tháng 4 năm 2020 về việc sửa đổi, bổ sung một số điều của nghị định số 38/2016/nđ-cp ngày 15 tháng 5 năm 2016 của chính phủ quy định chi tiết một số điều của luật khí tượng thủy văn;

- Nghị định số 73/2017/NĐ-CP ngày 14 tháng 06 năm 2017 về việc thu thập, quản lý, khai thác và sử dụng thông tin, dữ liệu tài nguyên và môi trường;

- Thông tư 26/2014/TT-BTNMT ngày 28 tháng 05 năm 2014 của Bộ Tài nguyên và Môi trường về việc ban hành Quy trình và Định mức kinh tế - kỹ thuật xây dựng cơ sở dữ liệu tài nguyên và môi trường;

- Thông tư số 07/2016/TT-BTNMT ngày 16 tháng 05 năm 2016 của Bộ Tài nguyên và Môi trường về việc quy định các bộ dữ liệu, chuẩn dữ liệu và xây dựng, quản lý cơ sở dữ liệu khí tượng thủy văn quốc gia;

- Thông tư số 40/2017/TT-BTNMT ngày 23 tháng 10 năm 2017 của Bộ Tài nguyên và Môi trường về việc quy định kỹ thuật chuẩn bộ dữ liệu về quan trắc, điều tra, khảo sát khí tượng, thủy văn, hải văn, môi trường không khí;

- Thông tư số 30/2018/TT-BTNMT ngày 26 tháng 12 năm 2018 của Bộ tài nguyên và Môi trường về việc quy định kỹ thuật về quan trắc và cung cấp thông tin, dữ liệu khí tượng thủy văn đối với trạm khí tượng thủy văn chuyên dùng;

- Thông tư số 18/2022/TT-BTNMT ngày 21 tháng 11 năm 2022 của Bộ Tài nguyên và Môi trường, quy định nội dung quan trắc khí tượng thủy văn đối với trạm thuộc mạng lưới trạm khí tượng thủy văn quốc gia.

#### *II.1.2.2. Quy trình, phương pháp thực hiện xác định danh mục dữ liệu chủ*

**Bước 1:** Xác định các danh mục dữ liệu chủ thông qua văn bản quy phạm pháp luật do Chính phủ và Bộ Tài nguyên và Môi trường đã ban hành.

- Công việc thực hiện: Thu thập, nghiên cứu các văn bản quy phạm pháp luật đã được ban hành, hướng dẫn về việc xác định các bộ dữ liệu, danh mục dữ liệu, các thông tin cần quản lý đối với các đối tượng quản lý KTTV Quốc gia từ đó đưa ra được sơ bộ về cấu trúc dữ liệu chi tiết cho CSDL Khí tượng thủy văn Quốc gia.

- Phương pháp thực hiện: Nghiên cứu văn pháp quy phạm pháp luật.

- Kết quả dự kiến cần đạt được:

(1) Xác định được các bộ dữ liệu, danh mục dữ liệu chủ của CSDL KTTV quốc gia.

(2) Hình thành sơ bộ về cấu trúc chi tiết của CSDL KTTV quốc gia (Bao gồm các đối tượng cần quản lý; thuộc tính của các đối tượng).

**Bước 2:** Điều tra, khảo sát, thu thập thông tin về hiện trạng dữ liệu, nguồn dữ liệu, mục đích và nhu cầu sử dụng dữ liệu phục vụ việc xây dựng bộ dữ liệu chủ và khung kiến trúc tổng thể của dữ liệu;

- Công việc thực hiện:

- (1) Khảo sát, thu thập thông tin về hiện trạng dữ liệu;
- (2) Khảo sát, thu thập thông tin về các nguồn dữ liệu;
- (3) Khảo sát, thu thập thông tin về định dạng dữ liệu;
- (4) Khảo sát, thu thập thông tin về mục đích, nhu cầu sử dụng dữ liệu.

- Phương pháp thực hiện: Điều tra khảo sát, xin ý kiến chuyên gia

- Kết quả dự kiến cần đạt được:

- (1) Hiện trạng về dữ liệu;
- (2) Hiện trạng về các nguồn dữ liệu;
- (3) Hiện trạng về định dạng dữ liệu;
- (4) Mục đích, nhu cầu sử dụng dữ liệu.

**Bước 3:** Phân tích, đánh giá, xác định khung kiến trúc tổng thể của dữ liệu và danh mục dữ liệu chủ cho CSDL KTTV Quốc gia

- Công việc thực hiện:

(1) Phân tích thông tin tại bước 1 và bước 2 thực hiện đánh giá đề xuất bộ danh mục dữ liệu chủ KTTV Quốc gia;

(2) Phối hợp với đơn vị để thực hiện rà soát, đánh giá kết quả phân tích, xác định, thống nhất bộ danh mục dữ liệu chủ KTTV Quốc gia;

- Phương pháp thực hiện: Phân tích kết hợp xin ý kiến chuyên gia

- Kết quả dự kiến cần đạt được

- (1) Khung kiến trúc và Danh mục dữ liệu chủ của CSDL KTTV quốc gia.

### *II.1.2.3. Xác định danh mục dữ liệu chủ*

Việc xác định các bộ dữ liệu và danh mục dữ liệu chủ KTTV được thực hiện trên cơ sở các văn bản quy định pháp luật của Chính phủ và Bộ Tài nguyên và Môi trường đã ban hành.

Thông tư số 26/2014/TT-BTNMT ngày 28 tháng 05 năm 2014 của Bộ Tài nguyên và Môi trường về việc ban hành Quy trình và Định mức kinh tế - kỹ thuật xây dựng cơ sở dữ liệu tài nguyên và môi trường định nghĩa: *Bộ dữ liệu là tập hợp các tài liệu/dữ liệu ở dạng giấy, dạng số có cùng nội dung và tính chất để làm tài liệu/dữ liệu đầu vào phục vụ cho việc xây dựng cơ sở dữ liệu.*

Thông tư số 07/2016/TT-BTNMT ngày 28/05/2014 của Bộ Tài nguyên và Môi trường ban hành quy định các bộ dữ liệu, chuẩn dữ liệu và xây dựng, quản lý cơ sở dữ liệu khí tượng thủy văn Quốc gia, bao gồm các bộ dữ liệu sau:

- Bộ dữ liệu về quan trắc, Điều tra, khảo sát khí tượng, thủy văn, hải văn, môi trường không khí và nước;
- Bộ dữ liệu về dự báo, cảnh báo khí tượng thủy văn;
- Bộ dữ liệu về hồ sơ kỹ thuật trạm và giấy phép dự báo, cảnh báo khí tượng thủy văn;
- Bộ dữ liệu về biến đổi khí hậu
- Bộ dữ liệu nghiên cứu khoa học, chương trình, dự án về khí tượng thủy văn;
- Bộ dữ liệu văn bản quy phạm pháp luật về khí tượng thủy văn

Trong đó, Bộ dữ liệu về biến đổi khí hậu không thuộc phạm vi quản lý của Tổng cục KTTV. Ngoài ra, căn cứ theo hiện trạng thực tế việc xây dựng Bộ dữ liệu KTTV cần bổ sung thêm Bộ dữ liệu quan trắc phi truyền thống thực hiện quản lý các dữ liệu Radar, dữ liệu định vị sét.

**Từ phân tích nêu trên, CSDL KTTV quốc gia cần có Danh mục dữ liệu chủ bao gồm các Bộ dữ liệu sau:**

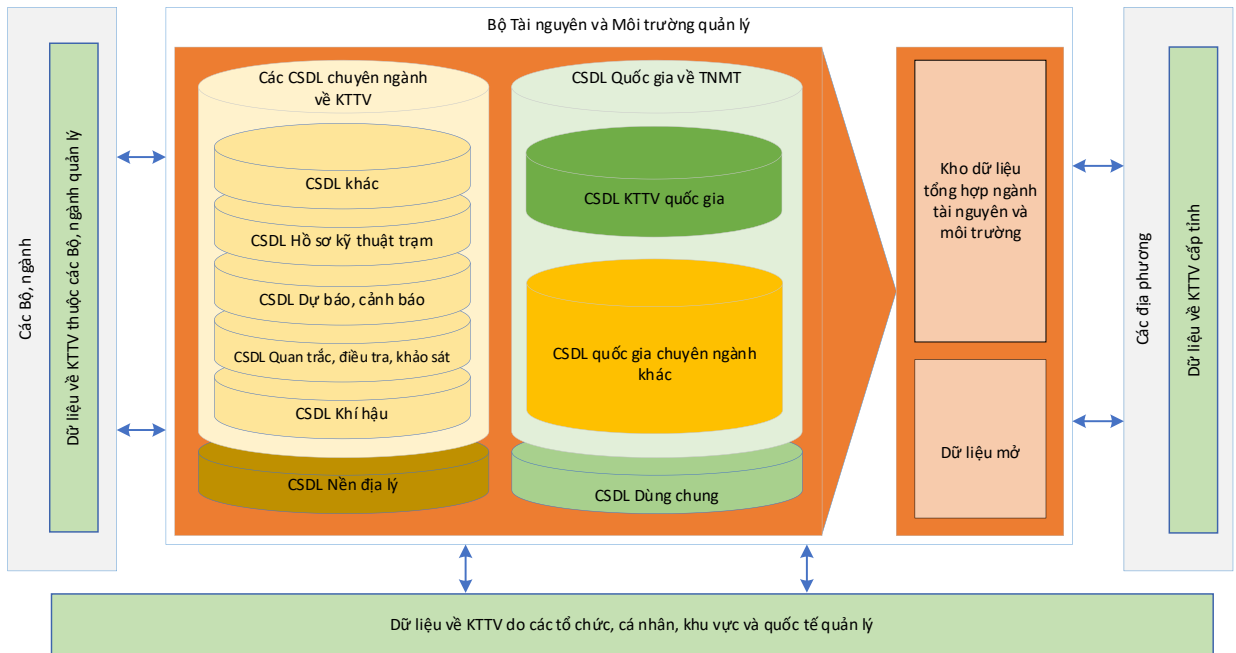
TT	Danh mục dữ liệu chủ
<b>I</b>	<b>Bộ dữ liệu Quan trắc, Điều tra, khảo sát khí tượng, thủy văn, hải văn, môi trường không khí và nước</b>
1	Dữ liệu khí tượng bề mặt
2	Dữ liệu khí tượng nông nghiệp
3	Dữ liệu thủy văn
4	Dữ liệu hải văn
5	Dữ liệu thám không vô tuyến
6	Dữ liệu gió trên cao (pilot)
7	Dữ liệu bức xạ
8	Dữ liệu ô dôn, bức xạ cực tím
9	Dữ liệu mưa
10	Dữ liệu môi trường không khí
11	Dữ liệu môi trường nước sông
12	Dữ liệu môi trường nước hồ
13	Dữ liệu môi trường nước biển ven bờ
14	Dữ liệu điều tra, khảo sát thủy văn
15	Dữ liệu điều tra, khảo sát khí tượng biển

<b>TT</b>	<b>Danh mục dữ liệu chủ</b>
16	Dữ liệu điều tra, khảo sát hải văn
17	Dữ liệu điều tra khảo sát về môi trường nước biển
18	Dữ liệu quan trắc, điều tra, khảo sát
19	Dữ liệu đơn vị thực hiện quan trắc, điều tra, khảo sát
20	Dữ liệu bản quyền thông tin khảo sát
21	Dữ liệu tài liệu thuyết minh
<b>II</b>	<b>Bộ dữ liệu Hồ sơ kỹ thuật trạm và giấy phép dự báo, cảnh báo khí tượng thủy văn</b>
22	Dữ liệu hồ sơ kỹ thuật của các trạm, công trình, phương tiện đo khí tượng thủy văn
23	Dữ liệu thiết bị quan trắc
24	Dữ liệu bảo dưỡng, sửa chữa, thay thế linh kiện, thiết bị quan trắc
25	Dữ liệu tổ chức/cá nhân được cấp phép hoạt động dự báo, cảnh báo khí tượng thủy văn
26	Dữ liệu hồ sơ cấp phép dự báo, cảnh báo
<b>III</b>	<b>Bộ dữ liệu Nghiên cứu khoa học, chương trình, dự án về khí tượng thủy văn</b>
27	Dữ liệu chương trình, đề án, dự án về khí tượng thủy văn, giám sát biến đổi khí hậu
28	Dữ liệu phim, ảnh về đối tượng nghiên cứu khí tượng thủy văn
29	Dữ liệu kế hoạch và kết quả thực hiện tác động vào thời tiết
<b>IV</b>	<b>Bộ dữ liệu Văn bản quy phạm pháp luật về khí tượng thủy văn</b>
30	Dữ liệu văn bản quy phạm pháp luật
<b>V</b>	<b>Bộ dữ liệu Quan trắc phi truyền thống</b>
31	Dữ liệu quan trắc radar thời tiết
32	Dữ liệu quan trắc định vị sét
<b>V</b>	<b>Bộ dữ liệu Dự báo, cảnh báo khí tượng thủy văn</b>
33	Dữ liệu quan trắc bản tin dự báo, cảnh báo khí tượng thủy văn
34	Dữ liệu quan trắc sản phẩm của các mô hình dự báo
35	Dữ liệu quan trắc dữ liệu mô hình dự báo
36	Dữ liệu quan trắc biểu đồ, bản đồ, ảnh thu từ vệ tinh

### ***II.1.3. Thiết kế cấu trúc dữ liệu chi tiết cho cơ sở dữ liệu khí tượng thủy văn quốc gia***

#### ***II.1.3.1. Mô hình dữ liệu tổng thể CSDL Khí tượng thủy văn Quốc gia***

- a) Mô hình kiến trúc tổng thể ngành Tài nguyên môi trường

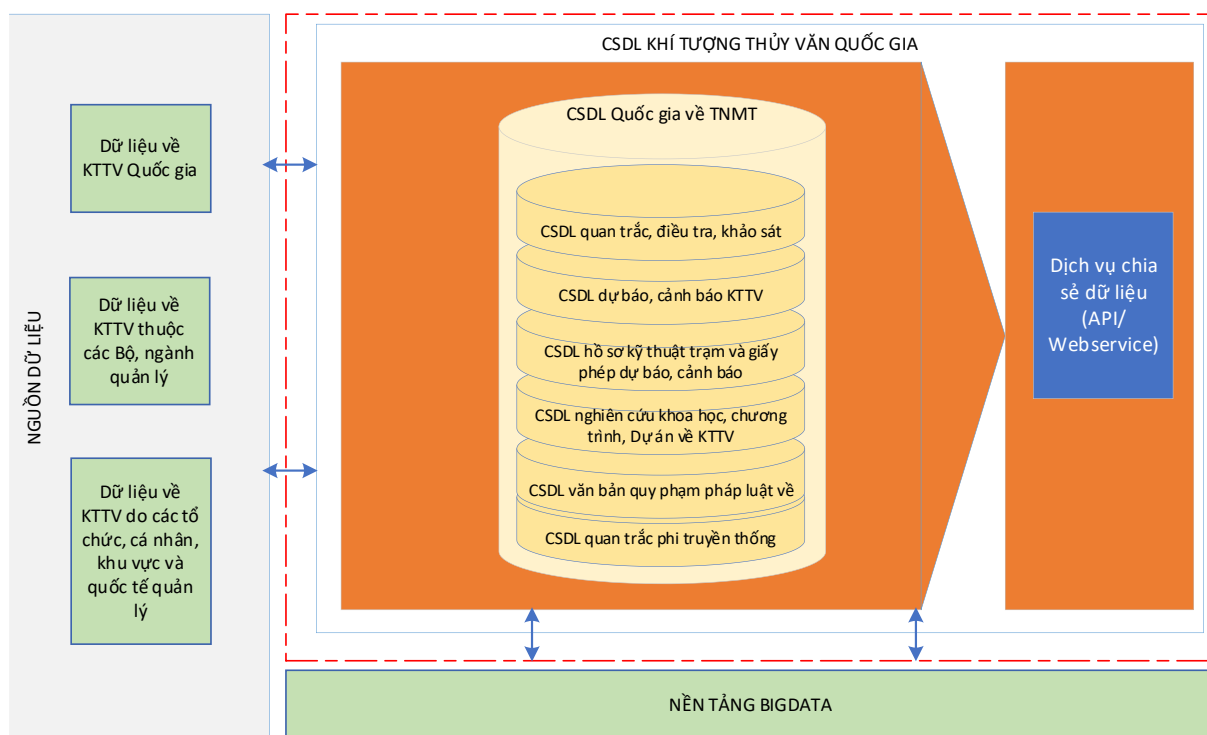


Mô hình Kiến trúc CSDL dữ liệu lĩnh vực KTTV Quốc gia được kế thừa từ mô hình Kiến trúc dữ liệu ngành tài nguyên và môi trường, trong đó thành phần cốt lõi là các thực thể thể hiện thông tin dữ liệu lĩnh vực KTTV Quốc gia do Bộ Tài nguyên và Môi trường quản lý có sự kết nối chia sẻ với các hệ CSDL KTTV Quốc gia do các tổ chức, cá nhân khác xây dựng, quản lý

CSDL KTTV Quốc gia tập hợp số liệu, thông tin, dữ liệu quốc gia về khí tượng thủy văn, bao gồm:

- Thông tin, dữ liệu về quan trắc, điều tra, khảo sát KTTV, hải văn, môi trường không khí và nước;
- Thông tin, dữ liệu về dự báo, cảnh báo khí tượng thủy văn;
- Thông tin, dữ liệu về hồ sơ kỹ thuật trạm và giấy phép dự báo, cảnh báo khí tượng thủy văn;
- Thông tin, dữ liệu về nghiên cứu khoa học, chương trình, Dự án về KTTV;
- Bộ dữ liệu Quan trắc phi truyền thống;
- Thông tin, dữ liệu văn bản quy phạm pháp luật về khí tượng thủy văn.

#### b) Mô hình kiến trúc CSDL Khí tượng thủy văn Quốc Gia



### Mô hình kiến trúc CSDL Khí tượng thủy văn Quốc gia

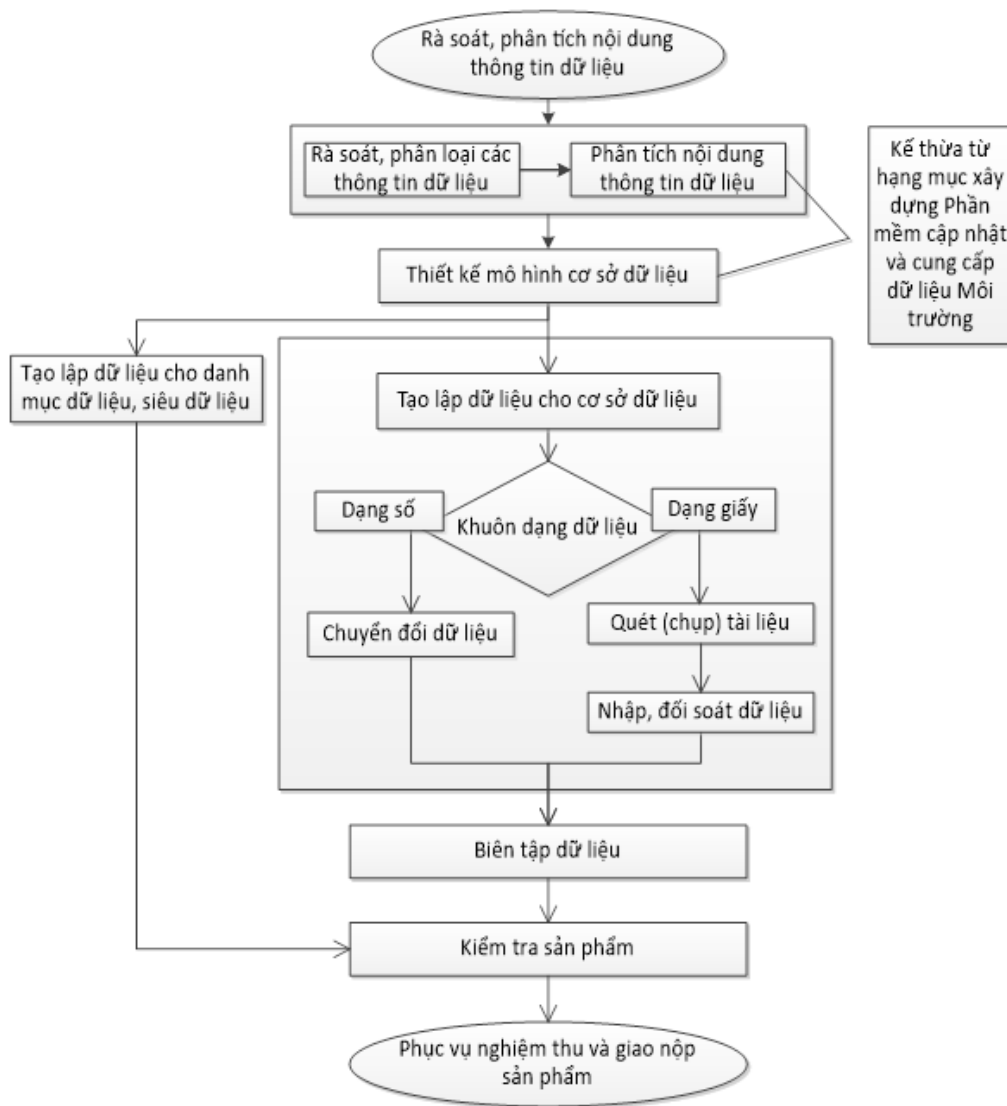
Kiến trúc nghiệp vụ CSDL KTTV Quốc gia gồm 2 thành phần chính bao gồm:

- Các CSDL thành phần: Bộ dữ liệu Quan trắc, Điều tra, khảo sát khí tượng, thủy văn, hải văn, môi trường không khí và nước; Bộ dữ liệu Hồ sơ kỹ thuật trạm và giấy phép dự báo, cảnh báo khí tượng thủy văn; Bộ dữ liệu Nghiên cứu khoa học, chương trình, dự án về khí tượng thủy văn; Bộ dữ liệu Văn bản quy phạm pháp luật về khí tượng thủy văn; Bộ dữ liệu Quan trắc phi truyền thống; Bộ dữ liệu Dự báo, cảnh báo khí tượng thủy văn.

- Dịch vụ chia sẻ dữ liệu (Api/Web Service) sẵn sàng chia sẻ dữ liệu khi có yêu cầu.

#### II.1.3.2. Quy trình thực hiện

Căn cứ quy trình xây dựng CSDL Tài nguyên và Môi trường được quy định tại Thông tư số 26/2014/TT-BTNMT của Bộ Tài nguyên và Môi trường ngày 28 tháng 05 năm 2014 của Bộ Tài nguyên và Môi trường ban hành Quy trình và Định mức kinh tế - kỹ thuật xây dựng cơ sở dữ liệu tài nguyên, môi trường, quy trình thực hiện xây dựng, thiết kế cấu trúc dữ liệu chi tiết cho cơ sở dữ liệu khí tượng thủy văn quốc gia bao gồm các nội dung sau:



*Sơ đồ quy trình xây dựng danh mục dữ liệu chủ và Xây dựng, thiết kế cấu trúc dữ liệu chi tiết cho cơ sở dữ liệu khí tượng thủy văn quốc gia*

## **Bước 1: Rà soát và phân tích nội dung thông tin dữ liệu**

### **Rà soát, phân loại các thông tin dữ liệu**

#### a) Mục đích

Rà soát, phân loại và đánh giá chi tiết các thông tin dữ liệu phục vụ xây dựng CSDL phù hợp với yêu cầu.

#### b) Các bước thực hiện

- Rà soát, đánh giá và phân loại chi tiết dữ liệu đã được chuẩn hóa và chưa được chuẩn hóa.

- Chuẩn bị dữ liệu mẫu.

#### c) Sản phẩm

- Báo cáo rà soát, phân loại và đánh giá các thông tin dữ liệu.
- Bộ dữ liệu mẫu.

### **Phân tích nội dung thông tin dữ liệu**

#### a) Mục đích

Phân tích, xác định chi tiết các thông tin dữ liệu phục vụ thiết kế và lập dự toán xây dựng CSDL.

#### b) Các bước thực hiện

- Xác định danh mục các đối tượng quản lý (ĐTQL).
- Xác định chi tiết các thông tin cho từng ĐTQL.
- Xác định chi tiết các quan hệ giữa các ĐTQL.
- Xác định chi tiết các tài liệu quét (tài liệu đính kèm) và các tài liệu dạng giấy cần nhập vào CSDL từ bàn phím.
- Xác định khung danh mục dữ liệu, siêu dữ liệu sử dụng trong CSDL.
- Xác định các yếu tố ảnh hưởng đến việc xây dựng CSDL.
- Quy đổi ĐTQL.

Ghi chú: Đối tượng quản lý phục vụ tính dự toán là ĐTQL đã được quy đổi theo các yếu tố ảnh hưởng (số lớp, bảng dữ liệu; kiểu dữ liệu; số trường thông tin, số lượng các quan hệ).

#### c) Sản phẩm

- Danh mục ĐTQL và các thông tin chi tiết (danh mục ĐTQL, các thông tin chi tiết cho từng ĐTQL, các quan hệ và các yếu tố ảnh hưởng đến việc xây dựng CSDL).
- Danh mục chi tiết các tài liệu quét và giấy cần nhập vào CSDL.
- Báo cáo quy định khung danh mục dữ liệu, siêu dữ liệu.
- Báo cáo Quy đổi ĐTQL.

### **Bước 2: Thiết kế mô hình cơ sở dữ liệu**

Trường hợp dự án hiện tại có cả hai nội dung xây dựng CSDL và xây dựng ứng dụng phần mềm thì các bước “Thiết kế mô hình CSDL” và “Nhập dữ liệu mẫu để kiểm tra mô hình CSDL” chỉ thực hiện một lần ở bước này.

#### a) Mục đích

- Thiết kế mô hình danh mục dữ liệu (data catalogue), siêu dữ liệu (Metadata) theo (chuẩn dữ liệu, khung dữ liệu) dựa trên kết quả rà soát, phân tích.

- Thiết kế mô hình CSDL dựa trên kết quả rà soát, phân tích.

b) Các bước thực hiện

- Thiết kế mô hình danh mục dữ liệu, siêu dữ liệu.

- Thiết kế mô hình CSDL:

- + Thiết kế mô hình CSDL.

- + Nhập dữ liệu mẫu để kiểm tra mô hình CSDL.

c) Sản phẩm

- Mô hình CSDL, mô hình danh mục dữ liệu, siêu dữ liệu dưới dạng XML.

- Báo cáo thuyết minh mô hình danh mục dữ liệu, siêu dữ liệu.

- Báo cáo thuyết minh mô hình CSDL.

- Báo cáo kết quả kiểm tra mô hình CSDL trên dữ liệu mẫu.

### **Bước 3: Tạo lập dữ liệu cho danh mục dữ liệu, siêu dữ liệu**

a) Mục đích

Tạo lập nội dung dữ liệu cho danh mục dữ liệu, siêu dữ liệu dựa trên kết quả rà soát, phân tích và thiết kế.

b) Các bước thực hiện

- Tạo lập nội dung cho danh mục dữ liệu.

- Tạo lập nội dung cho siêu dữ liệu.

c) Sản phẩm

- Cơ sở dữ liệu danh mục dữ liệu, siêu dữ liệu đã nhập đủ nội dung.

- Báo cáo kết quả thực hiện.

### **Bước 4: Tạo lập dữ liệu cho cơ sở dữ liệu**

#### **Chuyển đổi dữ liệu**

a) Mục đích

Chuyển đổi dữ liệu dạng số (không gian và phi không gian) đã được chuẩn hóa vào CSDL.

b) Các bước thực hiện

- Đối với dữ liệu không gian dạng số chưa được chuẩn hóa thì việc chuẩn hóa dữ liệu được thực hiện theo các quy định của từng chuyên ngành trước khi thực hiện chuyển đổi vào CSDL (biên tập bản đồ, chuyển đổi hệ tọa độ,...).

- Đối với dữ liệu phi không gian dạng số chưa được chuẩn hóa:
  - + Chuẩn hóa phông chữ theo tiêu chuẩn TCVN 6909 (nếu có).
  - + Chuẩn hóa dữ liệu phi không gian theo thiết kế mô hình CSDL.
- Chuyển đổi dữ liệu dạng số đã chuẩn hóa vào CSDL.

c) Sản phẩm

- Dữ liệu dạng số trước khi chuyển đổi.
- Dữ liệu phi không gian trước khi chuẩn hóa.
- Cơ sở dữ liệu đã được chuyển đổi.
- Báo cáo kết quả thực hiện chuyển đổi dữ liệu.

**Quét (chụp) tài liệu**

a) Mục đích

Quét (chụp) các tài liệu để phục vụ đính kèm vào các trường thông tin cho các lớp, bảng dữ liệu của ĐTQL.

b) Các bước thực hiện

- Quét (chụp) các tài liệu.
- Xử lý và đính kèm tài liệu quét.

c) Sản phẩm

- Danh mục các tài liệu quét và đã được đính kèm vào các lớp, bảng dữ liệu của các ĐTQL.

Ghi chú: Đối với dự án hiện tại, phạm vi thực hiện ưu tiên các loại dữ liệu là các sản phẩm đã được kỹ thuật số hoá, tin học hoá. Do đó, công tác Quét (chụp) tài liệu này chỉ thực hiện khi đã hoàn tất toàn bộ các dữ liệu được ưu tiên.

**Nhập, đối soát dữ liệu**

a) Mục đích

Nhập, đối soát các dữ liệu từ dạng giấy vào CSDL đã được thiết kế. Dữ liệu sau khi nhập vào CSDL phải được đối chiếu, kiểm soát để đảm bảo tính chính xác dữ liệu.

## b) Các bước thực hiện

- Đối với các dữ liệu không gian dạng giấy: số hóa theo quy định chuyên ngành sau đó thực hiện bước “Chuyển đổi dữ liệu”.

- Đối với nhập dữ liệu dạng giấy (phi không gian):

+ Nhập dữ liệu có cấu trúc cho đối tượng phi không gian.

+ Nhập dữ liệu có cấu trúc cho đối tượng không gian.

+ Nhập dữ liệu phi cấu trúc cho đối tượng phi không gian.

+ Nhập dữ liệu phi cấu trúc cho đối tượng không gian.

Ghi chú: Đối việc cập nhật dữ liệu của những trường hợp chỉ cập nhật bổ sung dữ liệu thì yêu cầu cập nhật bổ sung thông tin theo Mẫu đính kèm Thông tư 26/2014/TT-BTNMT để phân loại dữ liệu cần cập nhật bổ sung tương ứng theo các bước đã nêu ở trên.

- Đối soát dữ liệu:

+ Dữ liệu có cấu trúc đã nhập cho đối tượng phi không gian.

+ Dữ liệu có cấu trúc đã nhập cho đối tượng không gian.

+ Dữ liệu phi cấu trúc đã nhập cho đối tượng phi không gian.

+ Dữ liệu phi cấu trúc đã nhập cho đối tượng không gian.

## c) Sản phẩm

- Dữ liệu dạng giấy dùng để nhập dữ liệu (được lưu trữ ở đơn vị thi công phục vụ kiểm tra, nghiệm thu của chủ đầu tư khi có yêu cầu).

- Báo cáo đối soát dữ liệu và các vấn đề phát sinh trong quá trình nhập dữ liệu.

- CSDL đã được nhập đầy đủ nội dung.

- Danh mục dữ liệu để cung cấp, khai thác, sử dụng phục vụ yêu cầu quản lý nhà nước.

## **Bước 5: Biên tập dữ liệu**

### a) Mục đích

Biên tập cơ sở dữ liệu theo quy định.

### b) Các bước thực hiện

- Đối với dữ liệu không gian.

- + Tuyên bố đối tượng.
- + Sửa lỗi tương quan của dữ liệu không gian (topology).
- Đối với dữ liệu phi không gian: Hiệu đính nội dung.
- Trình bày hiển thị dữ liệu không gian.

c) Sản phẩm

- CSDL đã được biên tập.
- File trình bày hiển thị dữ liệu không gian.

**Bước 6: Kiểm tra sản phẩm**

a) Mục đích

Kiểm tra CSDL đã được tạo lập đảm bảo tính đầy đủ, chính xác, phù hợp với nội dung đã được phê duyệt.

b) Các bước thực hiện

- Kiểm tra mô hình CSDL.
- Kiểm tra nội dung CSDL.
- + Kiểm tra dữ liệu không gian.
- + Kiểm tra dữ liệu phi không gian.
- Kiểm tra danh mục dữ liệu, siêu dữ liệu.

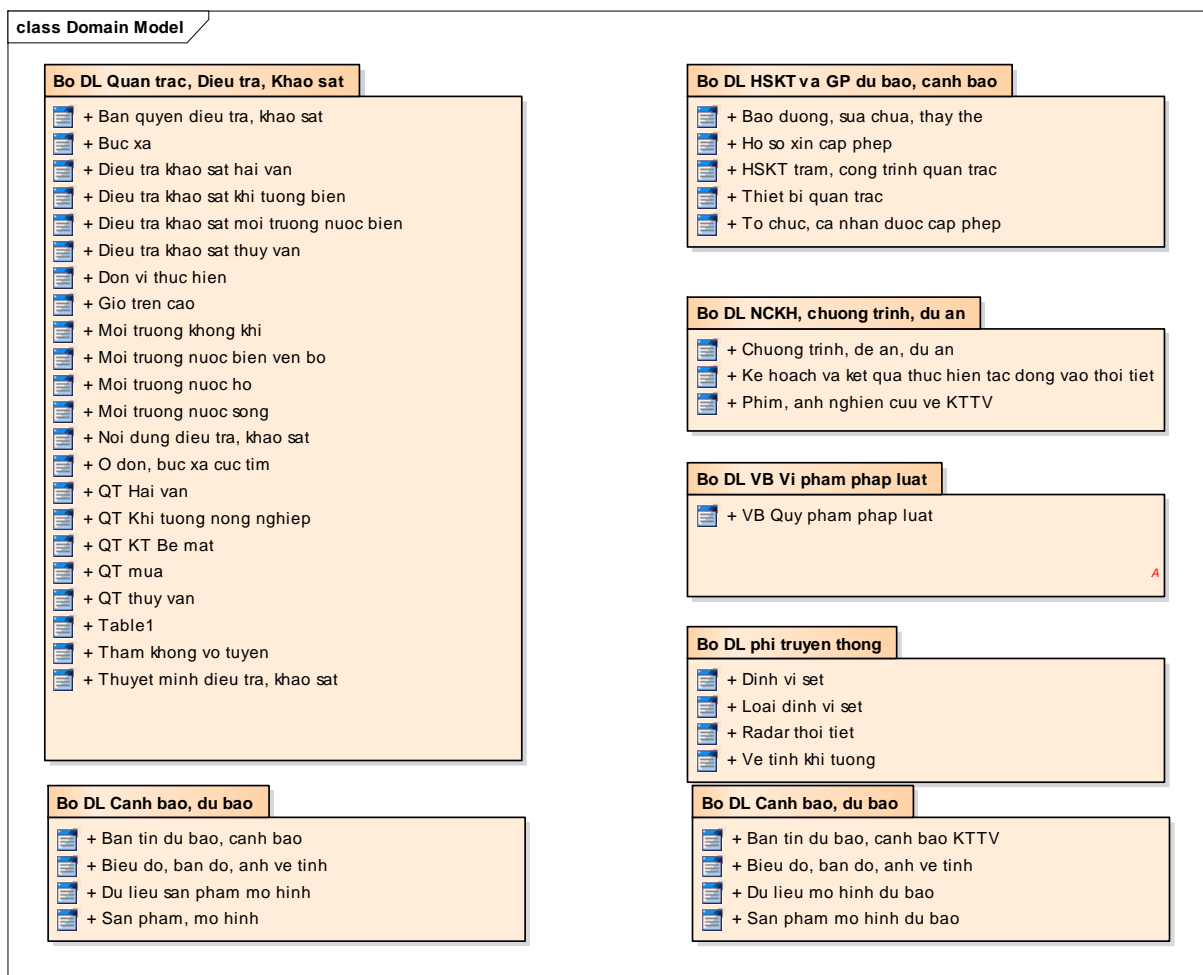
c) Sản phẩm

- Báo cáo kết quả kiểm tra sản phẩm.
- Báo cáo kết quả sửa chữa.
- Báo cáo kiểm tra, nghiệm thu chất lượng, khối lượng.

Việc thiết kế cấu trúc dữ liệu chi tiết cho cơ sở dữ liệu khí tượng thủy văn quốc gia thuộc dự án hiện tại **không bao gồm Bước 4 - Tạo lập dữ liệu cho cơ sở dữ liệu. Đối với Bước 5 - Biên tập dữ liệu, chỉ thực hiện công việc Tuyên bố đối tượng.**

*II.1.3.3. Mô hình dữ liệu chi tiết cho cơ sở dữ liệu KTTV quốc gia*

Mô hình dữ liệu cho cơ sở dữ liệu khí tượng thủy văn quốc gia thể hiện như sau:



#### II.1.3.4. Thiết kế chi tiết kiến trúc dữ liệu

a) Bộ dữ liệu về quan trắc, Điều tra, khảo sát khí tượng, thủy văn, hải văn, môi trường không khí và nước

(1) Thông tin đối tượng “Quan trắc khí tượng bề mặt”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>QT KT Be mat (Quan trắc khí tượng bề mặt)</b>
<b>Mô tả</b>	Là thông tin dữ liệu về Khí tượng bề mặt
<b>Kiểu cơ sở</b>	Table
<b>Tên các thuộc tính</b>	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc,

	<p>Nhiệt độ, Nhiệt độ tối cao, Nhiệt độ tối thấp, Nhiệt độ ướt, Nhiệt độ điểm sương, Sức trương hơi nước (Âm độ tuyệt đối), Sức trương hơi nước bão hòa, Âm độ tương đối, Độ hụt bão hòa, Khí áp mực trạm, Khí áp tối cao mực trạm, Khí áp tối thấp mực trạm, Khí áp mực biển, Khí áp tối cao mực biển, Khí áp tối cao mực biển, Hướng gió, Tốc độ gió, Hướng gió mạnh nhất trong ngày, Tốc độ gió mạnh nhất trong ngày, Hướng gió trung bình trong hai phút, Tốc độ gió trung bình trong hai phút, Hướng gió giật mạnh nhất trong 2 giây, Tốc độ gió giật mạnh nhất trong 2 giây, Lượng bốc hơi, Nhiệt độ thường mặt đất, Nhiệt độ đất tối cao, Nhiệt độ đất tối thấp, Nhiệt độ đất tại các lớp đất sâu (5,...,300), Trạng thái mặt đất, Lượng giáng thủy, Cường độ giáng thủy, Kích thước (rắn), Độ dày (tuyết), Lượng mây tổng quan,</p>
--	---

	Lượng mây dưới, Mây giữa, Mây trên, Mây dưới, Loại mây, Dạng mây, Tính mây, dạng mây phụ, Độ cao chân mây, Hiện tượng khí tượng, Thời gian nắng, Tầm nhìn ngang
Kiểu dữ liệu	Phi không gian

(2) Thông tin đối tượng “Quan trắc khí tượng nông nghiệp”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>QT KT nông nghiệp (Quan trắc khí tượng nông nghiệp)</b>
Mô tả	Là đối tượng thông tin dữ liệu về Khí tượng Nông nghiệp
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Nhiệt độ nước trên ruộng, Độ ẩm đất đo theo phương pháp định lượng (khoan và hòm sấy), Độ ẩm đất đo theo phương pháp định tính (ước lượng bằng mắt), Xáo trộn không khí tại các lớp không khí gần mặt đất, Lượng nước trong đất tính ra %, Lượng nước trong đất tính ra mm Lượng nước có ích
Kiểu dữ liệu	Phi không gian

### (3) Thông tin đối tượng “Quan trắc thủy văn”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>QT thủy văn (Quan trắc thủy văn)</b>
Mô tả	Là đối tượng thông tin dữ liệu về quan trắc thủy văn
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Nhiệt độ nước tối cao, Nhiệt độ nước, Nhiệt độ nước tối thấp, Lượng mưa ngày, Mức nước từng giờ, Độ sâu, Độ rộng mặt nước, Chiều dài lưu vực, Diện tích lưu vực, Diện tích mặt cắt ngang, Tốc độ, Lưu lượng nước, Độ dốc, Hệ số nhám lòng sông, Tổng lượng dòng chảy, Lượng triều, Hàm lượng chất lơ lửng, Lưu lượng chất lơ lửng, Tổng lượng chất lơ lửng
Kiểu dữ liệu	Phi không gian

### (4) Thông tin đối tượng “Quan trắc hải văn”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>QT thuy van (Quan trắc hải văn)</b>
Mô tả	Là đối tượng thông tin dữ liệu về quan trắc hải văn
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Tốc độ gió, Hướng gió, Tầm nhìn xa phía biển, Mực nước biển, Mực nước biển TB, Thủy triều, Dao động dư, Loại sóng biển, Độ cao sóng, Biên độ sóng, Độ dài sóng, Độ dốc sóng, Chu kỳ sóng, Tốc độ truyền sóng, Tần số sóng, Hướng sóng, Dạng sóng, Cấp sóng, Trạng thái mặt biển, Nhiệt độ bề mặt biển, Độ muối, Kiểu sáng biển, Cường độ sáng biển,

	Các hiện tượng khí tượng hải văn nguy hiểm
Kiểu dữ liệu	Phi không gian

(5) Thông tin đối tượng “Thám không vô tuyến”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Tham không vô tuyến</b>
Mô tả	Là đối tượng thông tin dữ liệu về thám không vô tuyến
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Khí áp, Nhiệt độ, Độ ẩm, Hướng gió, Tốc độ gió, Độ cao địa thế vị, Tỷ ẩm, Nhiệt độ thế vị, Nhiệt độ thế vị tương đương, Nhiệt độ thế vị ảo
Kiểu dữ liệu	Phi không gian

(6) Thông tin đối tượng “Gió trên cao (Pilot)”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Gió trên cao</b>
Mô tả	Là đối tượng thông tin dữ liệu về gió trên cao (Pilot)
Kiểu cơ sở	Table

Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Hướng gió, Tốc độ gió
Kiểu dữ liệu	Phi không gian

(7) Thông tin đối tượng Bức xạ

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Buc xa</b>
Mô tả	Là đối tượng thông tin dữ liệu về bức xạ
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Cường độ bức xạ trực tiếp, Cường độ trực xạ trên mặt phẳng ngang, Cán cân bức xạ (được tính), Cán cân bức xạ sóng ngắn, Cán cân bức xạ không có bức xạ trực tiếp, Cán cân bức xạ sóng dài, Trị số phản xạ sóng ngắn, Trị số bức xạ khuếch tán, Anbedo, Tổng xạ, Thời sai,

	Độ cao mặt trời, Màu sắc bầu trời, Trạng thái mặt đốm, Trạng thái đĩa mặt trời, Độ trong suốt khí quyển, Độ vẩn đục khí quyển
Kiểu dữ liệu	Phi không gian

(8) Thông tin đối tượng “Ô dôn, bức xạ cực tím”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>O don, buc xa cuc tím</b>
Mô tả	Là đối tượng thông tin dữ liệu về ô dôn, bức xạ cực tím
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Điều kiện khí tượng, Dạng mây, Độ cao mặt trời, Loại quan trắc, Cường độ ánh sáng theo phin I, Cường độ ánh sáng theo phin II, Độ nhạy, Nhiệt độ, Hệ số nhiệt, Hệ số dạng mây, Tổng lượng ô dôn quan trắc theo ánh sáng mặt trời trực tiếp, Tổng lượng ô dôn quan trắc theo ánh sáng tán xạ của vùng trời thiên đỉnh, Tổng lượng ô dôn của lần đo,

	<p>Tổng lượng ô dôn của kỳ quan trắc,  Số lượng kỳ quan trắc theo ánh sáng trực tiếp của mặt trời trong ngày,  Số lượng kỳ quan trắc theo ánh sáng tán xạ của bầu trời thiên đỉnh,  Tổng lượng ô dôn trung bình ngày quan trắc theo ánh sáng trực tiếp của mặt trời,  Tổng lượng ô dôn trung bình ngày quan trắc theo ánh sáng tán xạ của bầu trời thiên đỉnh,  Cường độ bức xạ trong miền tử ngoại B,  Cường độ bức xạ trong miền tử ngoại A,  Cường độ bức xạ Eriten,  Hệ số kiểm định,  Hệ số suy giảm năng lượng,  Hệ số suy giảm năng lượng bức xạ Eriten trong miền tử ngoại A, B,  Hệ số nhiệt của dụng cụ,  Cường độ ánh sáng,  Thời gian tối ưu</p>
Kiểu dữ liệu	Phi không gian

(9) Thông tin đối tượng “Mưa”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>QT Mưa</b>
Mô tả	Là đối tượng thông tin dữ liệu về quan trắc lưu lượng mưa
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Lượng mưa

Kiểu dữ liệu	Phi không gian
--------------	----------------

(10) Thông tin đối tượng “Môi trường không khí”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Moi truong khong khi</b>
Mô tả	Là đối tượng thông tin dữ liệu về quan trắc môi trường không khí
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Bụi lắng tổng cộng tháng, Lượng nước sau mỗi lần lọc, Tổng lượng nước thu được sau lọc, Số bình lấy mẫu, Số trận mưa trong tuần, Tổng lượng mưa hứng được, Độ dẫn điện, Độ pH, Hàm lượng NH <sub>4</sub> <sup>+</sup> , Hàm lượng Na <sup>+</sup> , Hàm lượng K <sup>+</sup> , Hàm lượng Ca <sup>2+</sup> , Hàm lượng Mg <sup>2+</sup> , Hàm lượng Cl <sup>-</sup> , Hàm lượng NO <sub>3</sub> <sup>-</sup> , Hàm lượng SO <sub>4</sub> <sup>2-</sup> , Hàm lượng HCO <sub>3</sub> <sup>-</sup> , Tổng lượng ion, Sai số phân tích

Kiểu dữ liệu	Phi không gian
--------------	----------------

(11) Thông tin đối tượng “Môi trường nước sông”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Moi truong nuoc song</b>
Mô tả	Là đối tượng thông tin dữ liệu về quan trắc môi trường nước sông
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Trạng thái dòng chảy, Các vật nổi, hiện tượng khác thường, Mùi, Vị, Độ dẫn điện, Độ pH, Tổng lượng ion, Hàm lượng Na+, Hàm lượng K+, Hàm lượng Ca2+, Hàm lượng, Mg2+, Hàm lượng CO32- Hàm lượng, SO42- Hàm lượng HCO3- Hàm lượng, HCO3- Sai số phân tích,

	<p>Tổng sắt, Hàm lượng Silic điôxít, Độ kiềm toàn phần, Độ cứng toàn phần, Độ đục, Độ mặn, Nhu cầu oxy hóa học, Ôxy hòa tan, Hàm lượng, PO<sub>3-4</sub> (P), Hàm lượng F , Hàm lượng NO<sub>2</sub>-, Hàm lượng NCO<sub>3</sub>-, Hàm lượng NH<sub>4</sub>+, Hàm lượng tổng N, Hàm lượng tổng P, Hàm lượng Cr (VI), Hàm lượng Asen, Hàm lượng thủy ngân, Hàm lượng Kẽm, Hàm lượng Cadimi, Hàm lượng Chì, Hàm lượng đồng, Hàm lượng Niken, Hàm lượng Mangan, Hàm lượng Cr tổng, Dầu mỡ, Tổng Colifom, Dư lượng hóa chất bảo vệ thực vật - Clo hữu cơ, Aldrin + Dieldrin , Endrin , BHC, DDT, Endounfan,</p>
--	---

	Lindan, Chlordance, Heptachlor, Du lượng hóa chất bảo vệ thực vật - Phốt pho hữu cơ, Paration, Malation, Tổng thể tích hút được, Hàm lượng bụi tổng số, Hàm lượng NH <sub>4</sub> <sup>+</sup> , Hàm lượng SO <sub>4</sub> <sup>2-</sup> , Hàm lượng NCO <sub>3</sub> <sup>-</sup> , Hàm lượng Chì
Kiểu dữ liệu	Phi không gian

(12) Thông tin đối tượng “Môi trường nước hồ”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Moi truong nuoc ho</b>
Mô tả	Là đối tượng thông tin dữ liệu về môi trường nước hồ
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc, Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Sóng và cấp sóng, Trạng thái dòng chảy, Các vật trôi nổi, Sự phát triển của thủy sinh vật, Các hiện tượng khác thường, Độ trong, Mùi,

	V <sub>i</sub> , Độ pH, Oxy hòa tan, Độ dẫn điện, Độ mặn, Tổng lượng ion, Hàm lượng Na <sup>+</sup> , Hàm lượng K <sup>+</sup> , Hàm lượng Ca <sup>2+</sup> , Hàm lượng Mg <sup>2+</sup> , Hàm lượng CO <sub>3</sub> <sup>2-</sup> , Hàm lượng SO <sub>4</sub> <sup>2-</sup> , Hàm lượng HCO <sub>3</sub> <sup>-</sup> , Hàm lượng Cl <sup>-</sup> , Sai số phân tích, Tổng sắt, Hàm lượng Silic điôxít, Độ kiềm toàn phần, Độ cứng toàn phần, Độ đục, Độ mặn, Nhu cầu oxy hóa học, Oxy hòa tan
Kiểu dữ liệu	Phi không gian

(13) Thông tin đối tượng “Môi trường nước biển ven bờ”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Moi truong nuoc bien ve bo</b>
Mô tả	Là đối tượng thông tin dữ liệu về môi trường nước biển ven bờ
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID trạm quan trắc,

	Thời gian quan trắc, ID đơn vị cung cấp, Thời gian cung cấp, File dữ liệu gốc, Dòng chảy ven bờ, Hướng chảy, Cường độ chảy, Màu nước, Váng dầu, Phát triển tảo, Đột biến độ đục, Các vật trôi nổi, Độ pH, Độ dẫn điện, Độ mặn, Nhu cầu ô xy hóa sinh học BOD5, Hàm lượng NH4+, Hàm lượng NO3- Hàm lượng NO2- Hàm lượng PO4- Hàm lượng Silic, Hàm lượng Chì, Hàm lượng Đồng, Nhu cầu oxy hóa học, Ôxy hòa tan
Kiểu dữ liệu	Phi không gian

(14) Thông tin đối tượng “Điều tra khảo sát hải văn”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Điều tra khảo sát hải văn</b>
Mô tả	Là đối tượng thông tin dữ liệu về điều tra khảo sát hải văn
Kiểu cơ sở	Table

Tên các thuộc tính	ID, ID Nội dung quan trắc, điều tra, khảo sát, Thời gian cung cấp, Thời gian thực hiện điều tra, Địa điểm thực hiện điều tra, Họ và tên người chủ trì thực hiện, ID Sông, Thông tin mô tả nội dung điều tra, khảo sát, Điều tra và xác nhận vết lũ, Điều tra thời gian xuất hiện đỉnh lũ, Điều tra khái quát địa mạo đoạn sông
Kiểu dữ liệu	Phi không gian

(15) Thông tin đối tượng “Điều tra khảo sát khí tượng biển”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Điều tra khảo sát khí tượng biển</b>
Mô tả	Là đối tượng thông tin dữ liệu về điều tra khảo sát khí tượng biển
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID Nội dung quan trắc, điều tra, khảo sát, Thời gian cung cấp, Thời gian thực hiện điều tra, Vị trí thực hiện điều tra, khảo sát, Họ và tên người chủ trì thực hiện, Thông tin mô tả nội dung điều tra, khảo sát, Bức xạ mặt trời, Độ trong suốt nước biển
Kiểu dữ liệu	Phi không gian

(16) Thông tin đối tượng “Điều tra khảo sát hải văn”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Điều tra khảo sát hải văn</b>
Mô tả	Là đối tượng thông tin dữ liệu về điều tra khảo sát hải văn
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID Nội dung quan trắc, điều tra, khảo sát, Thời gian cung cấp, Thời gian thực hiện điều tra, Vị trí thực hiện điều tra, khảo sát, Hướng, Tốc độ
Kiểu dữ liệu	Phi không gian

(17) Thông tin đối tượng “Điều tra khảo sát môi trường nước biển”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Điều tra khảo sát môi trường nước biển</b>
Mô tả	Là đối tượng thông tin dữ liệu về điều tra khảo sát môi trường nước biển
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID Nội dung quan trắc, điều tra, khảo sát, Thời gian cung cấp, Thời gian thực hiện điều tra, Vị trí thực hiện điều tra, khảo sát, Tổng dầu tầng mặt , Hàm lượng cadimi, Hàm lượng đồng, Hàm lượng chì, Hàm lượng sắt, Hàm lượng kẽm, Hàm lượng Asen, Hàm lượng thủy ngân,

	Hàm lượng Mangan, Hàm lượng Sắt, Hàm lượng Niken, Muối Phốt phát, Nitrit, Nitrat, Ammoni, Silicat, Thực vật nổi, Động vật nổi, Sinh vật đáy
Kiểu dữ liệu	Phi không gian

(18) Thông tin đối tượng “Quan trắc, điều tra, khảo sát”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Nội dung điều tra, khảo sát</b>
Mô tả	Là đối tượng thông tin dữ liệu về nội dung quan trắc, điều tra, khảo sát
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Nhan đề, Loại dữ liệu quan trắc, điều tra, khảo sát khí tượng, Chủ đề chính, Chủ đề thông tin dữ liệu quan trắc, điều tra, khảo sát, Mô tả, tóm tắt dữ liệu quan trắc, điều tra, khảo sát, ID Đơn vị thực hiện quan trắc, điều tra, khảo sát, ID bản quyền tác giả, ID thuyết minh
Kiểu dữ liệu	Phi không gian

(19) Thông tin đối tượng “Đơn vị thực hiện quan trắc, điều tra, khảo sát”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Đơn vị thực hiện</b>
Mô tả	Là đối tượng thông tin dữ liệu về đơn vị thực hiện quan trắc, điều tra, khảo sát
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Tên đơn vị, Người đại diện, Chức vụ, Vai trò, Địa chỉ, Mô tả, Thời gian có thể liên hệ, Chỉ dẫn liên hệ, Điện thoại, Địa chỉ chi tiết, Huyện, Tỉnh, Quốc gia, Thư điện tử
Kiểu dữ liệu	Phi không gian

(20) Thông tin đối tượng “Bản quyền thông tin khảo sát”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Bản quyền điều tra, khảo sát</b>
Mô tả	Là đối tượng thông tin dữ liệu về Bản quyền thông tin khảo sát
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Tác giả, Tác giả phụ, Nhà xuất bản,

	Bản quyền, File đính kèm
Kiểu dữ liệu	Phi không gian

(21) Thông tin đối tượng “Tài liệu thuyết minh”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Thuyet minh dieu tra, khao sat</b>
Mô tả	Là đối tượng thông tin dữ liệu về tài liệu thuyết minh
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Ngày tháng, Nơi chứa, Liên kết, Định danh, Ngôn ngữ
Kiểu dữ liệu	Phi không gian

*b) Bộ dữ liệu về hồ sơ kỹ thuật trạm và giấy phép dự báo, cảnh báo khí tượng thủy văn*

(1) Thông tin đối tượng “Hồ sơ kỹ thuật của các trạm, công trình, phương tiện đo khí tượng thủy văn”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	
Mô tả	Là đối tượng thông tin dữ liệu về hồ sơ kỹ thuật của các trạm, công trình, phương tiện đo khí tượng thủy văn
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Mã hồ sơ, Loại trạm, Mã trạm, Tên trạm,

	Ảnh trạm, Sơ đồ vị trí trạm, Bản vẽ thiết kế trạm, Quyết định thành lập trạm, Quyết định/thỏa thuận giao đất, Vị trí địa lý, Địa chỉ trạm, Mốc độ cao, Ngày bắt đầu hoạt động, Chức năng nhiệm vụ của trạm, Đơn vị quản lý, Cán bộ phụ trách, Số lượng cán bộ làm việc tại trạm, Mô tả đoạn sông đặt trạm, Vị trí công trình, Hành lang kỹ thuật, Sơ đồ hệ thống mốc giới hành lang kỹ thuật trạm, Sơ đồ công trình quan trắc, Công trình quan trắc, Cột/Tháp quan trắc, Sơ đồ lắp đặt máy, thiết bị, Yếu tố quan trắc, Tần suất truyền tin, Định dạng dữ liệu, Hệ thống chống sét, Hệ thống truyền tin, Hệ thống lưu trữ số liệu
Kiểu dữ liệu	Phi không gian

(2) Thông tin đối tượng “Thiết bị quan trắc”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Thiết bị quan trắc</b>
Mô tả	Là đối tượng thông tin dữ liệu về thông tin thiết bị quan trắc

Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID Trạm quan trắc, Tên thiết bị, Thông số kỹ thuật, Loại thiết bị, Năm sản xuất, Ngày kiểm định/hiệu chuẩn, Ngày hết hạn, kiểm định/hiệu chuẩn, Ngày đưa vào sử dụng, Thời hạn sử dụng
Kiểu dữ liệu	Phi không gian

(3) Thông tin đối tượng “Bảo dưỡng, sửa chữa, thay thế linh kiện, thiết bị quan trắc”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Bao duong, sua chua, thay the</b>
Mô tả	Là đối tượng thông tin dữ liệu về thông tin bảo dưỡng, sửa chữa, thay thế linh kiện, thiết bị quan trắc
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID Trạm quan trắc, ID thiết bị, Người thực hiện, Thời gian, Nội dung sửa chữa, bảo dưỡng, thay thế
Kiểu dữ liệu	Phi không gian

(4) Thông tin đối tượng “Tổ chức/cá nhân được cấp phép hoạt động dự báo, cảnh báo khí tượng thủy văn”

<b>Kiểu đối tượng:</b>	
------------------------	--

<b>Tên</b>	<b>To chuc, ca nhan duoc cap phep</b>
Mô tả	Là đối tượng thông tin dữ liệu về thông tin tổ chức/cá nhân được cấp phép hoạt động dự báo, cảnh báo khí tượng thủy văn
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Loại đơn vị, Loại hồ sơ, Tên đơn vị/cá nhân, Số giấy phép, Ngày cấp phép, Người ký, Chức vụ, Lần cấp phép/gia hạn, Phạm vi hoạt động cảnh báo, dự báo, Loại hoạt động cảnh báo, dự báo được cấp phép, Đối tượng cung cấp thông tin dự báo, cảnh báo, Thời hạn của giấy phép, Địa chỉ , Người đại diện, Chức vụ, Địa chỉ, Điện thoại, File đính kèm
Kiểu dữ liệu	Phi không gian

(5) Thông tin đối tượng “Hồ sơ cấp phép dự báo, cảnh báo”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Ho so xin cap phep</b>
Mô tả	Là đối tượng thông tin dữ liệu về hồ sơ cấp phép dự báo, cảnh báo
Kiểu cơ sở	Table
Tên các thuộc tính	ID,

	ID giáp phép, Tên tài liệu/ hồ sơ, File hồ sơ
Kiểu dữ liệu	Phi không gian

c) Bộ dữ liệu nghiên cứu khoa học, chương trình, dự án về khí tượng thủy văn

(1) Thông tin đối tượng “Chương trình, đề án, dự án về khí tượng thủy văn, giám sát biến đổi khí hậu”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Chương trình, đề án, dự án</b>
Mô tả	Là đối tượng thông tin dữ liệu về chương trình, đề án, dự án về khí tượng thủy văn, giám sát biến đổi khí hậu
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Mã chương trình, đề án, dự án, Tên chương trình, đề án, dự án, Mục tiêu chương trình, đề án, dự án, Thời gian thực hiện, Tổng vốn đầu tư, Cấp thực hiện, Số quyết định phê duyệt chương trình, đề án, dự án, Ngày phê duyệt, Người phê duyệt, Chức vụ
Kiểu dữ liệu	Phi không gian

(2) Thông tin đối tượng “Phim, ảnh về đối tượng nghiên cứu khí tượng thủy văn”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Phim, ảnh nghiên cứu về KTTV</b>

Mô tả	Là đối tượng thông tin dữ liệu về Phim, ảnh về đối tượng nghiên cứu khí tượng thủy văn
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Tên phim, ảnh về đối tượng nghiên cứu KTTV, Ngày thu thập, File phim, ảnh về đối tượng nghiên cứu khí tượng thủy văn, Ghi chú
Kiểu dữ liệu	Phi không gian

(3) Thông tin đối tượng “Kế hoạch và kết quả thực hiện tác động vào thời tiết”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Ke hoạch va ket qua thuc hien tac dong moi truong vao thoi tiet</b>
Mô tả	Là đối tượng thông tin dữ liệu về Kế hoạch và kết quả thực hiện tác động vào thời tiết
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Số kế hoạch thực hiện tác động vào thời tiết, Ngày kế hoạch thực hiện tác động vào thời tiết, Tên kế hoạch thực hiện tác động vào thời tiết, Tên cơ quan, tổ chức lập, Mục tiêu, Thời gian thực hiện, Địa điểm thực hiện, Dự toán thực hiện, Nguồn vốn, Sự cần thiết của kế hoạch tác động vào thời tiết, Mục đích tác động vào thời tiết, Khu vực dự kiến thực hiện tác động vào thời tiết, Thời gian dự kiến thực hiện tác động vào thời tiết,

	Hình thức tác động, Mô tả giải pháp tác động, Công cụ, phương tiện tác động, Hóa chất, thiết bị sử dụng để tác động, Các bước tiến hành tác động, File kế hoạch đã được phê duyệt, File Báo cáo kết quả thực hiện tác động vào thời tiết
Kiểu dữ liệu	Phi không gian

*d) Bộ dữ liệu văn bản quy phạm pháp luật về khí tượng thủy văn*

(1) Thông tin đối tượng “Văn bản quy phạm pháp luật”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>VB Quy phạm pháp luật</b>
Mô tả	Là đối tượng thông tin dữ liệu về văn bản quy phạm pháp luật
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Loại văn bản, Tên loại văn bản, Số của văn bản, Ký hiệu của văn bản, Ngày, tháng, năm văn bản, Tên cơ quan, tổ chức ban hành văn bản, Trích yếu nội dung, Ngôn ngữ, Số trang của văn bản, Ghi chú, Chức vụ, họ tên người ký, Chức vụ của người ký văn bản, Họ và tên người ký văn bản, Mức độ khẩn, Thời gian lưu trữ, Mức độ mật,

	File đính kèm
Kiểu dữ liệu	Phi không gian

*e) Bộ dữ liệu quan trắc phi truyền thống*

(1) Thông tin đối tượng “Radar thời tiết”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Radar thời tiết</b>
Mô tả	Là đối tượng thông tin dữ liệu về Radar thời tiết
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Thời gian, Ngày, Góc nâng, Góc hướng (độ), Phản hồi vô tuyến (dBZ), Vận tốc xuyên tâm (m/s), Độ rộng phổ (m/s), Độ lệch phản hồi 2 kênh phân cực (dB), Hệ số tương quan, Độ lệch pha 2 kênh phân cực (độ), Độ lệch pha riêng (độ/km)
Kiểu dữ liệu	Phi không gian

(2) Thông tin đối tượng “Định vị sét”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Định vị sét</b>
Mô tả	Là đối tượng thông tin dữ liệu về định vị sét
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Thời gian,

	Ngày, Giá trị dòng điện, Vĩnh độ, Kinh độ, Số sensor tham gia đo, Số DOF- Số giải pháp đo, Loại sét
Kiểu dữ liệu	Phi không gian

(3) Thông tin đối tượng “Vệ tinh khí tượng”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Ve tinh khi tuong</b>
Mô tả	Là đối tượng thông tin dữ liệu về vệ tinh khí tượng
Kiểu cơ sở	Table
Tên các thuộc tính	ID Thời gian, Ngày, Loại chỉ số, Giá trị
Kiểu dữ liệu	Phi không gian

g) Bộ dữ liệu về dự báo, cảnh báo khí tượng thủy văn

(1) Thông tin đối tượng “Bản tin dự báo, cảnh báo thời tiết”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Ban tin du bao, canh bao KTTV</b>
Mô tả	Là đối tượng thông tin dữ liệu về bản tin dự báo, cảnh báo thời tiết
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Số bản tin dự báo, cảnh báo thời tiết,

	Ngày ban hành, Người ban hành, Loại bản tin, Nội dung dự báo, cảnh báo thời tiết, Cấp độ cảnh báo, Phạm vi, khu vực cảnh báo, File đính kèm
Kiểu dữ liệu	Phi không gian

(2) Thông tin đối tượng “Bản tin dự báo, cảnh báo khí tượng khí hậu”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Bản tin dự báo, cảnh báo KTTV</b>
Mô tả	Là đối tượng thông tin dữ liệu về bản tin dự báo, cảnh báo khí hậu
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Số bản tin dự báo, cảnh báo khí hậu, Ngày ban hành, Người ban hành, Loại bản tin, Nội dung dự báo, cảnh báo khí hậu, Cấp độ cảnh báo, Phạm vi, khu vực cảnh báo, File đính kèm
Kiểu dữ liệu	Phi không gian

(3) Thông tin đối tượng “Bản tin dự báo, cảnh báo thủy văn”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Bản tin dự báo, cảnh báo KTTV</b>
Mô tả	Là đối tượng thông tin dữ liệu về bản tin dự báo, cảnh báo thủy văn

Kiểu cơ sở	Table
Tên các thuộc tính	ID, Số bản tin dự báo, cảnh báo thủy văn, Ngày ban hành, Người ban hành, Loại bản tin, Nội dung dự báo, cảnh báo thủy văn, Cấp độ cảnh báo, Phạm vi, khu vực cảnh báo, File đính kèm
Kiểu dữ liệu	Phi không gian

(4) Thông tin đối tượng “Bản tin dự báo, cảnh báo khí tượng hải văn”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Bản tin dự báo, cảnh báo KTTV</b>
Mô tả	Là đối tượng thông tin dữ liệu về bản tin dự báo, cảnh báo hải văn
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Số bản tin dự báo, cảnh báo hải văn, Ngày ban hành, Ngày ban hành, Loại bản tin, Nội dung dự báo dự báo, cảnh báo hải văn, Cấp độ cảnh báo, Phạm vi, khu vực cảnh báo, File đính kèm
Kiểu dữ liệu	Phi không gian

(5) Thông tin đối tượng “Bản tin dự báo, cảnh báo xu thế thời tiết”

<b>Kiểu đối tượng:</b>	
------------------------	--

<b>Tên</b>	<b>Bản tin du báo, cảnh báo KTTV</b>
Mô tả	Là đối tượng thông tin dữ liệu về bản tin dự báo, cảnh báo xu thế thời tiết
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Số bản tin, Ngày ban hành, Ngày ban hành, Loại bản tin, Nội dung dự báo, Cấp độ cảnh báo, Phạm vi, khu vực cảnh báo, File đính kèm
Kiểu dữ liệu	Phi không gian

(6) Thông tin đối tượng “Sản phẩm của các mô hình dự báo”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Sản phẩm mô hình du báo</b>
Mô tả	Là đối tượng thông tin dữ liệu về Sản phẩm của các mô hình dự báo
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Tên mô hình dự báo, Loại mô hình dự báo, Tên sản phẩm mô hình dự báo, Thông tin mô tả mô hình, sản phẩm dự báo
Kiểu dữ liệu	Phi không gian

(7) Thông tin đối tượng “Dữ liệu mô hình dự báo”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Dữ liệu mô hình du báo</b>

Mô tả	Là đối tượng thông tin dữ liệu về Dữ liệu mô hình dự báo
Kiểu cơ sở	Table
Tên các thuộc tính	ID, ID mô hình dự báo, Ngày, Thời gian gửi dữ liệu, Giá trị
Kiểu dữ liệu	Phi không gian

(8) Thông tin đối tượng “Biểu đồ, bản đồ, ảnh thu từ vệ tinh”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Bieu do, ban do, anh ve tinh</b>
Mô tả	Là đối tượng thông tin dữ liệu về biểu đồ, bản đồ, ảnh thu từ vệ tinh
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Thời gian thu thập, Ngày thu thập, Hình ảnh thu thập, File đính kèm, Mô tả
Kiểu dữ liệu	Phi không gian

II.1.3.5. Danh mục dữ liệu

(1) Thông tin danh mục dữ liệu “Loại trạm”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Loai tram</b>
Mô tả	Là danh sách các loại trạm
Kiểu cơ sở	Table
Tên các thuộc tính	ID,

	Mã loại trạm, Tên loại trạm, Mô tả
Kiểu dữ liệu	Phi không gian

(2) Thông tin danh mục dữ liệu “Hướng gió”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Huong gio</b>
Mô tả	Là danh sách các hướng gió
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Kí hiệu hướng gió, Tên hướng gió, Mô tả
Kiểu dữ liệu	Phi không gian

(3) Thông tin danh mục dữ liệu “Dạng mây”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Dang may</b>
Mô tả	Là danh sách các Dạng mây
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Kí hiệu dạng mây, Tên dạng mây, Mô tả
Kiểu dữ liệu	Phi không gian

(4) Thông tin danh mục dữ liệu “Sông/Lưu vực sông/Tiêu lưu vực”

<b>Kiểu đối tượng:</b>
------------------------

<b>Tên</b>	<b>Sông/Luu vực sông/Tiểu lưu vực</b>
Mô tả	Là danh sách các Sông/Lưu vực sông/Tiểu lưu vực
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Phân loại, Ký hiệu, Tên danh mục sông/lưu vực/tiểu lưu vực
Kiểu dữ liệu	Phi không gian

(5) Thông tin danh mục dữ liệu “Loại sóng”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Loại sóng</b>
Mô tả	Là danh sách các loại sóng
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Kí hiệu loại sóng biển, Tên loại sóng biển, Mô tả
Kiểu dữ liệu	Phi không gian

(6) Thông tin danh mục dữ liệu “Loại văn bản”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Loại song</b>
Mô tả	Là danh sách các loại văn bản
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Kí hiệu loại văn bản, Tên loại văn bản, Mô tả
Kiểu dữ liệu	Phi không gian

(7) Thông tin danh mục dữ liệu “Loại sét”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Loai song</b>
Mô tả	Là danh sách các loại sét
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Kí hiệu loại sét, Tên loại sét, Mô tả
Kiểu dữ liệu	Phi không gian

(8) Thông tin danh mục dữ liệu “Loại bản tin”

<b>Kiểu đối tượng:</b>	
<b>Tên</b>	<b>Loai bản tin</b>
Mô tả	Là danh sách các loại bản tin
Kiểu cơ sở	Table
Tên các thuộc tính	ID, Kí hiệu loại bản tin, Tên loại bản tin, Mô tả
Kiểu dữ liệu	Phi không gian

*II.1.3.6. Nhập dữ liệu*

Thực hiện tạo lập danh mục dữ liệu cho CSDL bao gồm tạo lập nội dung cho danh mục dữ liệu và tạo lập nội dung cho siêu dữ liệu. Các loại dữ liệu (đối tượng) cần tạo lập như sau:

- 1) Loại trạm;
- 2) Đơn vị đo;
- 3) Hướng gió;

- 4) Dạng mây;
- 5) Sông/Lưu vực sông/Tiểu lưu vực;
- 6) Loại sóng.
- 7) Loại văn bản
- 8) Loại sét
- 9) Loại bản tin

Phương thức nhập liệu bằng cách nhập tự động hoặc nhập thủ công.

#### *II.1.3.7. Biên tập dữ liệu*

Thực hiện biên tập dữ liệu cho CSDL bao gồm tuyên bố đối tượng cho dữ liệu. Các loại dữ liệu (đối tượng) trong nội dung tuyên bố như sau:

- 1) Khí tượng bề mặt
- 2) Khí tượng nông nghiệp
- 3) Thủy văn
- 4) Hải văn
- 5) Thám không vô tuyến
- 6) Gió trên cao (Pilot)
- 7) Bức xạ
- 8) Ô dôn, bức xạ cực tím
- 9) Mưa
- 10) Môi trường không khí
- 11) Môi trường nước sông
- 12) Môi trường nước hồ
- 13) Môi trường nước biển ven bờ
- 14) Điều tra, khảo sát thủy văn
- 15) Điều tra, khảo sát khí tượng biển
- 16) Điều tra, khảo sát Hải văn
- 17) Điều tra khảo sát về môi trường nước biển
- 18) Quan trắc, điều tra, khảo sát
- 19) Đơn vị thực hiện quan trắc, điều tra, khảo sát
- 20) Bản quyền thông tin khảo sát
- 21) Tài liệu thuyết minh
- 22) Hồ sơ kỹ thuật của các trạm, công trình, phương tiện đo khí tượng thủy văn
- 23) Thiết bị quan trắc
- 24) Bảo dưỡng, sửa chữa, thay thế linh kiện, thiết bị quan trắc

- 25) Tổ chức/cá nhân được cấp phép hoạt động dự báo, cảnh báo khí tượng thủy văn
- 26) Hồ sơ cấp phép dự báo, cảnh báo
- 27) Chương trình, đề án, dự án về khí tượng thủy văn, giám sát biến đổi khí hậu
- 28) Phim, ảnh về đối tượng nghiên cứu khí tượng thủy văn
- 29) Kế hoạch và kết quả thực hiện tác động vào thời tiết
- 30) Văn bản quy phạm pháp luật
- 31) Radar thời tiết
- 32) Định vị sét
- 33) Vệ tinh khí tượng
- 34) Bản tin dự báo, cảnh báo thời tiết
- 35) Bản tin dự báo, cảnh báo khí hậu
- 36) Bản tin dự báo, cảnh báo thủy văn
- 37) Bản tin dự báo, cảnh báo hải văn
- 38) Bản tin dự báo, cảnh báo xu thế thời tiết
- 39) Sản phẩm của các mô hình dự báo
- 40) Dữ liệu mô hình dự báo
- 41) Biểu đồ, bản đồ, ảnh thu từ vệ tinh

#### II.1.3.8. Yêu cầu khối lượng công việc

STT	Nội dung	Đơn vị tính	Khó khăn	Khối lượng
<b>I</b>	<b>RÀ SOÁT VÀ PHÂN TÍCH NỘI DUNG THÔNG TIN DỮ LIỆU</b>			
<b>1</b>	<b>Rà soát, phân loại các thông tin dữ liệu cần đưa vào cơ sở dữ liệu</b>			
1.1	Rà soát, đánh giá và phân loại chi tiết dữ liệu đã được chuẩn hóa và chưa được chuẩn hóa	Bộ dữ liệu	1-3	1
1.2	Chuẩn bị dữ liệu mẫu	Bộ dữ liệu	1-3	1
<b>2</b>	<b>Phân tích nội dung thông tin dữ liệu</b>			
2.1	Xác định danh mục các đối tượng quản lý	ĐTQL	3	17,4
2.2	Xác định chi tiết các thông tin cho từng đối tượng quản lý	ĐTQL	3	17,4
2.3	Xác định chi tiết các ràng buộc giữa các đối tượng quản lý	ĐTQL	3	17,4
2.4	Xác định các yếu tố ảnh hưởng đến việc xây dựng cơ sở dữ liệu	CSDL	1-3	1

2.5	Xác định chi tiết các tài liệu quét (tài liệu đính kèm) và các tài liệu dạng giấy cần nhập vào cơ sở dữ liệu từ bàn phím	Bộ dữ liệu	3	1
2.6	Xác định khung danh mục dữ liệu, siêu dữ liệu sử dụng trong cơ sở dữ liệu	CSDL	3	1
2.7	Quy đổi đối tượng quản lý	ĐTQL	1-3	17,4
<b>II</b>	<b>THIẾT KẾ MÔ HÌNH CƠ SỞ DỮ LIỆU</b>			
1	Thiết kế mô hình danh mục dữ liệu, siêu dữ liệu.	CSDL	3	1
2	<b>Thiết kế mô hình cơ sở dữ liệu</b>	<b>ĐTQL</b>	<b>3</b>	<b>17,4</b>
3	Nhập dữ liệu mẫu để kiểm tra mô hình dữ liệu	ĐTQL	3	17,4
<b>III</b>	<b>TẠO LẬP DỮ LIỆU CHO DANH MỤC DỮ LIỆU, SIÊU DỮ LIỆU</b>			
1	Tạo lập nội dung cho danh mục dữ liệu	ĐTQL	1-3	17,4
2	Tạo lập nội dung cho siêu dữ liệu	ĐTQL	1-3	17,4
<b>V</b>	<b>BIÊN TẬP DỮ LIỆU</b>			
1	Tuyên bố đối tượng	ĐTQL	2	17,4
<b>VI</b>	<b>KIỂM TRA SẢN PHẨM</b>			
1	Kiểm tra mô hình dữ liệu	ĐTQL	1	17,4
2	Kiểm tra nội dung CSDL	ĐTQL	1	17,4
3	Kiểm tra danh mục dữ liệu, siêu dữ liệu	ĐTQL	1	17,4
<b>VII</b>	<b>PHỤC VỤ NGHIỆM THU VÀ GIAO NỘP SẢN PHẨM</b>			
1	Lập báo cáo tổng kết kỹ thuật và phục vụ nghiệm thu sản phẩm	ĐTQL	1-3	17,4
2	Đóng gói các sản phẩm dạng giấy và dạng số	ĐTQL	1-3	17,4
3	Giao nộp sản phẩm	CSDL	1-3	1

#### ***II.1.4. Dịch vụ chia sẻ dữ liệu KTTV Quốc gia***

Việc quản lý kết nối và chia sẻ dữ liệu KTTV quốc gia được thực hiện theo quy định và hướng dẫn tại Điều 5 Nghị định số 47/2020/NĐ-CP. Cụ thể như sau:

- Dữ liệu hình thành trong hoạt động của cơ quan nhà nước được chia sẻ phục vụ các hoạt động của cơ quan nhà nước hướng tới phục vụ người dân, doanh nghiệp tuân thủ quy định của pháp luật trong việc tạo lập, quản lý và sử dụng dữ liệu.

- Cơ quan nhà nước phải có trách nhiệm chia sẻ dữ liệu với cơ quan, tổ chức, cá nhân khác theo quy định của pháp luật; không cung cấp thông tin qua hình thức văn bản đối với thông tin đã được khai thác qua hình thức kết nối, chia sẻ dữ liệu giữa các hệ thống thông tin; không được thu thập, tổ chức thu thập lại

dữ liệu hoặc yêu cầu người dân, doanh nghiệp cung cấp thông tin, dữ liệu trong quá trình giải quyết các thủ tục hành chính nếu dữ liệu này đã được cơ quan nhà nước khác cung cấp, sẵn sàng cung cấp thông qua kết nối, chia sẻ dữ liệu, trừ trường hợp dữ liệu đó không bảo đảm yêu cầu về chất lượng theo tiêu chuẩn, quy chuẩn chuyên ngành hoặc pháp luật có quy định khác.

- Việc chia sẻ dữ liệu giữa các cơ quan nhà nước không làm ảnh hưởng tới quyền lợi và trách nhiệm của tổ chức, cá nhân có liên quan, không được xâm phạm quyền về đời sống riêng tư, bí mật cá nhân, bí mật gia đình trừ trường hợp pháp luật có quy định khác.

- Dữ liệu chia sẻ phải bảo đảm tính cập nhật, chính xác theo quy định của pháp luật.

- Dữ liệu chia sẻ giữa các cơ quan nhà nước là không thu phí, trừ trường hợp khai thác, sử dụng dữ liệu thuộc danh mục được quy định trong Luật Phí và lệ phí.

- Dữ liệu của cơ quan nhà nước chia sẻ cho tổ chức, cá nhân thực hiện theo các nguyên tắc sau:

+ Tổ chức, cá nhân được quyền khai thác dữ liệu của mình hoặc dữ liệu của tổ chức, cá nhân khác khi được tổ chức, cá nhân đó chấp nhận, trừ trường hợp pháp luật có quy định khác;

+ Các trường hợp ngoài quy định tại điểm a khoản này, dữ liệu của cơ quan nhà nước được chia sẻ cho tổ chức, cá nhân được thực hiện theo quy định của Luật Tiếp cận thông tin và các quy định của pháp luật hiện hành.

Kết nối, chia sẻ dữ liệu KTTV quốc gia theo hình thức mặc định, bao gồm các dữ liệu được quy định tại các Điều 6, Điều 7, Điều 8 của Nghị định này, trừ trường hợp khai thác, sử dụng dữ liệu thuộc danh mục được quy định trong Luật Phí và lệ phí.

Kết nối, chia sẻ dữ liệu KTTV quốc gia theo yêu cầu đặc thù, tuân thủ quy định tại Nghị định số 47/2020/NĐ-CP.

Phương thức kết nối, chia sẻ dữ liệu KTTV quốc gia:

- Sử dụng dịch vụ dữ liệu (API/Web service) làm phương thức để kết nối, chia sẻ dữ liệu giữa Cơ sở dữ liệu KTTV quốc gia với các hệ thống thông tin khác.

- Cấu trúc gói tin sử dụng định dạng JSON (JavaScript Object Notation).

#### *II.1.4.1. Quy trình thực hiện*

Dịch vụ chia sẻ dữ liệu KTTV Quốc gia xây dựng tuân thủ theo thông tư 14/2020/TT-BTNMT về ban hành quy trình và định mức kinh tế - kỹ thuật xây dựng, duy trì, vận hành hệ thống thông tin ngành tài nguyên và môi trường, gồm các bước như sau:

##### ***- Bước 1: Xác định yêu cầu***

- ✓ Thu thập yêu cầu phần mềm
- ✓ Xác định yêu cầu chức năng
- ✓ Đặc tả dữ liệu
- ✓ Xác định yêu cầu phi chức năng

##### ***- Bước 2: Phân tích và thiết kế***

- ✓ Xác định lại các quy trình nghiệp vụ được tin học hóa
- ✓ Xác định danh sách chức năng hệ thống
- ✓ Xác định các yêu cầu về thông tin dữ liệu
- ✓ Xác định các yêu cầu về giao diện của phần mềm
- ✓ Xác định các yêu cầu phi chức năng của phần mềm
- ✓ Thiết kế kiến trúc phần mềm
- ✓ Thiết kế biểu đồ trường hợp sử dụng
- ✓ Thiết kế biểu đồ tuần tự
- ✓ Thiết kế biểu đồ lớp
- ✓ Thiết kế mô hình cơ sở dữ liệu
- ✓ Thiết kế giao diện phần mềm

##### ***- Bước 3: Lập trình, viết mã nguồn***

- ✓ Viết mã nguồn
- ✓ Tích hợp mã nguồn

##### ***- Bước 4: Kiểm tra, kiểm thử***

- ✓ Kiểm tra mã nguồn theo quy tắc lập trình
- ✓ Kiểm tra mức thành phần
- ✓ Kiểm tra mức hệ thống

##### ***- Bước 5: Hoàn thiện, đóng gói sản phẩm***

- ✓ Viết tài liệu mô tả giới thiệu phần mềm
- ✓ Viết tài liệu hướng dẫn cài đặt phần mềm
- ✓ Xây dựng tài liệu hướng dẫn sử dụng phần mềm

✓ Đóng gói phần mềm

**- Bước 6: Cài đặt, chuyển giao, hướng dẫn sử dụng**

- ✓ Cài đặt, chuyển giao, hướng dẫn sử dụng
- ✓ Đào tạo, hướng dẫn người dùng sử dụng phần mềm
- ✓ Bàn giao tài liệu hướng dẫn cài đặt và sử dụng phần mềm

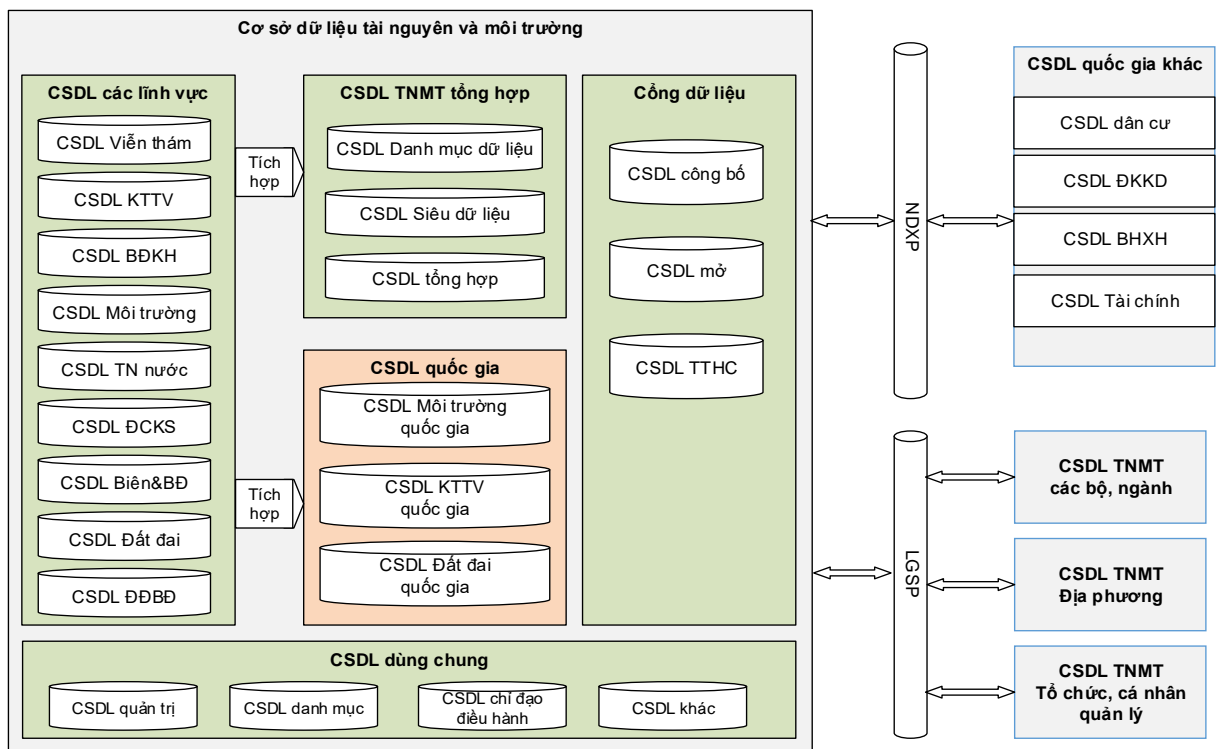
**- Bước 7: Bảo trì, bảo hành phần mềm**

- ✓ Bảo trì, bảo hành phần mềm
- ✓ Phát hành các bản vá lỗi
- ✓ Xử lý sự cố liên quan đến dữ liệu

**II.1.4.2. Mô hình kiến trúc dịch vụ chia sẻ dữ liệu KTTV Quốc gia**

**a) Mô hình kiến trúc hệ thống CSDL của Bộ Tài nguyên và Môi trường**

Bộ Tài nguyên và Môi trường đã định hướng mô hình tổng thể tổ chức CSDL như sau:



*CSDL Khí tượng thủy văn quốc gia trong mô hình tổng thể tổ chức dữ liệu tài nguyên và môi trường.*

Mô hình thể hiện các thông tin dữ liệu thuộc phạm vi quản lý của Bộ TN&MT chi tiết đến từng lĩnh vực và không phụ thuộc vào cơ quan, đơn vị tạo lập. Dữ liệu ngành tài nguyên và môi trường được thiết kế tối đa tính chia sẻ,

không thu thập xây dựng trùng lặp và phù hợp, đáp ứng được các yêu cầu nghiệp vụ.

Nhóm CSDL tài nguyên và môi trường do Bộ Tài nguyên và Môi trường quản lý, bao gồm 04 thành phần:

- Thành phần CSDL chuyên ngành tài nguyên và môi trường: Các CSDL lĩnh vực đo đạc và bản đồ, đất đai, địa chất và khoáng sản, biển và hải đảo, tài nguyên nước, môi trường, khí tượng thủy văn, biến đổi khí hậu, viễn thám; Các CSDL đặc thù, đa lĩnh vực: CSDL quan trắc về tài nguyên và môi trường, CSDL quy hoạch về tài nguyên và môi trường...

- Thành phần CSDL dùng chung bao gồm CSDL danh mục dùng chung, CSDL báo cáo tổng hợp, CSDL văn bản quy phạm pháp luật...

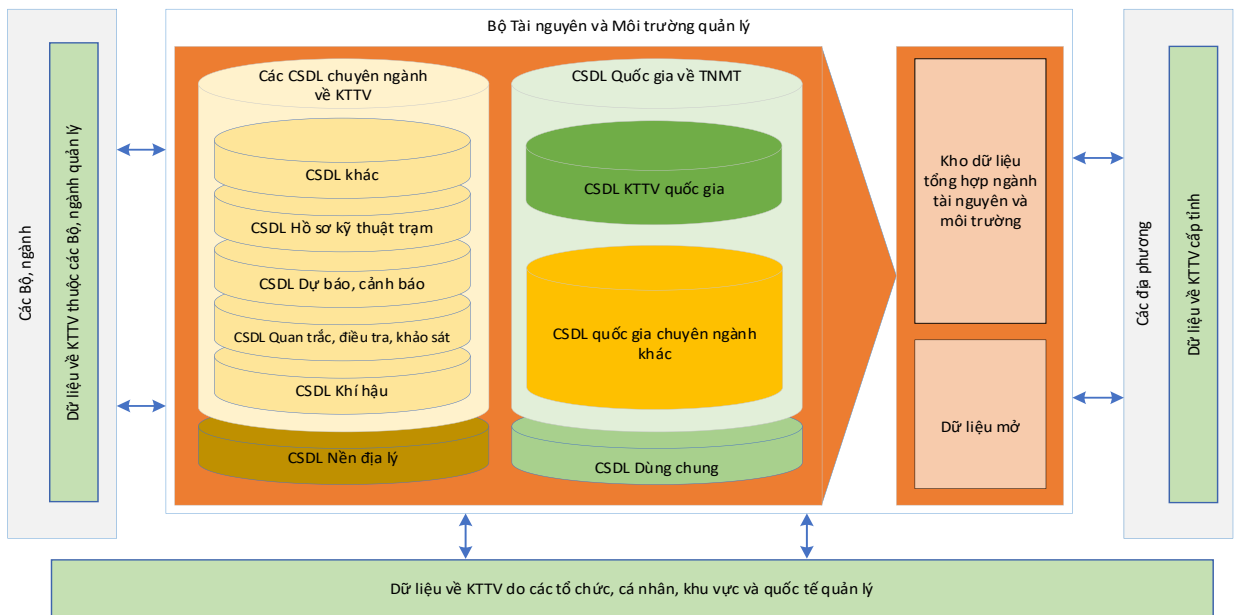
- Thành phần dữ liệu mở ngành tài nguyên và môi trường.

+ Nhóm CSDL tài nguyên và môi trường do các bộ, ngành, địa phương quản lý.

+ Nhóm CSDL tài nguyên và môi trường do tổ chức, cá nhân, khu vực, quốc tế quản lý.

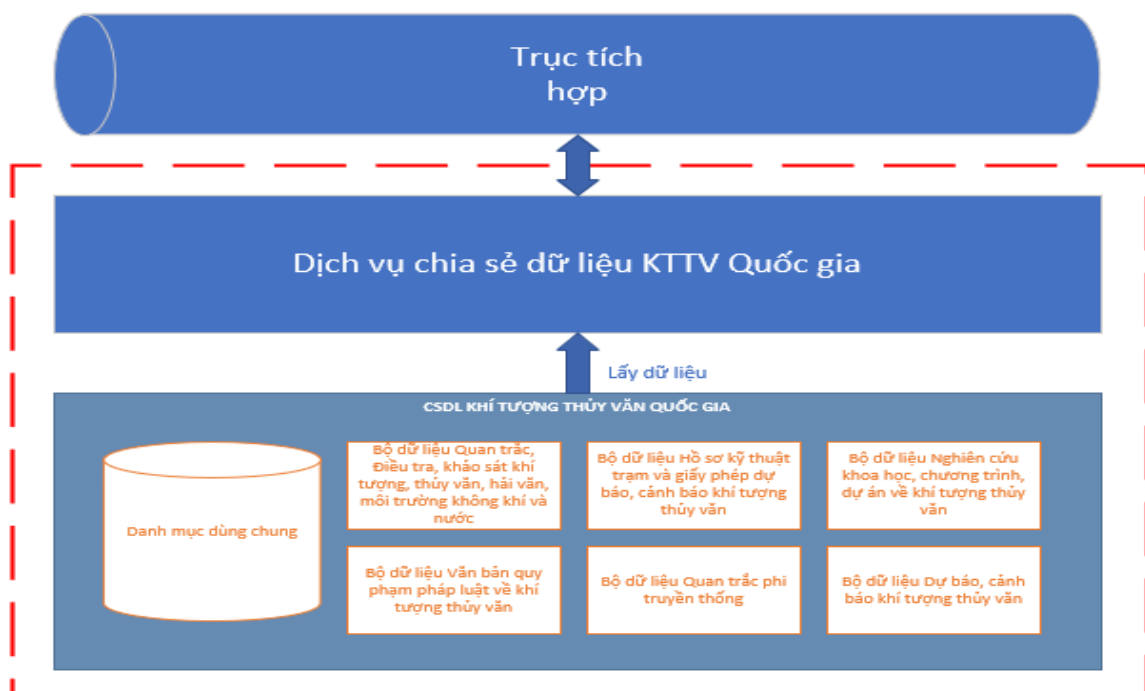
+ Thông tin, dữ liệu thu thập, thu nhận trực tiếp từ các thiết bị, email, mạng xã hội...

- Thành phần CSDL tài nguyên và môi trường quốc gia.



Mô hình các CSDL KTTV Quốc gia trong mô hình tổng thể CSDL tài nguyên và môi trường do Bộ TNMT quản lý

### b) Mô hình kiến trúc dịch vụ chia sẻ dữ liệu KTTV Quốc gia



Mô hình kiến trúc dịch vụ chia sẻ dữ liệu KTTV Quốc gia

Kiến trúc dịch vụ chia sẻ dữ liệu KTTV Quốc gia gồm 2 thành phần chính bao gồm:

- Các CSDL thành phần của CSDL KTTV Quốc gia: Bộ dữ liệu Quan trắc, Điều tra, khảo sát khí tượng, thủy văn, hải văn, môi trường không khí và nước;

Bộ dữ liệu Hồ sơ kỹ thuật trạm và giấy phép dự báo, cảnh báo khí tượng thủy văn; Bộ dữ liệu Nghiên cứu khoa học, chương trình, dự án về khí tượng thủy văn; Bộ dữ liệu Văn bản quy phạm pháp luật về khí tượng thủy văn; Bộ dữ liệu Quan trắc phi truyền thống; Bộ dữ liệu Dự báo, cảnh báo khí tượng thủy văn.

- Các dịch vụ chia sẻ dữ liệu KTTV Quốc gia (API/Web Service) sẵn sàng chia sẻ dữ liệu khi có yêu cầu.

#### II.1.4.3. Danh sách các dịch vụ chia sẻ dữ liệu KTTV Quốc gia

STT	Tên dịch vụ chia sẻ dữ liệu
<b>I</b>	<b>Dịch vụ chia sẻ Bộ dữ liệu Quan trắc, Điều tra, khảo sát khí tượng, thủy văn, hải văn, môi trường không khí và nước</b>
1	Dịch vụ chia sẻ thông tin dữ liệu khí tượng bề mặt
2	Dịch vụ chia sẻ thông tin dữ liệu khí tượng nông nghiệp
3	Dịch vụ chia sẻ thông tin dữ liệu thủy văn
4	Dịch vụ chia sẻ thông tin dữ liệu hải văn
5	Dịch vụ chia sẻ thông tin dữ liệu thám không vô tuyến
6	Dịch vụ chia sẻ thông tin dữ liệu gió trên cao (pilot)
7	Dịch vụ chia sẻ thông tin dữ liệu bức xạ
8	Dịch vụ chia sẻ thông tin dữ liệu ô dôn, bức xạ cực tím
9	Dịch vụ chia sẻ thông tin dữ liệu mưa
10	Dịch vụ chia sẻ thông tin dữ liệu môi trường không khí
11	Dịch vụ chia sẻ thông tin dữ liệu môi trường nước sông
12	Dịch vụ chia sẻ thông tin dữ liệu môi trường nước hồ
13	Dịch vụ chia sẻ thông tin dữ liệu môi trường nước biển ven bờ
14	Dịch vụ chia sẻ thông tin dữ liệu điều tra, khảo sát thủy văn
15	Dịch vụ chia sẻ thông tin dữ liệu điều tra, khảo sát khí tượng biển
16	Dịch vụ chia sẻ thông tin dữ liệu điều tra, khảo sát hải văn
17	Dịch vụ chia sẻ thông tin dữ liệu điều tra khảo sát về môi trường nước biển
18	Dịch vụ chia sẻ thông tin dữ liệu quan trắc, điều tra, khảo sát
19	Dịch vụ chia sẻ thông tin dữ liệu đơn vị thực hiện quan trắc, điều tra, khảo sát
20	Dịch vụ chia sẻ thông tin dữ liệu bản quyền thông tin khảo sát
21	Dịch vụ chia sẻ thông tin dữ liệu tài liệu thuyết minh
<b>II</b>	<b>Dịch vụ chia sẻ Bộ dữ liệu Hồ sơ kỹ thuật trạm và giấy phép dự báo, cảnh báo khí tượng thủy văn</b>

<b>STT</b>	<b>Tên dịch vụ chia sẻ dữ liệu</b>
22	Dịch vụ chia sẻ thông tin dữ liệu hồ sơ kỹ thuật của các trạm, công trình, phương tiện đo khí tượng thủy văn
23	Dịch vụ chia sẻ thông tin dữ liệu thiết bị quan trắc
24	Dịch vụ chia sẻ thông tin dữ liệu bảo dưỡng, sửa chữa, thay thế linh kiện, thiết bị quan trắc
25	Dịch vụ chia sẻ thông tin dữ liệu tổ chức/cá nhân được cấp phép hoạt động dự báo, cảnh báo khí tượng thủy văn
26	Dịch vụ chia sẻ thông tin dữ liệu hồ sơ cấp phép dự báo, cảnh báo
<b>III</b>	<b>Dịch vụ chia sẻ Bộ dữ liệu Nghiên cứu khoa học, chương trình, dự án về khí tượng thủy văn</b>
27	Dịch vụ chia sẻ thông tin dữ liệu chương trình, đề án, dự án về khí tượng thủy văn, giám sát biến đổi khí hậu
28	Dịch vụ chia sẻ thông tin dữ liệu phim, ảnh về đối tượng nghiên cứu khí tượng thủy văn
29	Dịch vụ chia sẻ thông tin dữ liệu kế hoạch và kết quả thực hiện tác động vào thời tiết
<b>IV</b>	<b>Dịch vụ chia sẻ Bộ dữ liệu Văn bản quy phạm pháp luật về khí tượng thủy văn</b>
30	Dịch vụ chia sẻ thông tin dữ liệu văn bản quy phạm pháp luật
<b>V</b>	<b>Dịch vụ chia sẻ Bộ dữ liệu Quan trắc phi truyền thống</b>
31	Dịch vụ chia sẻ thông tin dữ liệu quan trắc radar thời tiết
32	Dịch vụ chia sẻ thông tin dữ liệu quan trắc định vị sét
<b>V</b>	<b>Dịch vụ chia sẻ Bộ dữ liệu Dự báo, cảnh báo khí tượng thủy văn</b>
33	Dịch vụ chia sẻ thông tin dữ liệu quan trắc bản tin dự báo, cảnh báo khí tượng thủy văn
34	Dịch vụ chia sẻ thông tin dữ liệu quan trắc sản phẩm của các mô hình dự báo
35	Dịch vụ chia sẻ thông tin dữ liệu quan trắc dữ liệu mô hình dự báo
36	Dịch vụ chia sẻ thông tin dữ liệu quan trắc biểu đồ, bản đồ, ảnh thu từ vệ tinh

#### II.1.4.4. Các yêu cầu phi chức năng

##### a) Yêu cầu về quản lý, khai thác

Căn cứ theo Điều 5 Nghị định số 73/2017/NĐ-CP. Việc thu thập, quản lý, khai thác và sử dụng thông tin, dữ liệu tài nguyên và môi trường phải tuân theo các nguyên tắc sau đây:

1. Bảo đảm tính chính xác, trung thực, khoa học, khách quan và kế thừa.
2. Tập hợp, quản lý được đầy đủ các nguồn thông tin, dữ liệu của các cơ quan, tổ chức, cá nhân.
3. Lưu trữ, bảo quản đáp ứng yêu cầu sử dụng lâu dài, thuận tiện.
4. Tổ chức có hệ thống phục vụ thuận tiện khai thác và sử dụng, phát huy hiệu quả, tiết kiệm chung cho xã hội, nâng cao năng lực cạnh tranh quốc gia.
5. Được công bố công khai và được quyền tiếp cận theo quy định của Luật tiếp cận thông tin trừ các thông tin, dữ liệu thuộc về bí mật nhà nước và hạn chế sử dụng.
6. Phục vụ kịp thời công tác quản lý nhà nước, đáp ứng yêu cầu của phát triển kinh tế - xã hội, bảo đảm quốc phòng, an ninh.
7. Sử dụng đúng mục đích, phải trích dẫn nguồn và tuân thủ các quy định của pháp luật về bảo vệ bí mật nhà nước, sở hữu trí tuệ.
8. Khai thác và sử dụng thông tin, dữ liệu phải thực hiện nghĩa vụ tài chính theo quy định của pháp luật.

#### **b) Yêu cầu về kỹ thuật công nghệ xây dựng dịch vụ chia sẻ dữ liệu**

- Mô hình ứng dụng: Web-based Application
- Công cụ phát triển ứng dụng: Microsoft Visual Studio phiên bản 2015 trở lên
- Công nghệ nền tảng: .NET Framework 4.5
- Ngôn ngữ phát triển ứng dụng: C#, ASP.NET, HTML, XML
- Hệ điều hành máy chủ: Windows Server 2012 trở lên
- Hệ quản trị cơ sở dữ liệu: MS SQL Server 2014 trở lên
- Web/Application Server: IIS 7.5 trở lên
- Môi trường mạng truy cập: mạng LAN/ WAN/ Internet

#### **c) Yêu cầu cần đáp ứng đối với cơ sở dữ liệu**

Thông tin, dữ liệu được thiết kế và lưu trữ trên hệ quản trị cơ sở dữ liệu Microsoft SQL Server/Oracle. Các bảng dữ liệu được thiết kế với các liên kết và ràng buộc đảm bảo tính toàn vẹn của dữ liệu.

Việc truy xuất dữ liệu có thể thực hiện trực tiếp trên các bảng hoặc thông qua các View để tăng tính linh hoạt của hệ thống.

#### d) Yêu cầu an toàn thông tin

Đảm bảo tuân thủ các tiêu chuẩn về an ninh bảo mật, an toàn dữ liệu đối với hệ thống. Tuân thủ các tiêu chuẩn bảo mật về mức hệ thống, mức ứng dụng, mức CSDL. Đồng thời đảm bảo an toàn hệ thống thông tin theo cấp độ.

Phương án đảm bảo an toàn hệ thống thông tin, giải pháp ứng cứu sự cố an toàn thông tin phải phù hợp, tuân thủ các quy định hiện hành.

- Bảo mật mức hệ thống:

+ Thiết lập tường lửa để quản lý việc truy cập Internet cũng như bảo vệ các máy tính khỏi sự tấn công của virus, hacker.

+ Thiết lập an ninh cho các máy chủ Web: nâng cấp và cấu hình bảo mật cho phần mềm web server, tăng tính bảo mật cho hệ điều hành máy chủ qua việc quản lý cập nhật bản sửa lỗi.

- Bảo mật mức ứng dụng:

+ Mã nguồn hệ thống có tính bảo mật cao, tuân thủ các tiêu chuẩn quốc tế về các khuyến nghị bảo mật thường gặp (sql injection, XSS ...)

- Bảo mật mức CSDL:

+ Trao đổi thông tin, dữ liệu theo các chuẩn về an toàn thông tin SSL v3.0, HTTPS,...

+ Có cơ chế theo dõi và giám sát, lưu vết tất cả các hoạt động cho mỗi kênh thông tin và toàn hệ thống.

+ Toàn bộ các dữ liệu cần quản lý, phải được lưu trong CSDL được mã hóa và phân quyền truy cập chặt chẽ.

#### e) Các ràng buộc đối với dịch vụ chia sẻ dữ liệu gồm: ràng buộc môi trường, sự phụ thuộc vào hệ thống nền tảng

Dịch vụ chia sẻ dữ liệu cần đảm bảo các yêu cầu cơ bản sau:

- Không có yêu cầu đặc biệt về cấu hình máy chủ cài đặt hệ thống.

- Có khả năng cài đặt được và hoạt động ổn định ngay cả trên các máy chủ vật lý cũng như máy chủ ảo hóa.

#### **g) Yêu cầu về kiểm thử**

Các dịch vụ chia sẻ dữ liệu (API/Web Service) KTTV Quốc gia sẽ được thực hiện kiểm thử tại ít nhất một đơn vị thụ hưởng theo hình thức tự thực hiện. Việc kiểm thử phải đảm bảo tuân thủ nội dung và trình tự công tác thực hiện theo đúng quy định về quản lý đầu tư ứng dụng công nghệ thông tin sử dụng nguồn vốn ngân sách nhà nước.

Trước khi triển khai vận hành chính thức hệ thống, đơn vị phát triển hệ thống phải phối hợp với Chủ đầu tư để lập kế hoạch và thực hiện kiểm thử chất lượng, vận hành thử nghiệm hệ thống, đảm bảo hệ thống sau khi bàn giao sẽ hoạt động an toàn, ổn định và hiệu quả theo đúng nhu cầu người dùng. Việc kiểm thử phải bao gồm các hạng mục sau:

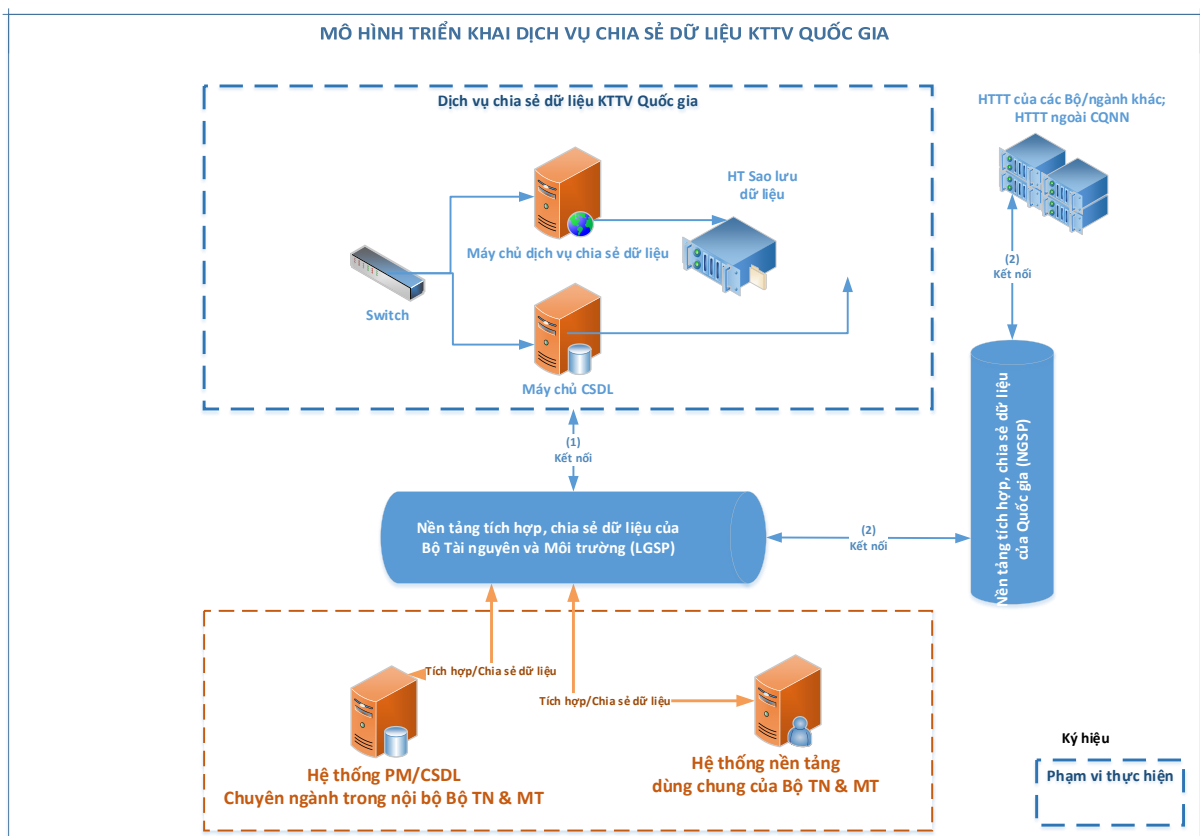
- Kiểm thử chức năng;
- Kiểm thử bảo mật;
- Kiểm thử hiệu năng.

#### *II.1.4.5. Cài đặt, chuyển giao, hướng dẫn sử dụng*

Dịch vụ chia sẻ dữ liệu (API/webservice) KTTV Quốc gia sau khi xây dựng xong sẽ được triển khai trên nền tảng hạ tầng của BigData.

Dịch vụ chia sẻ dữ liệu (API/Web Service) KTTV Quốc gia có thể giao tiếp với các hệ thống phần mềm, ứng dụng tích hợp thông qua mạng LAN/WAN nội bộ trong ngành TN&MT và Trục tích hợp dữ liệu LGSP của ngành tài nguyên.

Mô hình triển khai hệ thống được đề xuất như sau:



*Mô hình triển khai dịch vụ chia sẻ dữ liệu KTTV Quốc Gia*

**Máy chủ dịch vụ chia sẻ dữ liệu:** Là máy chủ có nhiệm vụ xử lý các yêu cầu đến từ người dùng, thực hiện truy vấn các dữ liệu từ vùng máy chủ CSDL và trả lại kết quả cho người dùng.

**Máy chủ CSDL:** Là máy chủ có nhiệm vụ tổ chức, lưu trữ dữ liệu của toàn hệ thống và xử lý các yêu cầu truy vấn, khai thác dữ liệu.

**Thiết bị sao lưu/phục hồi dữ liệu:** Là thiết bị có nhiệm vụ cung cấp dịch vụ sao lưu hàng ngày đối với dữ liệu trong máy chủ ứng dụng và máy chủ CSDL. Đảm bảo toàn bộ source code hệ thống phần mềm, dữ liệu trong CSDL đều được sao lưu, tránh mất mát thông tin khi gặp sự cố và khôi phục hệ thống trong thời gian sớm nhất nhằm đảm bảo tính thông suốt, sẵn sàng của phần mềm trong quá trình khai thác, sử dụng.

**Switch:** Thiết bị kết nối mạng trong hệ thống trung tâm dữ liệu, giúp kết nối giữa các máy chủ sử dụng và thiết bị mạng trong trung tâm dữ liệu.

**Firewall:** Thiết bị tường lửa triển khai đảm bảo tính an toàn thông tin thiết bị, hạn chế các mã độc, virus xâm nhập hệ thống phần mềm thông qua đường truyền internet.

Việc đảm bảo an toàn thông tin cho hệ thống được thực hiện thông qua các giải pháp về kỹ thuật và chính sách. Trong đó, các giải pháp kỹ thuật (tường lửa, IPS, chống spam, ...) là các giải pháp được trang bị tại Trung tâm dữ liệu. Để đảm bảo hệ thống được vận hành liên tục, an toàn về thông tin dữ liệu, về hạ tầng phần cứng, hệ thống mạng cần được bố trí đáp ứng các yêu cầu sau, tuy nhiên sẽ được yêu cầu nâng cấp khi có nhu cầu khai thác cao hơn trong các giai đoạn mở rộng tiếp theo:

<b>TT</b>	<b>Tên thiết bị</b>	<b>Số lượng</b>
<b>I</b>	<b>Máy chủ dịch vụ chia sẻ dữ liệu</b>	<b>1</b>
	Máy chủ: RAM 16 GB, CPU 16 vCore, SSD 256GB, 01 IP Address public, Băng thông 100 Mbps	<b>1</b>
<b>II</b>	<b>Máy chủ cơ sở dữ liệu Cloud</b>	
	Máy chủ Cloud: RAM 16 GB, CPU 16 vCore, SSD 300 GB, 01 IP Address, Băng thông 100 Mbps	<b>1</b>
<b>III</b>	<b>Thiết bị mạng</b>	<b>1</b>
	Switch	
	Firewall	
<b>IV</b>	<b>Phần mềm bản quyền</b>	<b>1</b>
1	Hệ quản trị cơ sở dữ liệu SQL Server Std 2019 trở lên	1
2	Hệ điều hành máy chủ Windows Server 2019 trở lên	2
3	Phần mềm diệt virus	2

### **II.1.5. Sản phẩm giao nộp**

Sản phẩm giao nộp hạng mục “Xây dựng danh mục dữ liệu chủ và xây dựng, thiết kế cấu trúc dữ liệu chi tiết cho cơ sở dữ liệu khí tượng thủy văn quốc gia” bao gồm:

- Tài liệu biểu mẫu quy định theo Thông tư 26/2014/TT-BTNMT ngày 28 tháng 05 năm 2014 của Bộ Tài nguyên và Môi trường ban hành quy trình và định mức kinh tế - kỹ thuật xây dựng cơ sở dữ liệu tài nguyên và môi trường;

- Tài liệu biểu mẫu quy định theo Thông tư 14/2020/TT-BTNMT ngày 27 tháng 11 năm 2020 của Bộ Tài nguyên và Môi trường ban hành quy trình và định mức kinh tế - kỹ thuật xây dựng, duy trì, vận hành hệ thống thông tin ngành tài nguyên và môi trường.

### **II.1.6. Tổ chức thực hiện**

Đơn vị thi công được giao nhiệm vụ thực hiện hạng mục “Xây dựng danh mục dữ liệu chủ và thiết kế cấu trúc dữ liệu chi tiết cho cơ sở dữ liệu khí tượng thủy văn quốc gia” tổ chức thực hiện theo các nội dung chính sau:

- Lập phương án kỹ thuật thi công chi tiết trên cơ sở Thiết kế kỹ thuật thi công đã được phê duyệt bao gồm các nội dung (phân tích nguồn dữ liệu sử dụng để thi công, phương pháp thi công, quy trình thi công, các tiêu chuẩn kỹ thuật cần tuân theo, các sản phẩm cần giao nộp và tiến độ thực hiện).

- Trình phương án kỹ thuật thi công để chủ đầu tư phê duyệt.

- Thực hiện thi công theo nội dung của phương án kỹ thuật thi công đã được chủ đầu tư phê duyệt.

- Giao nộp sản phẩm đủ số lượng, theo đúng mẫu, đảm bảo chất lượng cho chủ đầu tư. Thực hiện việc bổ sung, chỉnh sửa và hoàn thiện theo ý kiến của bộ phận kiểm tra, nghiệm thu.

- Nghiệm thu, bàn giao đưa cơ sở dữ liệu vào vận hành, thực hiện nghĩa vụ bảo hành, bảo trì theo quy định.

## **II.2. Xây dựng, triển khai mô hình cơ sở dữ liệu ứng dụng công nghệ Big Data cho Trung tâm dữ liệu tại Tổng cục KTTV**

### **II.2.1. Các chỉ tiêu kỹ thuật áp dụng trong triển khai lắp đặt, cài đặt, kiểm tra và hiệu chỉnh thiết bị**

Hệ thống được triển khai phải tuân thủ Danh mục tiêu chuẩn kỹ thuật về ứng dụng công nghệ thông tin trong cơ quan nhà nước được ban hành trong thông tư 39/2017/TT-BTTTT ngày 15 tháng 12 năm 2017 như sau:

Số TT	Loại tiêu chuẩn	Ký hiệu tiêu chuẩn	Tên đầy đủ của tiêu chuẩn	Quy định áp dụng
<b>1</b>	<b>Tiêu chuẩn về kết nối</b>			
1.1	Truyền siêu văn bản	HTTP v1.1	Hypertext Transfer Protocol version 1.1	Bắt buộc áp dụng
		HTTP v2.0	Hypertext Transfer Protocol version 2.0	Khuyến nghị áp dụng
1.2	Truyền tệp tin	FTP	File Transfer Protocol	

		HTTP v1.1	Hypertext Transfer Protocol version 1.1	Bắt buộc áp dụng một hoặc cả hai tiêu chuẩn
		HTTP v2.0	Hypertext Transfer Protocol version 2.0	Khuyến nghị áp dụng
		WebDAV	Web-based Distributed Authoring and Versioning	Khuyến nghị áp dụng
1.3	Truyền, phát luồng âm thanh/ hình ảnh	RTSP	Real-time Streaming Protocol	Khuyến nghị áp dụng
		RTP	Real-time Transport Protocol	Khuyến nghị áp dụng
		RTCP	Real-time Control Protocol	Khuyến nghị áp dụng
1.4	Truy cập và chia sẻ dữ liệu	OData v4	Open Data Protocol version 4.0	Khuyến nghị áp dụng
1.5	Truyền thư điện tử	SMTP/ MIME	Simple Mail Transfer	Bắt buộc áp dụng
			Protocol/Multipurpose	
			Internet Mail Extensions	
1.6	Cung cấp dịch vụ truy cập hộp thư điện tử	POP3	Post Office Protocol version 3	Bắt buộc áp dụng cả hai tiêu chuẩn đối với máy chủ
		IMAP 4rev1	Internet Message Access Protocol version 4 revision 1	
1.7	Truy cập thư mục	LDAP v3	Lightweight Directory Access Protocol version 3	Bắt buộc áp dụng
1.8	Dịch vụ tên miền	DNS	Domain Name System	Bắt buộc áp dụng
1.9	Giao vận mạng có kết nối	TCP	Transmission Control Protocol	Bắt buộc áp dụng
1.10	Giao vận mạng không kết nối	UDP	User Datagram Protocol	Bắt buộc áp dụng
1.11	Liên mạng LAN/WAN	IPv4	Internet Protocol version 4	Bắt buộc áp dụng
		IPv6	Internet Protocol version 6	Bắt buộc áp dụng đối với các thiết bị có kết nối Internet
1.12	Mạng cục bộ không dây	IEEE 802.11g	Institute of Electrical and Electronics Engineers Standard (IEEE) 802.11g	Bắt buộc áp dụng

		IEEE 802.11n	Institute of Electrical and Electronics Engineers Standard (IEEE) 802.11n	Khuyến nghị áp dụng
1.13	Truy cập Internet với thiết bị không dây	WAP v2.0	Wireless Application Protocol version 2.0	Bắt buộc áp dụng
1.14	Dịch vụ Web dạng SOAP	SOAP v1.2	Simple Object Access Protocol version 1.2	Bắt buộc áp dụng một, hai hoặc cả ba tiêu chuẩn
		WSDL V2.0	Web Services Description Language version 2.0	
		UDDI v3	Universal Description, Discovery and Integration version 3	
1.15	Dịch vụ Web dạng RESTful	RESTful web service	Representational state transfer	Khuyến nghị áp dụng
1.16	Dịch vụ đặc tả Web	WS BPEL v2.0	Web Services Business Process Execution Language Version 2.0	Khuyến nghị áp dụng
		WS-I Simple SOAP Binding Profile Version 1.0	Simple SOAP Binding Profile Version 1.0	Khuyến nghị áp dụng
		WS- Federation v1.2	Web Services Federation Language Version 1.2	Khuyến nghị áp dụng
		WS- Addressing v1.0	Web Services Addressing 1.0	Khuyến nghị áp dụng
		WS-Coordination Version 1.2	Web Services Coordination Version 1.2	Khuyến nghị áp dụng
		WS-Policy v1.2	Web Services Coordination Version 1.2	Khuyến nghị áp dụng
		OASIS Web Services Business Activity Version 1.2	Web Services Business Activity Version 1.2	Khuyến nghị áp dụng
		WS- Discovery Version 1.1	Web Services Dynamic Discovery Version 1.1	Khuyến nghị áp dụng

		WS-MetadataExchange	Web Services Metadata Exchange	Khuyến nghị áp dụng
1.17	Dịch vụ đồng bộ thời gian	NTPv3	Network Time Protocol version 3	Bắt buộc áp dụng một trong hai tiêu chuẩn
		NTPv4	Network Time Protocol version 4	
<b>2</b>	<b>Tiêu chuẩn về tích hợp dữ liệu</b>			
2.1	Ngôn ngữ định dạng văn bản	XML v1.0 (5 <sup>th</sup> Edition)	Extensible Markup Language version 1.0 (5 <sup>th</sup> Edition)	Bắt buộc áp dụng một trong hai tiêu chuẩn
		XML v1.1 (2 <sup>nd</sup> Edition)	Extensible Markup Language version 1.1	
2.2	Ngôn ngữ định dạng văn bản cho giao dịch điện tử	ISO/TS 15000:2014	Electronic Business	Bắt buộc áp dụng
			Extensible Markup Language (ebXML)	
2.3	Định nghĩa các lược đồ trong tài liệu XML	XML Schema V1.1	XML Schema version 1.1	Bắt buộc áp dụng
2.4	Biến đổi dữ liệu	XSL	Extensible Stylesheet Language	Bắt buộc áp dụng phiên bản mới nhất.
2.5	Mô hình hóa đối tượng	UML v2.5	Unified Modelling Language version 2.5	Khuyến nghị áp dụng
2.6	Mô tả tài nguyên dữ liệu	RDF	Resource Description Framework	Khuyến nghị áp dụng
		OWL	Web Ontology Language	Khuyến nghị áp dụng
2.7	Trình diễn bộ ký tự	UTF-8	8-bit Universal Character Set (UES)/Unicode Transformation Format	Bắt buộc áp dụng
2.8	Khuôn thức trao đổi thông tin địa lý	GML v3.3	Geography Markup Language version 3.3	Bắt buộc áp dụng
2.9	Truy cập và cập nhật các thông tin địa lý	WMS v1.3.0	OpenGIS Web Map Service version 1.3.0	Bắt buộc áp dụng
		WFS v1.1.0	Web Feature Service version 1.1.0	Bắt buộc áp dụng
2.10	Trao đổi dữ liệu đặc tả tài liệu XML	XMI v2.4.2	XML Metadata Interchange version 2.4.2	Khuyến nghị áp dụng

2.11	Sổ đăng ký siêu dữ liệu (MDR)	ISO/IEC 11179:2015	Sổ đăng ký siêu dữ liệu (Metadata registries - MDR)	Khuyến nghị áp dụng
2.12	Bộ phần tử siêu dữ liệu Dublin Core	ISO 15836-1:2017	Bộ phần tử siêu dữ liệu Dublin Core	<a href="#">Khuyến nghị áp dụng(*)</a>
2.13	Định dạng trao đổi siêu dữ liệu mô tả đối tượng dạng kịch bản JavaScript	JSON RFC 7159	JavaScript Object Notation	Khuyến nghị áp dụng
2.14	Ngôn ngữ mô hình quy trình nghiệp vụ	BPMN 2.0	Business Process Model and Notation version 2.0	Khuyến nghị áp dụng
<b>3</b>	<b>Tiêu chuẩn về truy cập thông tin</b>			
3.1	Chuẩn nội dung Web	HTML v4.01	Hypertext Markup Language version 4.01	Bắt buộc, áp dụng
		WCAG 2.0	W3C Web Content Accessibility Guidelines (WCAG) 2.0	Khuyến nghị áp dụng
		HTML 5	Hypertext Markup Language version 5	Khuyến nghị áp dụng
3.2	Chuẩn nội dung Web mở rộng	XHTML v1.1	Extensible Hypertext Markup Language version 1.1	Bắt buộc áp dụng
3.3	Giao diện người dùng	CSS2	Cascading Style Language Level 2	Bắt buộc áp dụng một trong ba tiêu chuẩn
		CSS3	Cascading Style Language Level 3	
		XSL	Extensible Stylesheet Language version	
3.4	Văn bản	(.txt)	Định dạng Plain Text (.txt): Dành cho các tài liệu cơ bản không có cấu trúc	Bắt buộc áp dụng
		(.rtf) v1.8, v1.9.1	Định dạng Rich Text (.rtf) phiên bản 1.8, 1.9.1: Dành cho các tài liệu có thể trao đổi giữa các nền khác nhau	Bắt buộc áp dụng

		(.docx)	Định dạng văn bản Word mở rộng của Microsoft (.docx)	Khuyến nghị áp dụng
		(.pdf) v1.4, v1.5, v1.6, v1.7	Định dạng Portable Document (.pdf) phiên bản 1.4, 1.5, 1.6, 1.7: Dành cho các tài liệu chỉ đọc	Bắt buộc áp dụng một, hai hoặc cả ba tiêu chuẩn
		(.doc)	Định dạng văn bản Word của Microsoft (.doc)	
		(.odt) v1.2	Định dạng Open Document Text (.odt) phiên bản 1.2	
3.5	Bảng tính	(.csv)	Định dạng Comma eparated Variable/Delimited (.csv): Dành cho các bảng tính cần trao đổi giữa các ứng dụng khác nhau.	Bắt buộc áp dụng
		(.xlsx)	Định dạng bảng tính Excel mở rộng của Microsoft (.xlsx)	Khuyến nghị áp dụng
		(.xls)	Định dạng bảng tính Excel của Microsoft (.xls)	Bắt buộc áp dụng một hoặc cả hai tiêu chuẩn
		(.ods) v1.2	Định dạng Open Document Spreadsheets (.ods) phiên bản 1.2	
3.6	Trình diễn	(.htm)	Định dạng Hypertext Document (.htm): cho các trình bày được trao đổi thông qua các loại trình duyệt khác nhau	Bắt buộc áp dụng
		(.pptx)	Định dạng PowerPoint mở rộng của Microsoft (.pptx)	Khuyến nghị áp dụng
		(.pdf)	Định dạng Portable Document (.pdf): cho các trình bày lưu dưới dạng chỉ đọc	Bắt buộc áp dụng một, hai hoặc cả ba tiêu chuẩn

		(.ppt)	Định dạng PowerPoint (.ppt) của Microsoft	
		(.odp) v1.2	Định dạng Open Document Presentation (.odp) phiên bản 1.2	
3.7	Ảnh đồ họa	JPEG	Joint Photographic Expert Group (.jpg)	Bắt buộc áp dụng một, hai, ba hoặc cả bốn tiêu chuẩn
		GIF v89a	Graphic Interchange (.gif) version 89a	
		TIFF	Tag Image File (.tif)	
		PNG	Portable Network Graphics (.png)	
3.8	Ảnh gắn với tọa độ địa lý	GEO TIFF	Tagged Image File Format for GIS applications	Bắt buộc áp dụng
3.9	Phim ảnh, âm thanh	MPEG-1	Moving Picture Experts Group-1	Khuyến nghị áp dụng
		MPEG-2	Moving Picture Experts Group-2	Khuyến nghị áp dụng
		MPEG-4	Moving Picture Experts Group-4	Khuyến nghị áp dụng
		MP3	MPEG-1 Audio Layer 3	Khuyến nghị áp dụng
		AAC	Advanced Audio Coding	Khuyến nghị áp dụng
3.10	Luồng phim ảnh, âm thanh	(.asf), (.wma), (.wmv)	Các định dạng của Microsoft Windows Media Player (.asf), (.wma), (.wmv)	Khuyến nghị áp dụng
		(.ra), (.rm), (.ram), (.rmm)	Các định dạng Real Audio/Real Video (.ra), (.rm), (.ram), (.rmm)	Khuyến nghị áp dụng
		(.avi), (.mov), (.qt)	Các định dạng Apple Quicktime (.avi), (.mov), (.qt)	Khuyến nghị áp dụng
3.11	Hoạt họa	GIF v89a	Graphic Interchange (.gif) version 89a	Khuyến nghị áp dụng
		(.swf)	Định dạng Macromedia Flash (.swf)	Khuyến nghị áp dụng
		(.swf)	Định dạng Macromedia Shockwave (.swf)	Khuyến nghị áp dụng

		(.avi), (.mov)	(.qt), Các định dạng Apple Quicktime (.avi),(qt),(mov)	Khuyến nghị áp dụng
3.12	Chuẩn nội dung cho thiết bị di động	WML v2.0	Wireless Markup Language version 2.0	Bắt buộc áp dụng
3.13	Bộ ký tự và mã hóa	ASCII	American Standard Code for Information Interchange	Bắt buộc áp dụng
3.14	Bộ ký tự và mã hóa cho tiếng Việt	TCVN 6909:2001	TCVN 6909:2001 “Công nghệ thông tin - Bộ mã ký tự tiếng Việt 16-bit”	Bắt buộc áp dụng
3.15	Nén dữ liệu	Zip	Zip (.zip)	Bắt buộc áp dụng một hoặc cả hai tiêu chuẩn
		.gz v4.3	GNU Zip (.gz) version 4.3	
3.16	Ngôn ngữ kịch bản phía trình khách	ECMA 262	ECMAScript version 6 (6 <sup>th</sup> Edition)	Bắt buộc áp dụng
3.17	Chia sẻ nội dung Web	RSS v1.0	RDF Site Summary version 1.0	Bắt buộc áp dụng một trong hai tiêu chuẩn
		RSS v2.0	Really Simple Syndication version 2.0	
		ATOM v1.0	ATOM version 1.0	Khuyến nghị áp dụng
3.18	Chuẩn kết nối ứng dụng công thông tin điện tử	JSR 168	Java Specification Requests 168 (Portlet Specification)	Bắt buộc áp dụng
		JSR286	Java Specification Requests 286 (Portlet Specification)	Khuyến nghị áp dụng
		WSRP v1.0	Web Services for Remote Portlets version 1.0	Bắt buộc áp dụng
		WSRP v2.0	Web Services for Remote Portlets version 2.0	Khuyến nghị áp dụng
<b>4</b>	<b>Tiêu chuẩn về an toàn thông tin</b>			
4.1	An toàn thư điện tử	S/MIME v3.2	Secure Multi-purpose Internet Mail Extensions version 3.2	Bắt buộc áp dụng
		OpenPGP	OpenPGP	Khuyến nghị áp dụng
4.2	An toàn tầng giao vận	SSH v2.0	Secure Shell version 2.0	Bắt buộc áp dụng

		TLS v1.2	Transport Layer Security version 1.2	Bắt buộc áp dụng
4.3	An toàn truyền tệp tin	HTTPS	Hypertext Transfer Protocol Secure	Bắt buộc áp dụng
		FTPS	File Transfer Protocol Secure	Khuyến nghị áp dụng
		SFTP	SSH File Transfer Protocol	Khuyến nghị áp dụng
4.4	An toàn truyền thư điện tử	SMTPTS	Simple Mail Transfer Protocol Secure	Bắt buộc áp dụng
4.5	An toàn dịch vụ truy cập hộp thư	POP3S	Post Office Protocol version 3 Secure	Bắt buộc áp dụng một hoặc cả hai tiêu chuẩn
		IMAPS	Internet Message Access Protocol Secure	
4.6	An toàn dịch vụ DNS	DNSSEC	Domain Name System Security Extensions	Khuyến nghị áp dụng
4.7	An toàn tầng mạng	IPsec - IP ESP	Internet Protocol security với IP ESP	Bắt buộc áp dụng
4.8	An toàn thông tin cho mạng không dây	WPA2	Wi-fi Protected Access 2	Bắt buộc áp dụng
4.9	Giải thuật mã hóa	TCVN 7816:2007	Công nghệ thông tin. Kỹ thuật mật mã thuật toán mã dữ liệu AES	Khuyến nghị áp dụng
		3DES	Triple Data Encryption Standard	Khuyến nghị áp dụng
		PKCS #1 V2.2	RSA Cryptography Standard - version 2.2	Khuyến nghị áp dụng, sử dụng lược đồ RSAES-OAEP để mã hóa
		ECC	Elliptic Curve Cryptography	Khuyến nghị áp dụng
4.10	Giải thuật chữ ký số	PKCS #1 V2.2	RSA Cryptography Standard - version 2.2	Bắt buộc áp dụng, sử dụng lược đồ RSASSA-PSS để ký
		ECDSA	Elliptic Curve Digital Signature Algorithm	Khuyến nghị áp dụng
4.11	Giải thuật băm cho chữ ký số	SHA-2	Secure Hash Algorithms-2	Khuyến nghị áp dụng
4.12	Giải thuật truyền khóa	RSA-KEM	Rivest-Shamir-Adleman - KEM (Key Encapsulation Mechanism) Key Transport Algorithm	Bắt buộc áp dụng

		ECDHE	Elliptic Curve Diffie Hellman Ephemeral	Khuyến nghị áp dụng
4.13	Giải pháp xác thực người sử dụng	SAML v2.0	Security Assertion Markup Language version 2.0	Khuyến nghị áp dụng
4.14	An toàn trao đổi bản tin XML	XML Encryption Syntax and Processing	XML Encryption Syntax and Processing	Bắt buộc áp dụng
		XML Signature Syntax and Processing	XML Signature Syntax and Processing	Bắt buộc áp dụng
4.15	Quản lý khóa công khai bản tin XML	XKMS v2.0	XML Key Management Specification version 2.0	Khuyến nghị áp dụng
4.16	Giao thức an toàn thông tin cá nhân	P3P v1.1	Platform for Privacy Preferences Project version 1.1	Khuyến nghị áp dụng
4.17	Hạ tầng khóa công khai			Khuyến nghị áp dụng
	Cú pháp thông điệp mật mã cho ký, mã hóa	PKCS#7 v1.5 (RFC 2315)	Cryptographic message syntax for file-based signing and encrypting version 1.5	
	Cú pháp thông tin thẻ mật mã	PKCS#15 v1.1	Cryptographic token information syntax version 1.1	
	Cú pháp thông tin khóa riêng	PKCS#8 V1.2 (RFC 5958)	Private-Key Information Syntax Standard version 1.2	
	Giao diện thẻ mật mã	PKCS#11 v2.20	Cryptographic token interface standard version 2.20	
	Cú pháp trao đổi thông tin cá nhân	PKCS#12 v1.1	Personal Information Exchange Syntax version 1.1	
	Khuôn dạng danh sách chứng thư số thu hồi	RFC 5280	Certificate Revocation List Profile	
	Khuôn dạng chứng thư số	RFC 5280	Public Key Infrastructure Certificate	

	Cú pháp yêu cầu chứng thực	PKCS#10 v1.7 (RFC 2986)	Certification Request Syntax Specification version 1.7	
	Giao thức trạng thái chứng thư trực tuyến	RFC 6960	On-line Certificate status protocol	
	Giao thức gắn tem thời gian	RFC 3161	Time stamping protocol	
	Dịch vụ tem thời gian	ISO/IEC 18014-1:2008	Information technology Security techniques - Time stamping services	
		ISO/IEC 18014-2:2009	Part 1: Framework	
		ISO/IEC 18014-3:2009	Part 2: Mechanisms producing independent tokens	
		ISO/IEC 18014-4:2015	Part 3: Mechanisms producing linked tokens	
			Part 4: Traceability of time sources	
4.18	An toàn cho dịch vụ Web	WS-Security v1.1.1	Web Services Security: SOAP Message Security Version 1.1.1	Khuyến nghị áp dụng
4.19	Khuôn dạng dữ liệu trao đổi sự cố an toàn mạng	RFC 7970	The Incident Object Description Exchange Format version 2 (IODEF)	Khuyến nghị áp dụng

## ***II.2.2. Danh mục thiết bị lắp đặt, cài đặt và các thông số kỹ thuật của thiết bị***

### ***II.2.2.1. Định cỡ hệ thống***

Để tính dung lượng ổ cứng lưu trữ cho mỗi máy data node trong mạng Hadoop sử dụng HDFS, cần lưu ý rằng HDFS lưu trữ dữ liệu theo cơ chế chia nhỏ và phân tán trên các node. Theo mặc định, Hadoop sử dụng cơ chế chia nhỏ dữ liệu thành các khối (block) có kích thước mặc định là 128MB (tuy nhiên, có thể điều chỉnh).

Dựa vào lượng dữ liệu là 20TB, ta cần tính số lượng khối dữ liệu và sau đó xác định dung lượng ổ cứng cần thiết cho mỗi máy data node. Trong trường hợp này, ta sẽ giả định mỗi khối có kích thước 128MB.

- Tính số lượng khối dữ liệu 1 TB = 1024 GB do vậy với lượng dữ liệu 20 TB =  $20 * 1024 \text{ GB} = 20480 \text{ GB}$

với khối có kích thước là 128MB = 0.128 GB

- Số lượng khối dữ liệu =  $20480 \text{ GB} / 0.128 \text{ GB} = 160,000$  khối

- Tính dung lượng ổ cứng cần thiết cho mỗi máy data node Vì HDFS lưu trữ các bản sao (replicas) của dữ liệu, mặc định là 2 bản sao. Điều này đảm bảo tính sẵn sàng và đáng tin cậy của dữ liệu trong môi trường phân tán.

Mỗi máy data node cần lưu trữ một bản sao của các khối dữ liệu này.

Dung lượng ổ cứng cần thiết cho mỗi máy data node = Tổng dung lượng dữ liệu / Số lượng máy data node. Dung lượng ổ cứng cần thiết cho mỗi máy data node =  $60\text{TB} / 5 = 12\text{TB}$  với 20% lưu trữ để trống để đảm bảo hiệu năng truy suất dữ liệu nên sẽ cần dung lượng cho mỗi node là 14,4TB. Vậy, để lưu trữ lượng dữ liệu 20TB trên 5 máy data node sử dụng HDFS với công nghệ Hadoop, mỗi máy data node cần có ít nhất 14,4TB dung lượng ổ cứng lưu trữ.

Các đề xuất về kích thước và điều chỉnh khác nhau dựa trên loại triển khai, dựa trên các yếu tố triển khai nhất định trong miền và môi trường Hadoop, ta có thể triển khai ở mức cơ bản với các tiêu chí triển khai tích hợp kỹ thuật dữ liệu như sau:

#### ***Số lượng người dùng hoạt động***

Số lượng người dùng đang làm việc trên kho dữ liệu mô hình trong quá trình thiết kế, sử dụng công cụ Analyst hoặc chạy các công việc kỹ thuật dữ liệu trong môi trường chạy thời gian gốc hoặc Hadoop tại bất kỳ thời điểm nào.

#### ***Số lượng đồng thời của các ảnh xạ được đẩy xuống***

Tổng số ánh xạ đang chạy trên các động cơ Blaze, Spark hoặc Hive mà được gửi đồng thời đến Dịch vụ tích hợp dữ liệu.

### ***Số lượng đối tượng trong kho dữ liệu mô hình***

Tổng số đối tượng thiết kế và chạy thời gian trong kho dữ liệu mô hình. Ví dụ, đối tượng dữ liệu, ánh xạ, luồng công việc và ứng dụng.

### ***Số lượng ứng dụng triển khai***

Tổng số ứng dụng được triển khai trên tất cả Dịch vụ tích hợp dữ liệu trong miền Hadoop.

### ***Số lượng đối tượng môi ứng dụng***

Tổng số đối tượng của tất cả các loại được triển khai là một phần của một ứng dụng duy nhất.

### ***Tổng lượng dữ liệu vận hành***

Tổng lượng dữ liệu được xử lý trong môi trường Hadoop tại bất kỳ thời điểm nào.

### ***Tổng số nút dữ liệu***

Tổng số nút dữ liệu trong cụm Hadoop.

### ***yarn.nodemanager.resource.cpu-vcores***

Một thuộc tính trong tệp yarn-site.xml trên cụm Hadoop xác định số lõi ảo cho các container.

### ***yarn.nodemanager.resource.memory-mb***

Một thuộc tính trong tệp yarn-site.xml trên cụm Hadoop xác định bộ nhớ vật lý tối đa có sẵn cho các container.

<b>Yếu tố triển khai</b>	<b>Triển khai ở mức tiêu chuẩn</b>
Số người dùng đang hoạt động đồng thời	10
Số lượng ánh xạ đẩy xuống đồng thời	1000 – 2000
Số đối tượng trong kho Model	< 20,000
Số lượng ứng dụng đã triển khai	< 100
Số đối tượng trên mỗi ứng dụng	50 -100
Tổng khối lượng dữ liệu vận hành trên cụm máy tính	>500 GB

Đối với các trường hợp sử dụng xử lý hàng loạt	1 triệu
--	---------

### ***Môi trường Hadoop***

Bảng sau đây chứa các hướng dẫn về các yếu tố triển khai trong môi trường Hadoop:

<b>Yếu tố triển khai</b>	<b>Triển khai ở mức tiêu chuẩn</b>
Tổng số nút dữ liệu	5
yarn.nodemanager.resource.cpu-vcores	24
yarn.nodemanager.resource.memory-mb	>49152 MB

### ***Khuyến nghị về định cỡ máy chủ***

Bảng sau liệt kê các yêu cầu phần cứng tối thiểu và tối ưu cho cụm Hadoop:

<b>Phần cứng</b>	<b>Triển khai ở mức tiêu chuẩn</b>
Tốc độ CPU	2 - 2.5 GHz
Lõi CPU logic hoặc ảo	24
Tổng bộ nhớ hệ thống	128 GB
Không gian đĩa cục bộ cho yarn.nodemanager.local-dirs1	500 GB
Kích thước khối DFS	128 MB
Hệ số sao chép HDFS	2
Dung lượng đĩa	3 TB
Tổng số đĩa cho HDFS	8 (sử dụng Raid5 với 1 ổ đĩa parity&1 ổ spare)
Tổng dung lượng HDFS trên mỗi nút	14,4 TB
Số node data	5
Tổng dung lượng HDFS trên cụm	60 TB

Dung lượng HDFS thực tế (có sao chép)	3 TB
/tmp mount	20 GB
Yêu cầu dung lượng đĩa cài đặt	22 GB
Băng thông mạng (Ethernet)	10 Gbps (bonded channel)

Danh mục đề xuất cấu hình máy chủ và thiết bị mạng cho hệ thống Bigdata gồm:

ST T	Phân vùng mạng	Tên máy chủ	Số lượng máy chủ	Ghi chú	RAM (GB)	CPU (Core)	Storage (TB)	RAID	Ethernet
	<b>TỔNG</b>		<b>11</b>						
<b>1</b>	<b>Máy chủ riêng phân vùng Database (DB zone)</b>		<b>7</b>						
1.1		Name Node 1	1	HDFS Name Node Mapreduce JobTracker ES Master node Hbase HMaster node Yarn Resource Manager ZooKeeper Ambari agent Apache Oozie	64	24	0.96	RAID 1	2 Port 10Gb

1.2		Name Node 2	1	HDFS Secondary Name Node Mapreduce JobTracker secondary ES Secondary Master node Hbase Secondary Master node	64	24	0.96	RAID 1	2 Port 10Gb
1.3		Data Node 1	1	HDFS data node Mapreduce TaskTracker ES Data node Yarn node Manager	256	48	20	RAID 5	2 Port 10Gb
1.4		Data Node 2	1	HDFS data node Mapreduce TaskTracker ES Data node Yarn node Manager	256	48	20	RAID 5	2 Port 10Gb
1.5		Data Node 3	1	HDFS data node Mapreduce TaskTracker ES Data node Yarn node Manager	256	48	20	RAID 5	2 Port 10Gb
1.6		Data Node 4	1	HDFS data node Mapreduce TaskTracker ES Data node Yarn node Manager	256	48	20	RAID 5	2 Port 10Gb
1.7		Data Node 5	1	HDFS data node Mapreduce TaskTracker ES Data node Yarn node Manager	256	48	20	RAID 5	2 Port 10Gb

2	Máy chủ phân vùng ứng dụng (APP Zone)		4						
2.1		APP	1	Apache Spark driver application HiveService Apache Drill Apache PIG Apache Mahout & SparkMlib Apache SQOOP ApacheFLUME Apache Storm Master Apache Spark cluster Manager Ambari web Tomcat	128	48	4	RAID 5	2 Port 10Gb
2.2		Worker 1	1	Kafka worker Apache Storm worker Apache Spark Worker Ambari agent	128	48	4	RAID 5	2 Port 1Gb
2.3		Worker 2	1	Kafka worker Apache Storm worker Apache Spark Worker Ambari agent	128	48	4	RAID 5	2 Port 10Gb
2.4		Worker 3	1	Kafka worker Apache Storm worker Apache Spark Worker Ambari agent	128	48	4	RAID 5	2 Port 10Gb

#### ***II.2.2.2. Máy chủ riêng phân vùng Database (DB zone)***

Phân vùng Database dành riêng cho việc lưu trữ dữ liệu và phân tách với các phân vùng mạng khác để đảm bảo tính bảo mật dữ liệu. Phân vùng dữ liệu chỉ cho phép truy cập từ Phân vùng ứng dụng, Phân vùng admin với một số dịch vụ cần truy cập vào dữ liệu thông qua firewall.

#### ***II.2.2.3. Máy chủ Name Node 1, 2***

Máy chủ Name node 1 cài đặt các ứng dụng Master chịu trách nhiệm thu nhận thông tin, điều phối và điều khiển các máy chủ data node. Đây là máy chủ quan trọng trong hệ thống Bigdata, nếu máy chủ này dừng hoạt động thì sẽ làm dừng toàn bộ hệ thống Bigdata, vì vậy máy chủ Name node 1 được dự phòng bởi máy chủ Name node 2 đóng vai trò là Slave hay Secondary. Khi máy chủ 1 dừng hoạt động thì máy chủ Name node 2 sẽ đóng vai trò làm Master và hệ thống vẫn có thể hoạt động. Các ứng dụng cài đặt trên máy chủ Name node 1 bao gồm:

- HDFS Name Node
- Mapreduce JobTracker
- ES Master node
- Hbase HMaster node
- Yarn Resource Manager
- ZooKeeper
- Ambari agent
- Apache Oozie

#### ***II.2.2.4. Các máy chủ DataNode***

Các máy chủ Data node đóng vai trò chính trong việc lưu trữ dữ liệu. Ngoài ra còn đảm nhiệm việc xử lý dữ liệu Mapreduce. Đồng thời các data node được cài đặt các ứng dụng worker để giám sát tài nguyên data node. Các ứng dụng được cài đặt bao gồm:

- HDFS data node
- Mapreduce TaskTracker
- ES Data node
- Yarn node Manager
- Hbase Region Server
- Ambari agent

Với nhiệm vụ chính lưu trữ dữ liệu thì Datanode cần thiết phải đảm bảo được các thông số sau:

- Storage có dung lượng lưu trữ 18TB dành riêng cho dữ liệu và 0.5TB dành cho OS và cài đặt ứng dụng.
- Storage đảm bảo chạy với bộ điều khiển RAID 5E để đảm bảo tính dự phòng của ổ cứng lưu trữ.
- Tần suất truy cập dữ liệu nhiều do chứa lượng lớn dữ liệu từ các ứng dụng database như HDFS, Hbase, ES, Mapreduce vì vậy IOPS cần đạt tối thiểu 20000.

Các ứng dụng database yêu cầu cao về CPU và RAM để đảm bảo các hoạt động nền và đáp ứng nhanh các yêu cầu truy vấn dữ liệu vì vậy cấu hình đề xuất CPU 24 Vcore và RAM 125 GB.

Kết nối mạng Ethernet cần đảm bảo yếu tố truy xuất nhanh với băng thông rộng 10GB/s và ít nhất 2 port, trong đó 1 port dự phòng.

#### ***II.2.2.5. Máy chủ phân vùng ứng dụng (DMZ Zone)***

##### ***Máy chủ ứng dụng (APP server)***

Máy chủ ứng dụng (APP server) đóng vai trò cung cấp các ứng dụng cơ sở và cho phép truy cập, triển khai và vận hành các ứng dụng lập trình. Các ứng dụng cơ sở để truy cập vào cơ sở dữ liệu như: HiveService, Sqoop, Flume hoặc triển khai các ứng dụng phân tích và xử lý dữ liệu như: Storm, Spark, Mahout, Drill, PIG. Hoặc ứng dụng giám sát hệ thống như Ambari web. Các ứng dụng cài đặt bao gồm:

- Apache Spark driver application
- HiveService
- Apache Drill
- Apache PIG
- Apache Mahout
- Apache SQOOP
- Apache FLUME
- Apache Storm Master
- Apache Spark cluster Manager
- Ambari web
- Apache Tomcat

Các ứng dụng chạy không đòi hỏi storage lưu trữ dữ liệu lớn chỉ lưu trữ log với 0.8T và 0.5T dành cho OS và cài đặt các ứng dụng. Storage đảm bảo chạy với bộ điều khiển RAID 5 để đảm bảo tính dự phòng của ổ cứng lưu trữ.

Kết nối mạng Ethernet cần đảm bảo yếu tố truy xuất nhanh với băng thông rộng 10GB/s và ít nhất 2 port, trong đó 1 port dự phòng.

Do chạy đồng thời nhiều ứng dụng vì vậy yêu cầu dung lượng RAM 128G và CPU 24 Vcore

### ***Máy chủ Worker***

Các máy chủ triển khai các ứng dụng worker đóng vai trò chạy các tác vụ xử lý dữ liệu, tính toán dạng Data Stream hoặc Data Flow. Các ứng dụng cài đặt bao gồm:

- Kafka worker
- Apache Storm worker
- Apache Spark Worker
- Ambari agent

Với nhiệm vụ đặc biệt xử lý và tính toán cao và đặc biệt là xử lý tính toán trên bộ nhớ RAM thì các máy chủ worker cần được trang bị CPU mạnh mẽ với 24 Vcore và dung lượng RAM 128G.

Các ứng dụng chạy không đòi hỏi storage lưu trữ dữ liệu lớn chỉ lưu trữ log với 0.8T và 0.5T dành cho OS và cài đặt các ứng dụng. Storage đảm bảo chạy với bộ điều khiển RAID 5 để đảm bảo tính dự phòng của ổ cứng lưu trữ.

Kết nối mạng Ethernet cần đảm bảo yếu tố truy xuất nhanh với băng thông rộng 10GB/s và ít nhất 2 port, trong đó 1 port dự phòng.

### ***II.2.3. Thống kê khối lượng các thiết bị, dịch vụ***

<b>TT</b>	<b>Nội dung công việc</b>	<b>Đơn vị tính (ĐVT)</b>	<b>Ghi chú</b>
<b>I</b>	<b>Triển khai nền tảng Big Data và xây dựng cơ sở dữ liệu tập trung</b>		
<b>1</b>	<b>Triển khai lớp lưu trữ HDFS (Hadoop Distributed File System)</b>	<b>Gói dịch vụ</b>	<b>39 dịch vụ</b>
	Cài đặt Master Name Node	Gói dịch vụ	
	Cấu hình Master Name node - Environment	Gói dịch vụ	
	Cấu hình Master Name node - Path	Gói dịch vụ	
	Cấu hình Master Name node - directory	Gói dịch vụ	
	Cấu hình Master Name node - Yarn Job Scheduler	Gói dịch vụ	
	Cấu hình Master Name node - Yarn option config	Gói dịch vụ	

	Cấu hình Master Name node - worker	Gói dịch vụ	
	Cấu hình Master Name node – RAM resource	Gói dịch vụ	
	Cài đặt Master Name node - Format	Gói dịch vụ	
	Cài đặt Secondary Name Node	Gói dịch vụ	
	Cấu hình Secondary Name node - Environment	Gói dịch vụ	
	Cấu hình Secondary Name node - Path		
	Cấu hình Secondary Name node - directory	Gói dịch vụ	
	Cấu hình Secondary Name node - Yarn Job Scheduler	Gói dịch vụ	
	Cấu hình Secondary Name node - Yarn option config	Gói dịch vụ	
	Cấu hình Secondary Name node - worker	Gói dịch vụ	
	Cấu hình Secondary Name node – RAM resource	Gói dịch vụ	
	Cài đặt Secondary Name node - Format	Gói dịch vụ	
	Cài đặt Data Node	Gói dịch vụ	
	Cấu hình Data node - Environment	Gói dịch vụ	
	Cấu hình Data node - Path	Gói dịch vụ	
	Cấu hình Data node - directory	Gói dịch vụ	
	Cấu hình Data node - Yarn option config	Gói dịch vụ	
	Cấu hình Data node - worker	Gói dịch vụ	
	Cấu hình Data node – RAM resource	Gói dịch vụ	
	Cài đặt Data node - Format	Gói dịch vụ	
	Cấu hình Hadoop in Non-Secure Mode	Gói dịch vụ	
	Cấu hình Environment of Hadoop Daemons	Gói dịch vụ	
	Cấu hình the Hadoop Daemons	Gói dịch vụ	
	Cấu hình History Server (Needs to be moved elsewhere)	Gói dịch vụ	
	Cấu hình MapReduce Applications	Gói dịch vụ	
	Cấu hình MapReduce JobHistory Server	Gói dịch vụ	
	Cấu hình Monitoring Health of NodeManagers	Gói dịch vụ	
	Cấu hình Slaves File	Gói dịch vụ	

	Cấu hình Hadoop Rack Awareness	Gói dịch vụ	
	Cấu hình Logging	Gói dịch vụ	
	Cấu hình Operating the Hadoop Cluster	Gói dịch vụ	
	Cấu hình Web Interfaces	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>2</b>	<b>Triển khai lớp quản lý tài nguyên Apache YARN (Yet-Another-Resource-Negotiator)</b>	<b>Gói dịch vụ</b>	<b>15 dịch vụ</b>
	Cấu hình YARN Resource Manager–RAM	Gói dịch vụ	
	Cấu hình YARN Resource Manager–RAM mapreduce	Gói dịch vụ	
	Cấu hình YARN NodeManager	Gói dịch vụ	
	Cài đặt queues	Gói dịch vụ	
	Cài đặt Queue Properties	Gói dịch vụ	
	Cài đặt application priority.	Gói dịch vụ	
	Cài đặt Capacity Scheduler container preemption	Gói dịch vụ	
	Cài đặt Reservation Properties	Gói dịch vụ	
	Cấu hình ReservationSystem with CapacityScheduler	Gói dịch vụ	
	Cấu hình Dynamic Auto-Creation and Management of Leaf Queues	Gói dịch vụ	
	Cập nhật Container (Experimental - API may change in the future)	Gói dịch vụ	
	Cấu hình Activities	Gói dịch vụ	
	Cấu hình Scheduler Activities	Gói dịch vụ	
	Cấu hình Application Activities	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>3</b>	<b>Triển khai lớp ứng dụng MapReduce</b>	<b>Gói dịch vụ</b>	<b>28 gói dịch vụ</b>
	Cài đặt MapReduce	Gói dịch vụ	
	Cấu hình MapReduce – JobTracker config	Gói dịch vụ	
	Cấu hình MapReduce – TaskTracker config	Gói dịch vụ	
	Cấu hình MapReduce – RAM Resource	Gói dịch vụ	
	Cấu hình MapReduce – Directory	Gói dịch vụ	
	Cấu hình MapReduce – Path	Gói dịch vụ	

	Cấu hình Shuffle/Reduce Parameters	Gói dịch vụ	
	Cấu hình Configured Parameters	Gói dịch vụ	
	Cấu hình Task Logs	Gói dịch vụ	
	Cấu hình Distributing Libraries	Gói dịch vụ	
	Cấu hình Job Submission and Monitoring	Gói dịch vụ	
	Cấu hình Job Control	Gói dịch vụ	
	Cấu hình Job Input	Gói dịch vụ	
	Cấu hình InputSplit	Gói dịch vụ	
	Cấu hình RecordReader	Gói dịch vụ	
	Cấu hình Job Output	Gói dịch vụ	
	Cấu hình OutputCommitter	Gói dịch vụ	
	Cấu hình Task Side-Effect Files	Gói dịch vụ	
	Cấu hình RecordWriter	Gói dịch vụ	
	Cấu hình các tính năng Useful	Gói dịch vụ	
	Cấu hình Submitting Jobs to Queues	Gói dịch vụ	
	Cấu hình Counters	Gói dịch vụ	
	Cấu hình DistributedCache	Gói dịch vụ	
	Cấu hình Profiling	Gói dịch vụ	
	Cấu hình Debugging	Gói dịch vụ	
	Cấu hình Data Compression	Gói dịch vụ	
	Cấu hình Skipping Bad Records	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>4</b>	<b>Triển khai thành phần xử lý dữ liệu bộ nhớ trong Spark</b>	<b>Gói dịch vụ</b>	<b>3 gói dịch vụ</b>
	Cài đặt thành phần xử lý dữ liệu bộ nhớ trong Spark	Gói dịch vụ	
	Cấu hình thành phần xử lý dữ liệu bộ nhớ trong Spark	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>5</b>	<b>Triển khai thành phần phân tích dữ liệu Pig</b>	<b>Gói dịch vụ</b>	<b>12 gói dịch vụ</b>
	Cài đặt Apache Pig	Gói dịch vụ	
	Cấu hình Apache Pig	Gói dịch vụ	
	Cấu hình Execution Modes	Gói dịch vụ	
	Cấu hình Interactive Mode	Gói dịch vụ	
	Cấu hình Batch Mode	Gói dịch vụ	

	Cấu hình Running jobs on a Kerberos secured cluster	Gói dịch vụ	
	Cấu hình Short lived jobs	Gói dịch vụ	
	Cấu hình Long lived jobs	Gói dịch vụ	
	Cấu hình Pig Latin Statements	Gói dịch vụ	
	Cấu hình Loading Data	Gói dịch vụ	
	Cấu hình Working with Data	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>6</b>	<b>Triển khai thành phần truy vấn dữ liệu Hive</b>	<b>Gói dịch vụ</b>	<b>21 gói dịch vụ</b>
	Cài đặt Hive	Gói dịch vụ	
	Cấu hình Hive	Gói dịch vụ	
	Cấu hình Compile Hive on master	Gói dịch vụ	
	Cấu hình Compile Hive on branch-1	Gói dịch vụ	
	Cấu hình Compile Hive Prior to 0.13 on Hadoop 0.20	Gói dịch vụ	
	Cấu hình Compile Hive Prior to 0.13 on Hadoop 0.23	Gói dịch vụ	
	Cấu hình Hive CLI	Gói dịch vụ	
	Cấu hình HiveServer2 and Beeline	Gói dịch vụ	
	Cấu hình HCatalog	Gói dịch vụ	
	Cấu hình WebHCat (Templeton)	Gói dịch vụ	
	Cấu hình Management Overview	Gói dịch vụ	
	Cấu hình Runtime	Gói dịch vụ	
	Cấu hình Hive, Map-Reduce and Local-Mode	Gói dịch vụ	
	Cấu hình DDL	Gói dịch vụ	
	Cấu hình Hive Tables	Gói dịch vụ	
	Cấu hình Browsing through Tables	Gói dịch vụ	
	Cấu hình Altering and Dropping Tables	Gói dịch vụ	
	Cấu hình Metadata Store	Gói dịch vụ	
	Cấu hình DML Operations	Gói dịch vụ	
	Cấu hình SQL Operations	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>7</b>	<b>Triển khai thành phần HBase tổ chức lưu trữ cơ sở dữ liệu NoSQL</b>	<b>Gói dịch vụ</b>	<b>31 dịch vụ</b>
	Cài đặt Apache HBase	Gói dịch vụ	

	Cấu hình Apache HBase	Gói dịch vụ	
	Cấu hình Files	Gói dịch vụ	
	Cấu hình HBase Distributed	Gói dịch vụ	
	Cấu hình Namespace	Gói dịch vụ	
	Cấu hình Table	Gói dịch vụ	
	Cấu hình Row	Gói dịch vụ	
	Cấu hình Column Family	Gói dịch vụ	
	Cấu hình Cells	Gói dịch vụ	
	Cấu hình Data Model Operations	Gói dịch vụ	
	Cấu hình Column Metadata	Gói dịch vụ	
	Cấu hình Joins	Gói dịch vụ	
	Cấu hình ACID	Gói dịch vụ	
	Cấu hình HBase và MapReduce	Gói dịch vụ	
	Cấu hình HBase, MapReduce, and the CLASSPATH	Gói dịch vụ	
	Cấu hình MapReduce Scan Caching	Gói dịch vụ	
	Cấu hình Bundled HBase MapReduce Jobs	Gói dịch vụ	
	Cấu hình HBase as a MapReduce Job Data Source and Data Sink	Gói dịch vụ	
	Cấu hình Writing HFiles Directly During Bulk Import	Gói dịch vụ	
	Cấu hình RowCounter Example	Gói dịch vụ	
	Cấu hình Map-Task Splitting	Gói dịch vụ	
	Cấu hình HBase MapReduce Examples	Gói dịch vụ	
	Cấu hình Accessing Other HBase Tables in a MapReduce Job	Gói dịch vụ	
	Cấu hình Speculative Execution	Gói dịch vụ	
	Cấu hình bảo mật Apache HBase	Gói dịch vụ	
	Cấu hình bảo mật Web UI	Gói dịch vụ	
	Cấu hình bảo mật Client Access to Apache HBase	Gói dịch vụ	
	Cấu hình bảo mật Simple User Access to Apache HBase	Gói dịch vụ	
	Cấu hình bảo mật truy cập đến HDFS and ZooKeeper	Gói dịch vụ	
	Cấu hình bảo mật truy cập đến dữ liệu	Gói dịch vụ	

	Hiệu chỉnh	Gói dịch vụ	
<b>8</b>	<b>Triển khai thành phần Apache Drill tổ chức lưu trữ cơ sở dữ liệu SQL</b>	<b>Gói dịch vụ</b>	<b>75 dịch vụ</b>
	Cài đặt Drill	Gói dịch vụ	
	Cài đặt Migrating Parquet Data	Gói dịch vụ	
	Cài đặt Drill in Embedded Mode	Gói dịch vụ	
	Cấu hình Drill	Gói dịch vụ	
	Cấu hình Migrating Parquet Data	Gói dịch vụ	
	Cấu hình Drill-on-YARN	Gói dịch vụ	
	Cấu hình Basic Drill Cluster	Gói dịch vụ	
	Cấu hình Drill Under YARN	Gói dịch vụ	
	Cấu hình Drill-on-YARN Command-Line Tool	Gói dịch vụ	
	Cấu hình Drill-on-YARN Web UI	Gói dịch vụ	
	Cấu hình Multiple Drill Clusters	Gói dịch vụ	
	Cấu hình Enabling Web UI Security	Gói dịch vụ	
	Cấu hình Drill in Embedded Mode	Gói dịch vụ	
	Cấu hình Drill in Distributed Mode	Gói dịch vụ	
	Cấu hình Starting the Web UI	Gói dịch vụ	
	Cấu hình Drill	Gói dịch vụ	
	Cấu hình Drill Introduction	Gói dịch vụ	
	Cấu hình Drill Memory	Gói dịch vụ	
	Cấu hình Securing Drill	Gói dịch vụ	
	Cấu hình Multitenant Cluster	Gói dịch vụ	
	Cấu hình Options	Gói dịch vụ	
	Cấu hình Ports Used by Drill	Gói dịch vụ	
	Cấu hình Drill Shell	Gói dịch vụ	
	Cấu hình cgroups to Control CPU Usage	Gói dịch vụ	
	Kết nối Data Source	Gói dịch vụ	
	Cấu hình Data Source Introduction	Gói dịch vụ	
	Cấu hình Storage Plugin Registration	Gói dịch vụ	
	Cấu hình Storage Plugin Configuration	Gói dịch vụ	
	Cấu hình File System Storage Plugin	Gói dịch vụ	
	Cấu hình HBase Storage Plugin	Gói dịch vụ	
	Cấu hình Hive Storage Plugin	Gói dịch vụ	

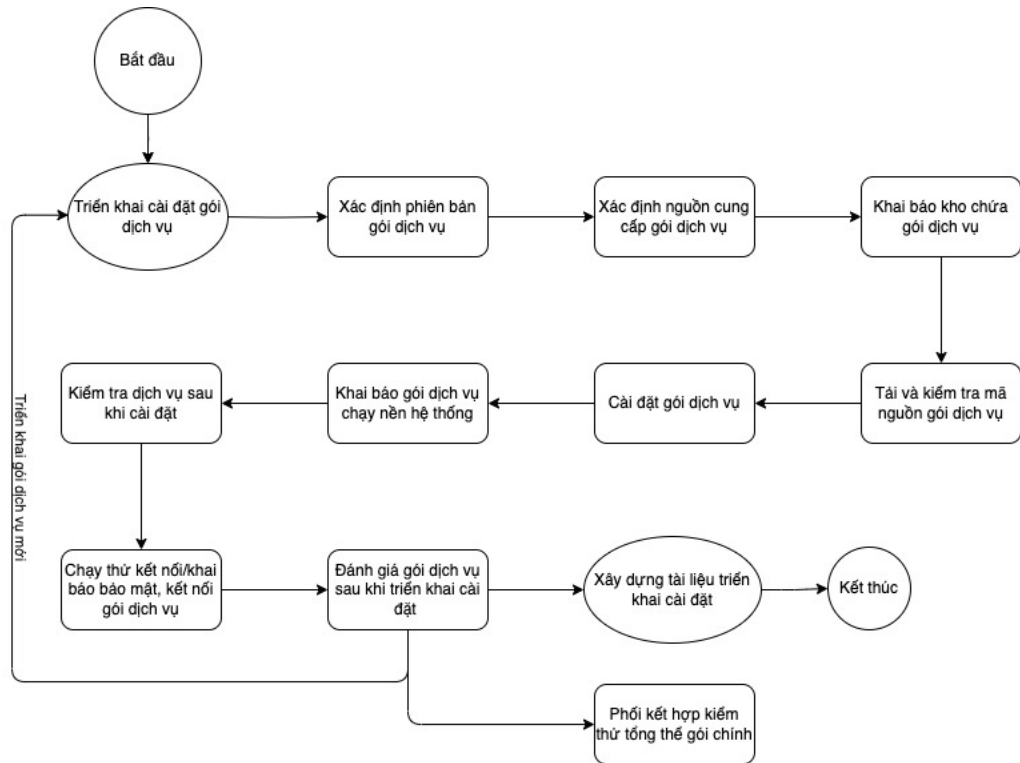
	Cấu hình RDBMS Storage Plugin	Gói dịch vụ	
	Cấu hình MongoDB Storage Plugin	Gói dịch vụ	
	Cấu hình MapR-DB Format	Gói dịch vụ	
	Cấu hình S3 Storage Plugin	Gói dịch vụ	
	Cấu hình OCI OS Storage Plugin	Gói dịch vụ	
	Cấu hình OpenTSDB Storage Plugin	Gói dịch vụ	
	Cấu hình Kafka Storage Plugin	Gói dịch vụ	
	Cấu hình Azure Blob Storage Plugin	Gói dịch vụ	
	Cấu hình Druid Storage Plugin	Gói dịch vụ	
	Cấu hình HTTP Storage Plugin	Gói dịch vụ	
	Cấu hình ElasticSearch Storage Plugin	Gói dịch vụ	
	Cấu hình Splunk Storage Plugin	Gói dịch vụ	
	Cấu hình Cassandra Storage Plugin	Gói dịch vụ	
	Cấu hình Dropbox Storage Plugin	Gói dịch vụ	
	Cấu hình ODBC/JDBC Interfaces	Gói dịch vụ	
	Cấu hình Interfaces Introduction	Gói dịch vụ	
	Cấu hình Using the JDBC Driver	Gói dịch vụ	
	Cấu hình Using JDBC with SQuirreL on Windows	Gói dịch vụ	
	Cấu hình the ODBC Driver	Gói dịch vụ	
	Cấu hình ODBC	Gói dịch vụ	
	Cấu hình Drill Explorer	Gói dịch vụ	
	Cấu hình Drill with BI Tools	Gói dịch vụ	
	Cấu hình Query Data	Gói dịch vụ	
	Cấu hình Querying a File System	Gói dịch vụ	
	Cấu hình Querying HBase	Gói dịch vụ	
	Cấu hình Querying Complex Data	Gói dịch vụ	
	Cấu hình Querying Hive	Gói dịch vụ	
	Cấu hình Querying the INFORMATION SCHEMA	Gói dịch vụ	
	Cấu hình Querying Indexes	Gói dịch vụ	
	Cấu hình Querying System Tables	Gói dịch vụ	
	Cấu hình Monitoring and Canceling Queries in the Drill Web UI	Gói dịch vụ	
	Cấu hình Drill Metastore	Gói dịch vụ	
	Cấu hình Performance Tuning	Gói dịch vụ	

	Cấu hình Partition Pruning	Gói dịch vụ	
	Cấu hình Asynchronous Parquet Reader	Gói dịch vụ	
	Cấu hình Optimizing Parquet Metadata Reading	Gói dịch vụ	
	Cấu hình Parquet Filter Pushdown	Gói dịch vụ	
	Cấu hình Hive Metadata Caching	Gói dịch vụ	
	Cấu hình Choosing a Storage Format	Gói dịch vụ	
	Cấu hình Query Plans and Tuning	Gói dịch vụ	
	Cấu hình Identifying Performance Issues	Gói dịch vụ	
	Cấu hình Performance Tuning Reference	Gói dịch vụ	
	Cấu hình Monitoring Metrics	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>9</b>	<b>Triển khai thành phần ZooKeeper quản lý cụm</b>	<b>Gói dịch vụ</b>	<b>9 dịch vụ</b>
	Cài đặt ZooKeeper	Gói dịch vụ	
	Cấu hình ZooKeeper	Gói dịch vụ	
	Thiết kế triển khai ZooKeeper	Gói dịch vụ	
	Cấu hình Parameters	Gói dịch vụ	
	Cấu hình Data File Management	Gói dịch vụ	
	Cấu hình The Data Directory	Gói dịch vụ	
	Cấu hình The Log Directory	Gói dịch vụ	
	Cấu hình File Management	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>10</b>	<b>Triển khai thành phần Oozie lên lịch công việc</b>	<b>Gói dịch vụ</b>	<b>27 dịch vụ</b>
	Cài đặt Oozie	Gói dịch vụ	
	Cấu hình Oozie	Gói dịch vụ	
	Cấu hình máy chủ Oozie	Gói dịch vụ	
	Cấu hình Oozie with an Alternate Tomcat	Gói dịch vụ	
	Cấu hình Database	Gói dịch vụ	
	Cấu hình Oozie	Gói dịch vụ	
	Cấu hình Oozie Properties	Gói dịch vụ	
	Cấu hình Logging	Gói dịch vụ	
	Cấu hình Oozie User Authentication	Gói dịch vụ	

	Cấu hình Oozie Hadoop Authentication	Gói dịch vụ	
	Cấu hình User ProxyUser	Gói dịch vụ	
	Cấu hình User Authorization	Gói dịch vụ	
	Cấu hình Oozie System ID	Gói dịch vụ	
	Cấu hình Filesystem	Gói dịch vụ	
	Cấu hình HCatalog	Gói dịch vụ	
	Cấu hình Notifications	Gói dịch vụ	
	Cấu hình Oozie with HTTPS (SSL)	Gói dịch vụ	
	Cấu hình Self-Signed Certificate	Gói dịch vụ	
	Cấu hình Certificate from a Certificate Authority	Gói dịch vụ	
	Cấu hình Oozie Server to use SSL (HTTPS)	Gói dịch vụ	
	Cấu hình Oozie Client to connect using SSL (HTTPS)	Gói dịch vụ	
	Cấu hình Oozie Web UI using SSL (HTTPS)	Gói dịch vụ	
	Cấu hình Oozie HA with SSL	Gói dịch vụ	
	Cấu hình Fine Tuning an Oozie Server	Gói dịch vụ	
	Cấu hình High Availability (HA)	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>11</b>	<b>Triển khai thành phần Flume đồng hóa dữ liệu phi cấu trúc</b>	<b>Gói dịch vụ</b>	<b>8 gói dịch vụ</b>
	Cài đặt Apache Flume	Gói dịch vụ	
	Cấu hình Flume	Gói dịch vụ	
	Cấu hình HDP and Flume	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>13</b>	<b>Triển khai thành phần Sqoop đồng hóa dữ liệu có cấu trúc</b>	<b>Gói dịch vụ</b>	
	Cài đặt Apache Sqoop	Gói dịch vụ	
	Cấu hình Sqoop	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>14</b>	<b>Triển khai hệ thống Kafka cho truyền nhận thông điệp</b>	<b>Gói dịch vụ</b>	<b>3 gói dịch vụ</b>
	Cài đặt Kafka	Gói dịch vụ	
	Cấu hình Kafka cluster	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	

<b>15</b>	<b>Triển khai thành phần Elasticsearch cho tìm kiếm dữ liệu</b>	<b>Gói dịch vụ</b>	<b>10 dịch vụ</b>
	Cài đặt Elasticsearch master node	Gói dịch vụ	
	Cài đặt Elasticsearch cluster	Gói dịch vụ	
	Cài đặt Elasticsearch data node	Gói dịch vụ	
	Cấu hình Elasticsearch directory	Gói dịch vụ	
	Cấu hình Elasticsearch Memory	Gói dịch vụ	
	Cấu hình Elasticsearch Port	Gói dịch vụ	
	Cấu hình Mapping Template	Gói dịch vụ	
	Cài đặt Elasticsearch Kibana	Gói dịch vụ	
	Cấu hình Elasticsearch Kibana	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>16</b>	<b>Triển khai thành phần giám sát và bảo trì cụm Ambari</b>	<b>Gói dịch vụ</b>	<b>11 dịch vụ</b>
	Cài đặt Ambari	Gói dịch vụ	
	Cấu hình Ambari	Gói dịch vụ	
	Cấu hình Ambari Server	Gói dịch vụ	
	Cấu hình Cluster	Gói dịch vụ	
	Cấu hình Version	Gói dịch vụ	
	Cấu hình Hosts	Gói dịch vụ	
	Cấu hình Services	Gói dịch vụ	
	Cấu hình Assign Masters	Gói dịch vụ	
	Cấu hình Assign Slaves and Clients	Gói dịch vụ	
	Cấu hình Services	Gói dịch vụ	
	Hiệu chỉnh	Gói dịch vụ	
<b>17</b>	<b>Tích hợp hệ thống với các dịch vụ dùng chung</b>	<b>Gói dịch vụ</b>	

## II.2.4. Chỉ dẫn biện pháp triển khai



*Quy trình triển khai cài đặt một gói dịch vụ*

Khi bắt đầu triển khai cài đặt một gói dịch vụ chuyên gia có trách nhiệm xác định phiên bản cài đặt gói dịch vụ đó đảm bảo phù hợp với tổng thể cũng như khả năng tương thích với các gói dịch vụ khác. Sau khi kiểm tra phiên bản gói dịch vụ chuyên gia cần xác định nguồn cung cấp gói dịch vụ đảm bảo tính nguyên vẹn, được xác định từ nhà cung cấp và khai báo kho chứa (repository) lên hệ điều hành.

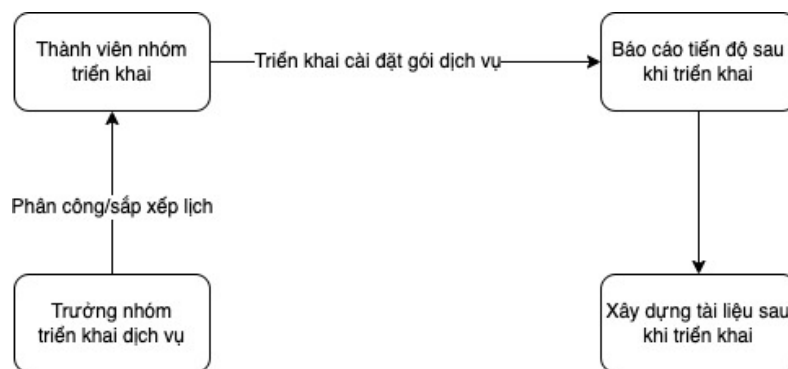
Khi tải xong gói dịch vụ chuyên gia cần kiểm tra lại lần nữa mã băm SHA mà nhà cung cấp thông báo công khai trùng khớp với mã băm SHA mà gói cài đặt được tải về đảm bảo tính nguyên vẹn không bị sửa xoá của gói dịch vụ.

Sau khi có được gói dịch vụ được tải về chuyên gia tiến hành cài đặt và khai báo cấu hình chi tiết cũng như thực hiện khai báo gói dịch vụ theo danh sách chạy nền của hệ điều hành (mỗi khi máy khởi động lại thì dịch vụ tự động tắt/bật).

Việc kiểm tra gói dịch vụ sau khi cài đặt sẽ được kết hợp với quy trình kiểm thử dữ liệu sau cài đặt và quy trình kiểm thử toàn bộ hệ thống, cũng như kiểm tra tính kết nối của gói dịch vụ xác định nguồn/đích kết nối đến gói dịch vụ đã cài đặt.

Đánh giá gói dịch vụ sau khi triển khai cài đặt là một bước bắt buộc trước khi thực hiện cài đặt một gói dịch vụ khác, cũng như chuyên gia phải thực hiện xây dựng tài liệu triển khai cài đặt bàn giao lại chủ đầu tư cũng như cho nhóm chuyên gia khác phối hợp thực hiện.

Trong toàn bộ quá trình triển khai việc lập nhật ký công tác triển khai thì thành viên được phân công có trách nhiệm thực hiện; Thực hiện giám sát công tác triển khai và lập nhật ký giám sát công tác triển khai là yêu cầu bắt buộc và được trưởng nhóm thực hiện.



*Quy trình quản lý triển khai các gói dịch vụ*

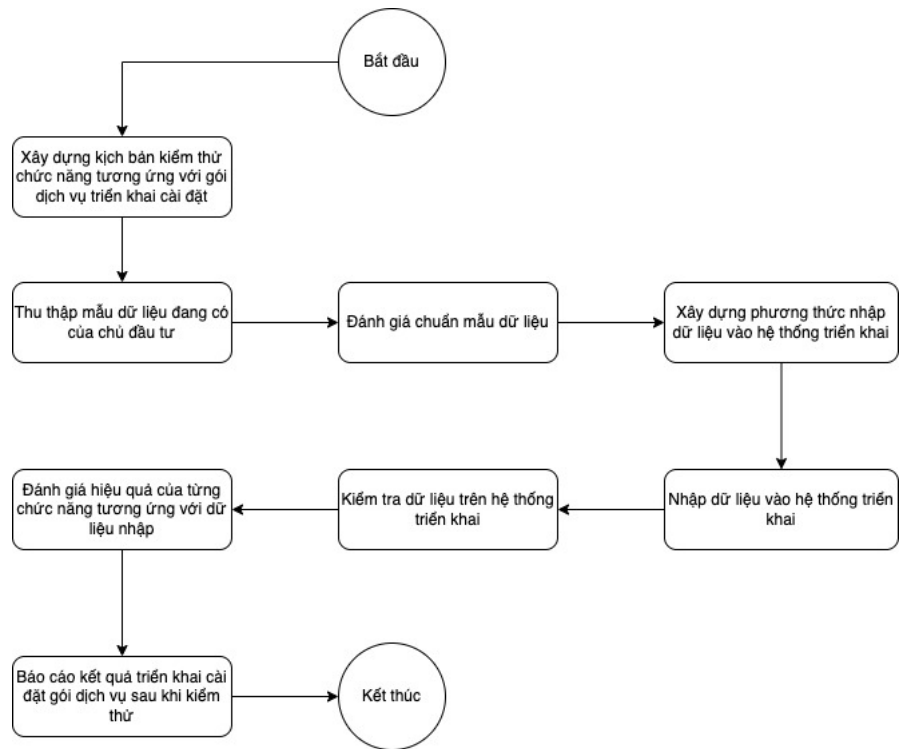
Trưởng nhóm triển khai có trách nhiệm phân công, sắp xếp lịch trình, giai đoạn, phân loại từng gói dịch vụ cài đặt khác nhau phù hợp khả năng của mỗi chuyên gia và cho phép mỗi gói dịch vụ sau khi cài đặt có tính liên kết để thực hiện kiểm thử.

Thành viên nhóm triển khai sau khi được phân công công việc có trách nhiệm xác định gói dịch vụ cài đặt mà mình được phân công có phù hợp với khả năng của mình hay không và báo cáo lại với trưởng nhóm triển khai về sự phù hợp về gói dịch vụ được phân công. Thành viên nhóm triển khai phải tuân thủ quy trình cài đặt triển khai cài đặt một gói dịch vụ cũng như phải báo cáo tiến độ mỗi ngày sau khi triển khai với trưởng nhóm để trưởng nhóm xây dựng phương án, kế hoạch phù hợp thời gian toàn bộ gói thầu.

Thành viên nhóm triển khai phải xây dựng tài liệu sau khi triển khai gồm:

- Tên, phiên bản gói dịch vụ triển khai
- Thời gian bắt đầu và kết thúc cài đặt gói dịch vụ
- Các tham số cài đặt, cấu hình gói dịch vụ
- Kết quả kiểm tra gói dịch vụ sau khi cài đặt cần có các nội dung như trạng thái gói dịch vụ trước và sau khi kiểm thử, kết quả đầu ra của dữ liệu khi chạy trên hệ thống triển khai, ....

Sau khi kết thúc một kịch bản kiểm thử đặt thành viên nhóm kiểm thử thực hiện kiểm thử kịch bản tiếp theo như đã được phân công.



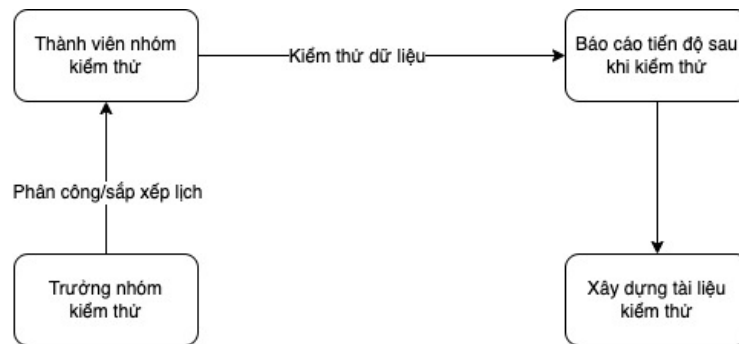
*Quy trình chuẩn bị dữ liệu và kiểm thử chức năng gói dịch vụ sau khi triển khai cài đặt*

Để thực hiện việc kiểm thử chức năng gói dịch vụ sau khi triển khai cài đặt trường nhóm chuyên gia kiểm thử phối hợp cùng chuyên gia kiểm thử xây dựng kịch bản kiểm thử chức năng tương ứng với gói dịch vụ triển khai cài đặt, các kịch bản kiểm thử sẽ cần xác định mẫu dữ liệu hiện có để đảm bảo hệ thống khi được triển khai sẽ hoạt động với theo nhu cầu của chủ đầu tư.

Việc thu thập mẫu dữ liệu hiện có là quá trình hai chiều gồm xác định mẫu dữ liệu đến kịch bản kiểm thử và ngược lại giúp cho các bước đánh giá chuẩn mẫu dữ liệu và phương thức nhập dữ liệu vào hệ thống triển khai đạt được kết quả tốt nhất, cũng như giúp nhóm triển khai cài đặt thực hiện bước cấu hình gói dịch vụ tối ưu tài nguyên xử lý dữ liệu.

Với dữ liệu sau khi nhập vào hệ thống triển khai, chuyên gia kiểm thử cần kiểm tra dữ liệu trên hệ thống triển khai có được kết quả như mong muốn với từng kịch bản kiểm thử đã xây dựng. Từ đó đánh giá hiệu quả của từng chức năng tương ứng với dữ liệu-kịch bản kiểm thử và hiệu năng tổng thể của gói hệ thống dịch vụ đã triển khai cài đặt.

Chuyên gia kiểm thử sau mỗi kịch bản kiểm thử kết hợp với chuyên gia triển khai cài đặt thực hiện lập báo cáo đánh giá kết quả triển khai cài đặt gói dịch vụ cần đáp ứng yêu cầu kỹ thuật triển khai.



*Quy trình quản lý kiểm thử chức năng gói dịch vụ cài đặt*

Trưởng nhóm kiểm thử có trách nhiệm phân công, sắp xếp lịch trình, giai đoạn, phân loại từng gói dịch vụ cài đặt cần xây dựng kịch bản phù hợp khả năng của mỗi chuyên gia.

Thành viên nhóm kiểm thử sau khi được phân công công việc có trách nhiệm xác định gói dịch vụ cài đặt mà mình được phân công xây dựng kịch bản kiểm thử có phù hợp với khả năng của mình hay không và báo cáo lại với trưởng nhóm kiểm thử về sự phù hợp về gói dịch vụ được phân công xây dựng kịch bản kiểm thử.

Thành viên nhóm kiểm thử phải tuân thủ quy trình chuẩn bị dữ liệu và kiểm thử chức năng gói dịch vụ sau khi triển khai cài đặt, cũng như phải báo cáo tiến độ mỗi ngày sau khi thực hiện kiểm thử với trưởng nhóm để trưởng nhóm xây dựng phương án, kế hoạch phù hợp thời gian toàn bộ gói thầu.

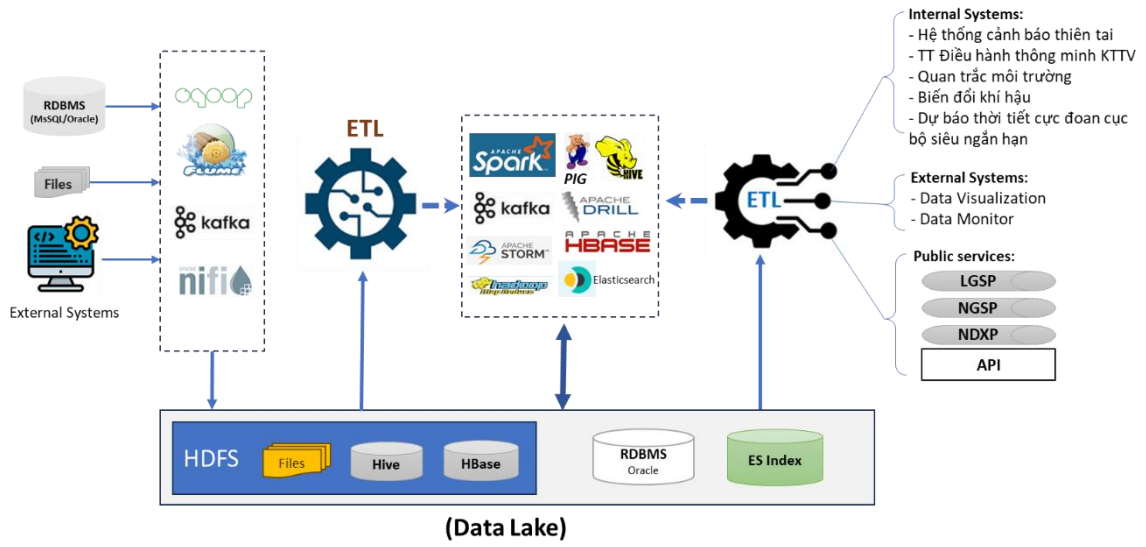
Thành viên nhóm kiểm thử phải xây dựng tài liệu sau khi kiểm thử gồm:

- Mã kịch bản kiểm thử
- Thời gian bắt đầu và kết thúc kiểm thử một kịch bản
- Các tham số đầu vào và kết quả đầu ra của dữ liệu sử dụng kiểm thử.
- Kết quả kiểm thử để kiểm tra gói dịch vụ sau khi cài đặt cần có các nội dung như: trạng thái dịch vụ hệ thống, công kết nối, thư mục cài đặt, quyền của nhóm tài khoản chạy dịch vụ,....

Sau khi kết thúc một gói cài đặt thành viên nhóm triển khai thực hiện triển khai cài đặt gói dịch vụ tiếp theo như kế hoạch cài đặt.

## ***II.2.5. Sơ đồ và thuyết minh tổng thể hệ thống***

### ***II.2.5.1. Sơ đồ luồng dữ liệu***



Dữ liệu được thu nhận vào Data Lake từ các nguồn khác nhau bao gồm các hệ thống phần mềm trong ngành (hệ thống dữ liệu qua trắc, dữ liệu vệ tinh, dữ liệu radar, dữ liệu tia sét, dữ liệu phân vùng rủi ro chuyên dùng ...), các dữ liệu đồng bộ từ CSDL hiện có, các phần mềm thu thập.

Việc thu nhận dữ liệu vào data lake được thực hiện sử dụng các công cụ trong hệ sinh thái Hadoop.

Apache Sqoop đảm nhận phần kết nối và tiếp nhận dữ liệu từ các CSDL quan hệ và lưu tập trung vào Data Lake.

Apache Flume hỗ trợ đồng bộ các loại dữ liệu phi cấu trúc và bán cấu trúc vào Data Lake

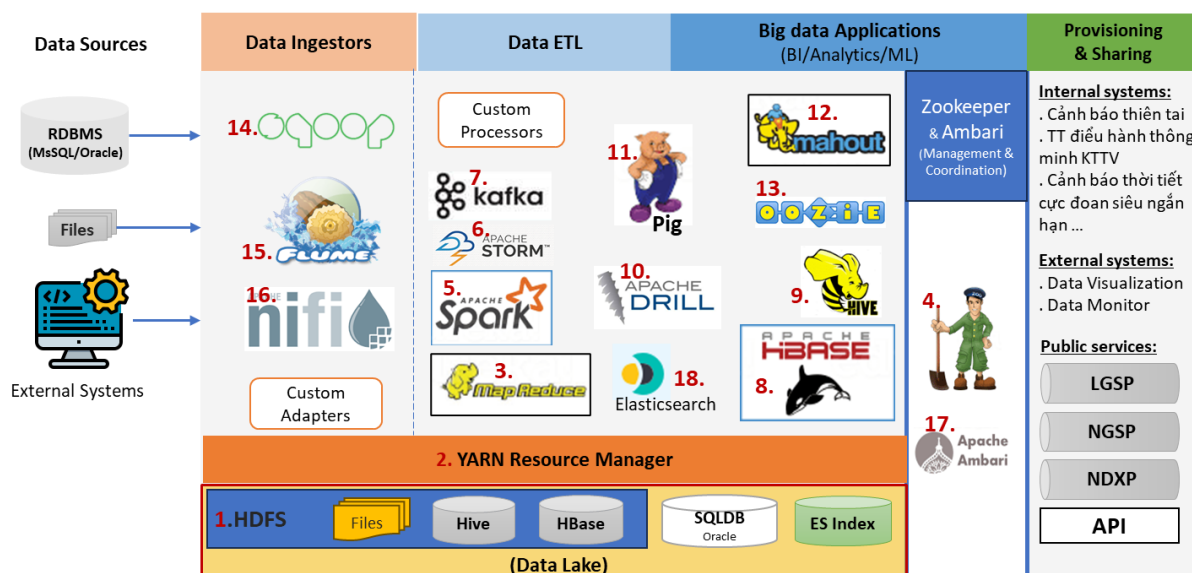
Apache NiFi với khả năng hỗ trợ kết nối tới tất cả các loại nguồn dữ liệu có thể được dùng thay thế cho Sqoop và Flume tùy theo nhu cầu xử lý. Một lý do khác chúng tôi đưa NiFi vào là do Sqoop đã được Apache tuyên bố ngừng phát triển từ tháng 6/2021. Trong trường hợp có vấn đề về tương thích trong tương lai, NiFi sẽ đảm nhiệm tốt các chức năng của Sqoop.

Sau khi dữ liệu được thu nhận vào Data Lake và lưu dưới dạng dữ liệu gốc, các module ETL sẽ thực hiện trích xuất, chuyển đổi, chuẩn hoá dữ liệu và lưu dữ liệu đã được chuẩn hoá vào Data Lake. Các framework hỗ trợ xử lý song song (Apache Spark, MapReduce), streaming (Apache Kafka, Storm) sẽ được dùng một cách phù hợp để đạt được hiệu quả xử lý tốt nhất, đảm bảo tốc độ xử lý và chất lượng dữ liệu đầu ra. Tùy theo từng loại dữ liệu gốc, dữ liệu đã được chuẩn hoá có thể được lưu vào HDFS, HBase, Oracle hay Elasticsearch index để thuận tiện cho việc cung cấp dữ liệu cho các hệ thống nghiệp vụ khai thác cũng như chia sẻ dữ liệu.

Tất cả dữ liệu trong Data Lake bao gồm dữ liệu gốc, dữ liệu đã được chuyển đổi, chuẩn hoá có thể được cung cấp, chia sẻ với các hệ thống nội bộ, hệ thống bên ngoài và các dịch vụ công. Hệ thống nội bộ (Internal systems) có thể kể đến các phần mềm khai thác như hệ thống cảnh báo thiên tai, TT Điều hành thông minh KTTV, Dự báo thời tiết cực đoan cục bộ siêu ngắn hạn, ... Các hệ thống bên ngoài (External systems) có thể là các hệ thống trực quan hoá dữ liệu (dữ liệu quan trắc mây, nhiệt, độ ẩm theo vùng) và các ứng dụng theo dõi dữ liệu. Khu dịch vụ công sẽ bao gồm trực kết nối nội bộ (LGSP), trực kết nối quốc gia (NGSP), nền tảng trao đổi dữ liệu quốc gia (NDXP), các API mở.

Việc trích xuất, tổng hợp dữ liệu từ Data Lake và đóng gói theo định dạng các chuẩn đầu ra được thực hiện với các module ETL thứ 2. Các module này có thể sẽ dùng tất cả các framework và công cụ đã được cài đặt như Apache HBase, Hive, Pig, Drill, ElasticSearch để truy xuất theo từng loại dữ liệu phù hợp.

### II.2.5.2. Kiến trúc tổng thể hệ thống



Mô hình kiến trúc tổng thể của hệ thống

Hệ thống được thiết kế theo mô hình tổng quát của hệ thống Big Data, có thể đáp ứng linh hoạt các nhu cầu từ thu nhận dữ liệu vào kho HDFS, chuyển đổi, chuẩn hoá dữ liệu và đưa vào DataLake để phục vụ các bài toán phân tích, xử lý nghiệp vụ và chia sẻ dữ liệu.

Nhìn theo trục dọc, nằm ở tầng thấp nhất là các thành phần có chức năng lưu trữ với HDFS là nền tảng. Các loại dữ liệu thô thu nhận về và các kết quả chuyển đổi, chuẩn hoá dữ liệu sẽ được lưu trữ vào các file, HMS (Hive Metastore), HBaseDB (NoSQL) trên nền HDFS. Các loại dữ liệu SQL được lưu vào hệ CSDL Postgres, các chỉ mục được quản lý bởi ElasticSearch,

Phía trên tầng lưu trữ là hệ thống quản lý tài nguyên YARN. Tiếp đó là các framework hỗ trợ phân tích, xử lý và lập trình các bài toán nghiệp vụ trên các tập dữ liệu lớn như MapReduce, Apache Spark, Kafka, Storm, Pig, Hive, Drill,...Mahout framework kết hợp với Spark Mlib hỗ trợ các bài toán về AI học máy.

Nhìn theo trục ngang (theo góc độ luồng xử lý dữ liệu), hệ thống hỗ trợ các công cụ đáp ứng linh hoạt các nhu cầu từ khi nhận dữ liệu từ các nguồn khác nhau với các định dạng khác nhau (như dữ liệu có cấu trúc, phi cấu trúc và bán cấu trúc). Thông thường các dữ liệu phi cấu trúc và bán cấu trúc sẽ được nhập vào hệ thống Data Lake thông qua Apache Flume. Các dữ liệu có cấu trúc sẽ do Apache Sqoop thực hiện việc nhập, xuất dữ liệu. Apache NiFi có thể đảm nhận tốt chức năng của cả Sqoop và Flume và có thể sử dụng khi cần. Trong trường hợp đặc biệt không thể dùng các framework có sẵn để thu nhận dữ liệu theo đúng yêu cầu, các Custom Adapter có thể được lập trình sử dụng các framework hỗ trợ xử lý như Spark, Kafka, Storm.

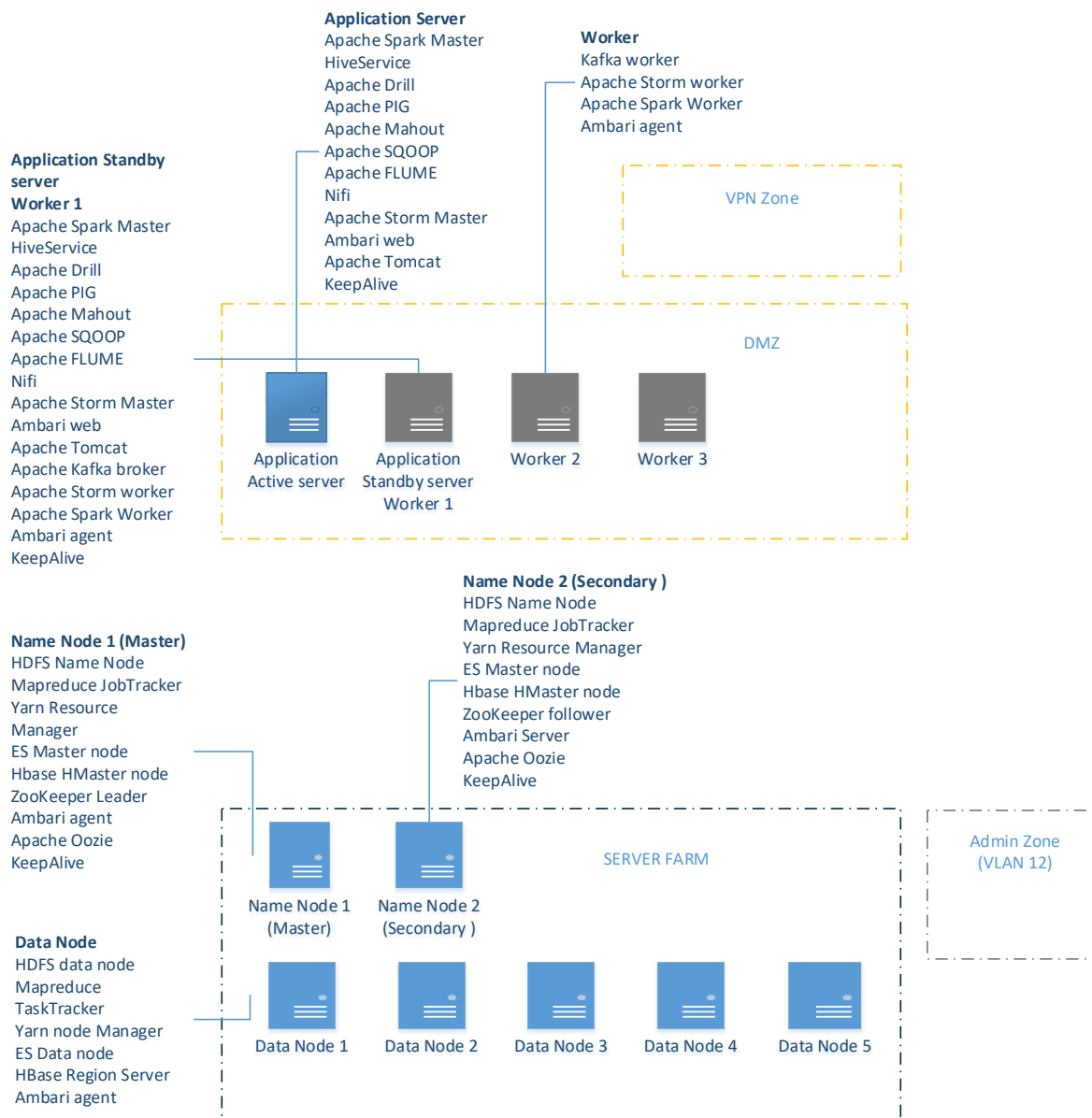
Dữ liệu từ các nguồn (data sources) sau khi được thu nhận và lưu vào HDFS sẽ được xử lý, chuẩn hoá trong giai đoạn Data ETL. Các dữ liệu đã được chuẩn hoá có thể được lưu thành dữ liệu có cấu trúc (SQL) hoặc NoSQL hoặc đánh chỉ mục tùy theo loại dữ liệu và nhu cầu sử dụng của các bài toán nghiệp vụ sau này. Các Custom Processor có thể được phát triển riêng theo yêu cầu chuyển đổi, chuẩn hoá dữ liệu nếu các framework có sẵn không thể xử lý.

Tùy tính chất của từng loại dữ liệu mà chúng có thể được xử lý song song hoặc tuần tự theo kiểu streaming. Hệ thống có sẵn sàng các framework hỗ trợ xử lý song song (Spark) và stream (Kafka & Storm) cũng như các framework làm việc với dữ liệu lớn trên HDFS như Hive, Hbase, Pig, Drill.

Trong trường hợp cần sử dụng dữ liệu cho các bài toán AI, Deep learning, hệ thống cũng sẵn sàng với Mahout và Spark Mllib.

Dữ liệu đã qua xử lý, chuẩn hoá, thậm chí cả các dữ liệu thô và các đầu ra của bài toán nghiệp vụ có thể được tích hợp vào trục kết nối nội bộ (LGSP) cũng như chia sẻ ra bên ngoài qua các kênh APIs.

### ***II.2.5.3. Sơ đồ tổng quan triển khai hệ thống***



Hệ thống phân chia 4 phân vùng mạng nội bộ (VLAN) để đảm bảo tính độc lập giữa các phân vùng. Khi một phân vùng có sự cố về mạng không bị ảnh hưởng tới phân vùng còn lại và hạn chế các truy cập không cần thiết giữa các phân vùng. Các phân vùng được kết nối với nhau thông qua hạ tầng mạng, thiết bị an ninh thông tin đảm bảo an toàn thông tin của Trung tâm thông tin và dữ liệu KTTV.

- **Phân vùng Database (Server Farm Zone)**

Phân vùng Database quy hoạch cho toàn bộ các ứng dụng lưu trữ dữ liệu, cần đảm bảo tính bảo mật cao vì vậy cần hạn chế tối đa các truy cập trái phép hoặc không cần thiết từ các phân vùng khác, chỉ cho phép các truy cập cần thiết từ phân vùng ứng dụng và phân vùng vận hành để phục vụ vận hành hệ thống.

- **Phân vùng ứng dụng (DMZ Zone)**

Phần vùng ứng dụng quy hoạch cho toàn bộ các server chạy các ứng dụng cơ sở cho người dùng hoặc các ứng dụng lập trình xử lý, tính toán dữ liệu. Phần vùng này cho phép truy cập từ bên ngoài bởi người dùng, từ phân vùng VPN, phân vùng Admin để truy cập vận hành các ứng dụng.

- ***Phần vùng vận hành (Admin zone)***

Phân vùng vận hành dành riêng cho việc truy cập cài đặt, vận hành hệ thống, từ phân vùng này có thể truy xuất tới tất cả các phân vùng.

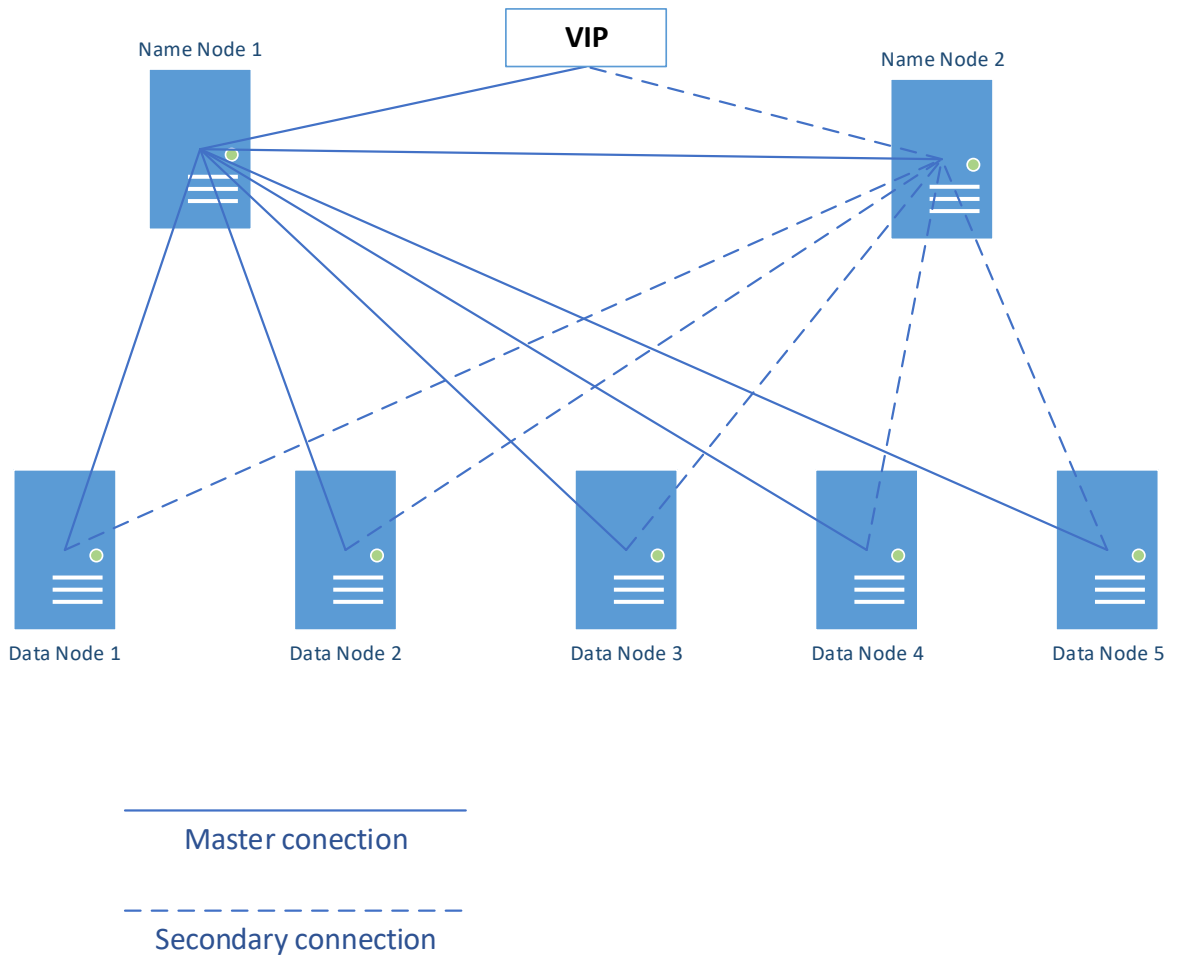
- ***Phân vùng VPN (VPN Zone)***

Phân vùng dành riêng cho việc truy cập VPN từ ngoài vào hệ thống. Từ phân vùng này có thể truy cập vào phân vùng ứng dụng để truy xuất vận hành các ứng dụng.

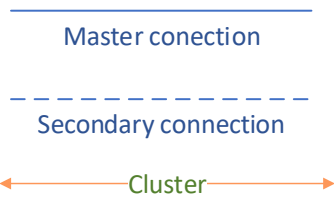
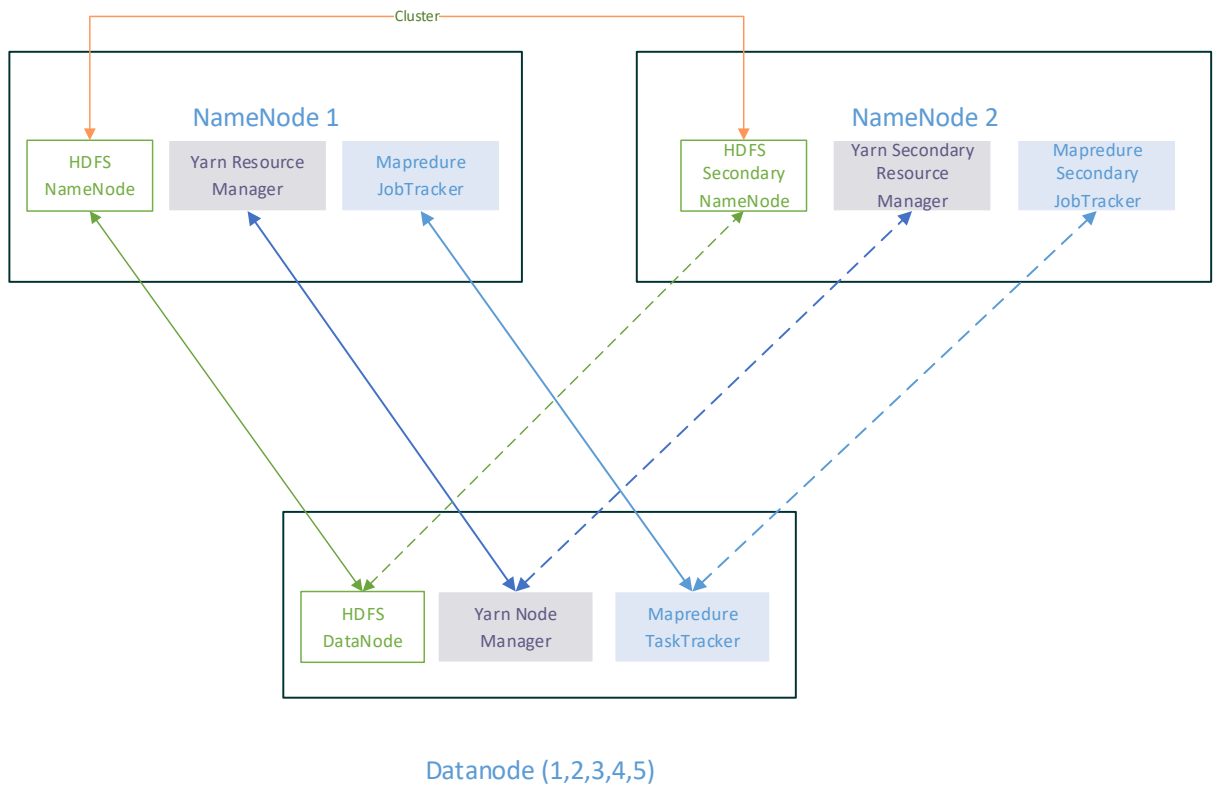
Nếu cần thiết thì bố trí một máy ảo có chức năng như một máy chủ Jumper và từ máy chủ này mới có thể truy cập vào các máy chủ hệ thống bên trong thông qua một ứng dụng quản tập trung đơn cử như: Password Manager Pro.

#### ***II.2.5.4. Kết nối Hadoop Cluster Architecture***

- ***Sơ đồ kết nối máy chủ Hadoop***



- *Sơ đồ kết nối ứng dụng Hadoop Cluster*



Tên ứng dụng	Dịch vụ	Port	Tên máy	Giao thức
<b>HDFS</b>	HDFS Master Name Node		Name Node 1	TCP
	HDFS Secondary Name Node		Name Node 2	TCP
	HDFS Data node		Datanode 1 Datanode 2 Datanode 3 Datanode 4 Datanode 5	TCP

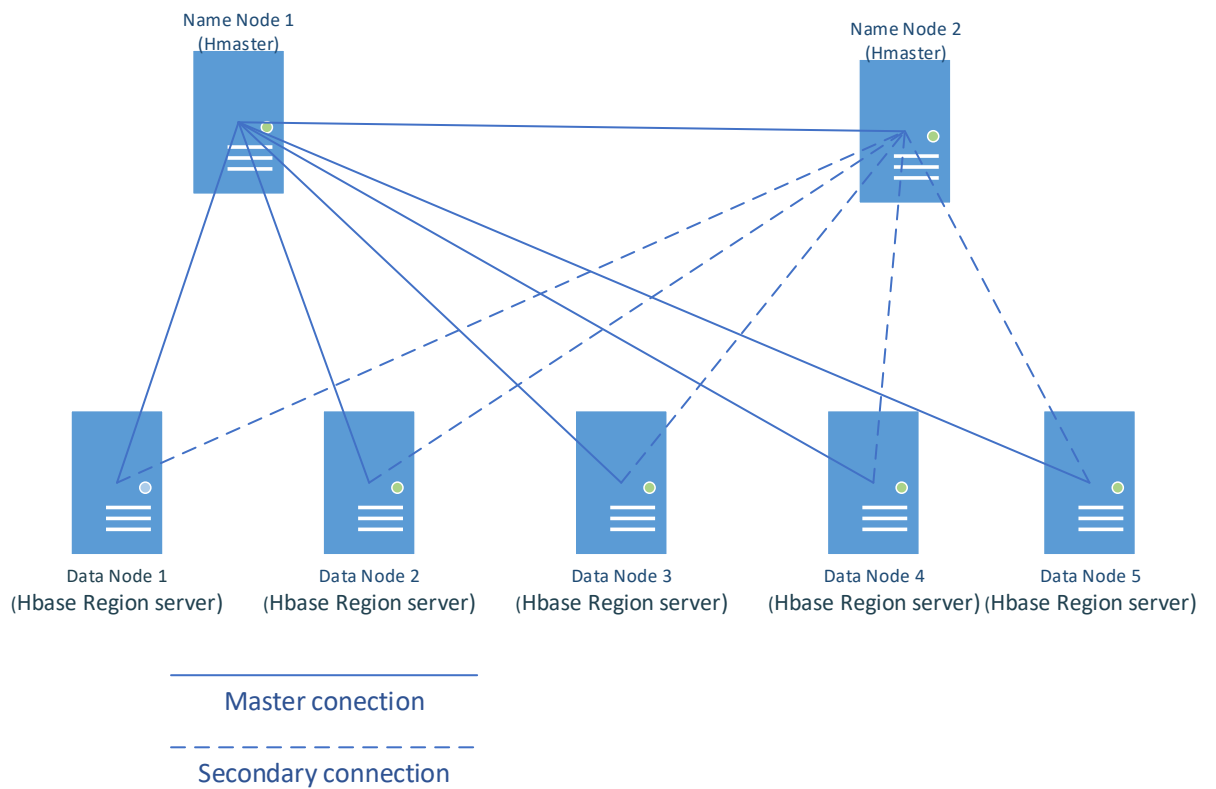
	Web UI of the Resource Manager		Name Node 1	HTTP
	Web UI of the NameNode daemon		Name Node 1	HTTP
	Web UI of the DataNode daemon		Name Node 1	HTTP
	HDFS NameNode web interface		Name Node 1	HTTP
<b>Mapreduce</b>	Mapreduce JobTracker		Datanode 1 Datanode 2 Datanode 3 Datanode 4 Datanode 5	TCP
	Mapreduce TaskTracker			TCP
<b>Yarn</b>	Yarn Resource Manager		Name Node 1 Name Node 2	TCP
	Yarn node Manager		Datanode 1 Datanode 2 Datanode 3 Datanode 4 Datanode 5	TCP

Ứng dụng Hadoop chạy trên 2 namenode và 5 datanode.

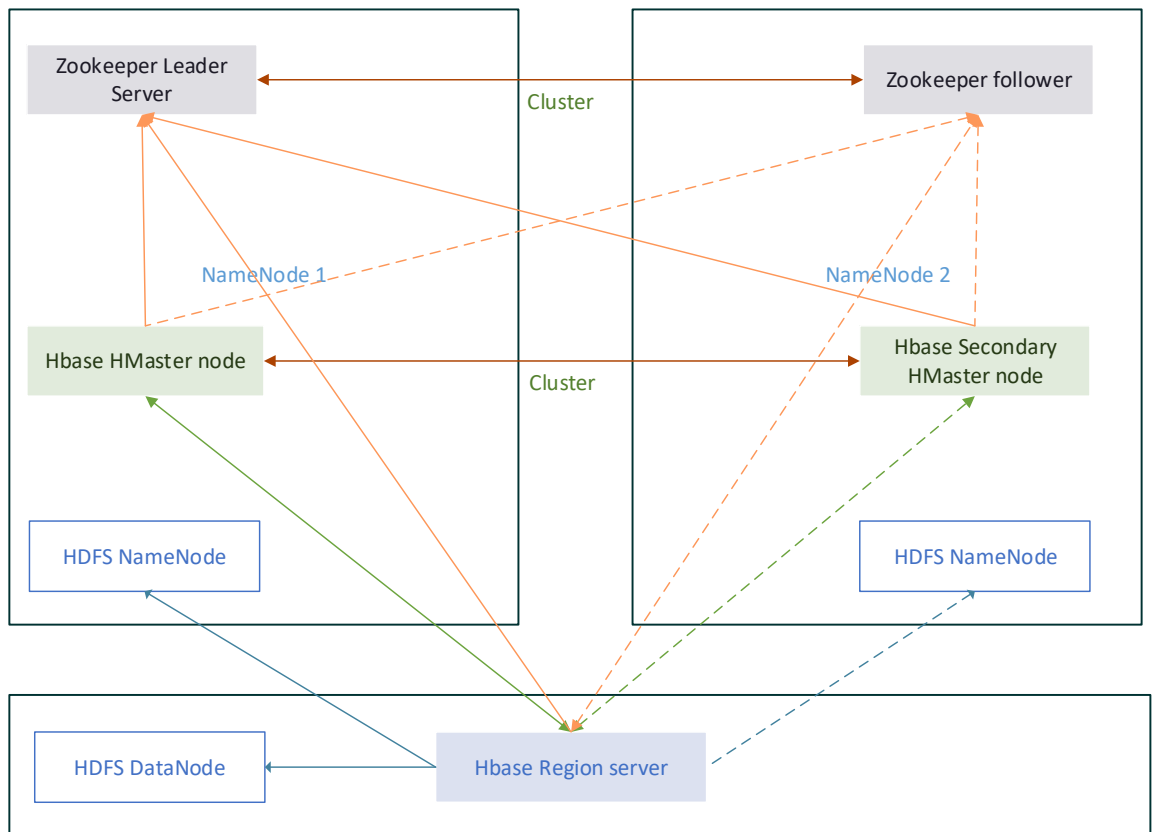
Namenode 1 và Namenode 2 chạy chế độ Cluster - Failover

### II.2.5.5. Kết nối Hbase Cluster

#### - Sơ đồ kết nối máy chủ Hbase



#### - Sơ đồ kết nối ứng dụng Hbase



Datanode (1,2,3,4,5)



Hbase chạy chế độ cluster, Hai server Hmaster false-over tự động.

Hbase Region server lưu trữ tệp data trên HDFS đảm bảo khả năng dự phòng cao khi có 1 node bị false sẽ tự động được điều tiết dữ liệu từ node khác.

Khi node hoạt động trở lại, hoặc thêm mới thì dữ liệu tại các node này sẽ dần được sử dụng cân bằng.

Để thêm node mới cần thực hiện:

- Sao chép cấu hình của Hbase đến node mới
- Sao chép cấu hình của ZooKeeper đến node mới
- Khởi động Hbase Region Server

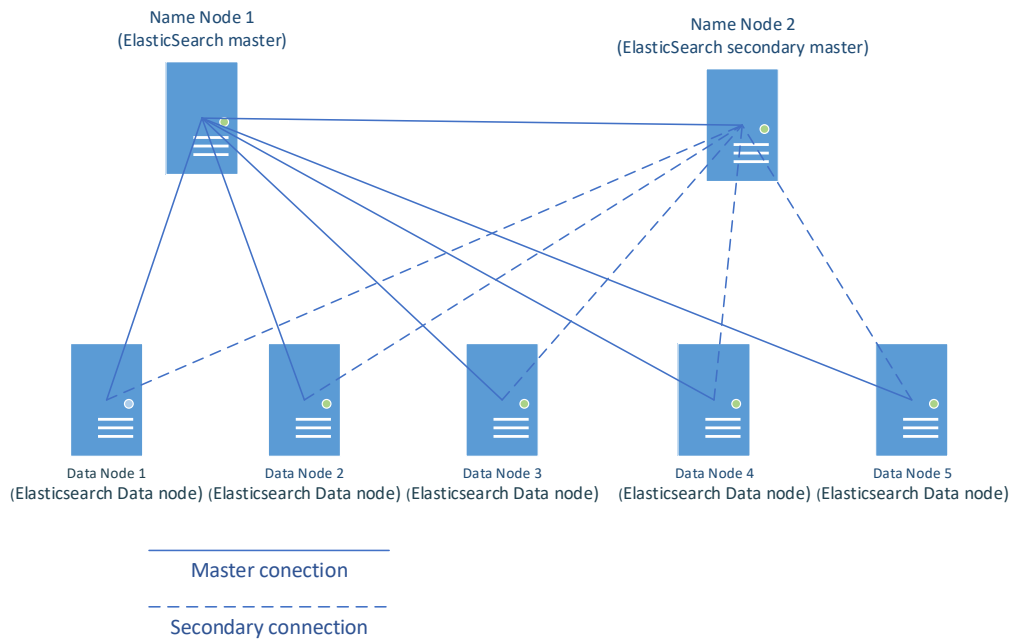
- Thêm hostname của node mới vào cấu hình Hbase trên Hbase Master node và reload cấu hình.

<b>Tên ứng dụng</b>	<b>Dịch vụ</b>	<b>Port</b>	<b>Tên máy</b>	<b>Giao thức</b>
<b>Hbase</b>	Hbase HMaster node		Name Node 1	TCP
	HMaster Info Web UI		Name Node 1	HTTP
	Hbase rest port		Name Node 1	HTTP
	Hbase rest info port		Name Node 1	HTTP
	hbase thrift info port		Name Node 1	TCP
	Hbase Secondary HMaster node		Name Node 2	TCP
	Hmaster Secondary Info Web UI		Name Node 2	HTTP
	Hbase Secondary rest port		Name Node 2	HTTP
	Hbase Secondary rest info port		Name Node 2	HTTP
	hbase Secondary thrift info port		Name Node 2	TCP
	HBase Region Server 1		Datanode 1	TCP
	HBase Region Server 2		Datanode 2	TCP
	HBase Region Server 3		Datanode 3	TCP

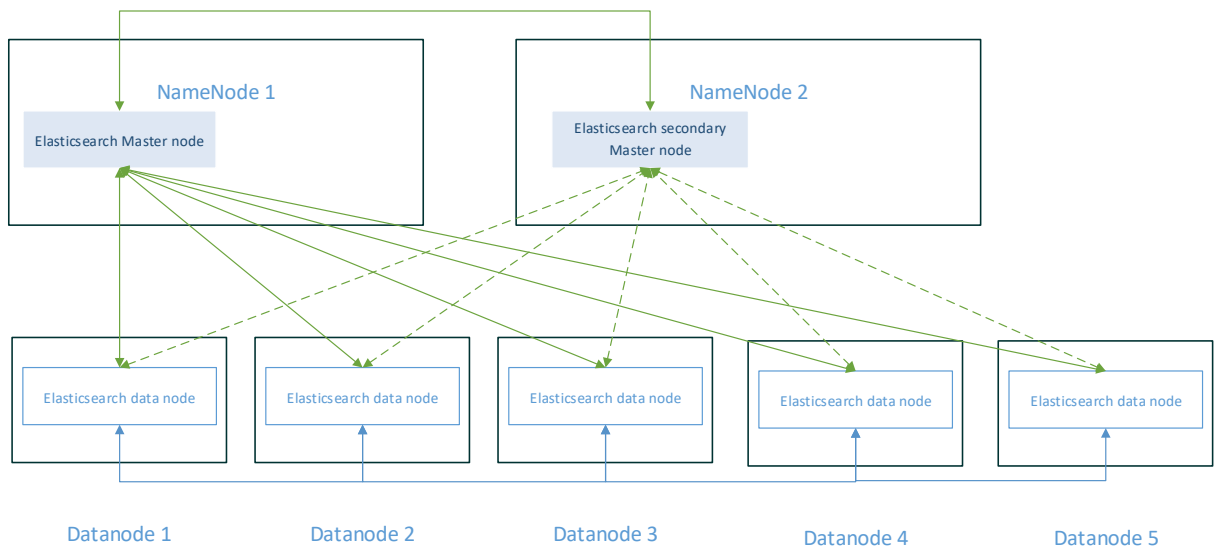
HBase Server 4	Region		Datanode 4	TCP
HBase Server 5	Region		Datanode 5	TCP
Hbase server thrift port	region		Datanode 1 Datanode 2 Datanode 3 Datanode 4 Datanode 5	TCP

### II.2.5.6. Kết nối Elasticsearch Cluster

- Sơ đồ kết nối máy chủ Elasticsearch



- Sơ đồ kết nối ứng dụng Elasticsearch



Elasticsearch chạy chế độ Cluster bao gồm 2 Master node và 5 datanode.

Khi Masternode chính ngừng hoạt động thì Secondary Master node sẽ tự động được các node còn lại bầu làm Master node.

Khi một datanode ngừng hoạt động thì dữ liệu từ datanode khác (repleca) sẽ được làm dữ liệu chính (Primary) nhờ vậy sẽ bảo toàn được dữ liệu.

Cluster có khả năng mở rộng theo chiều ngang bằng cách thêm các datanode vào cluster.

Các Masternode và Datanode kết nối với nhau thông qua TCP Transport.

Tên ứng dụng	Dịch vụ	Port	Tên máy	Giao thức
<b>Elastic Search</b>	ES Transport		NameNode 1 NameNode 1 Datanode 1 Datanode 2 Datanode 3 Datanode 4 Datanode 5	TCP

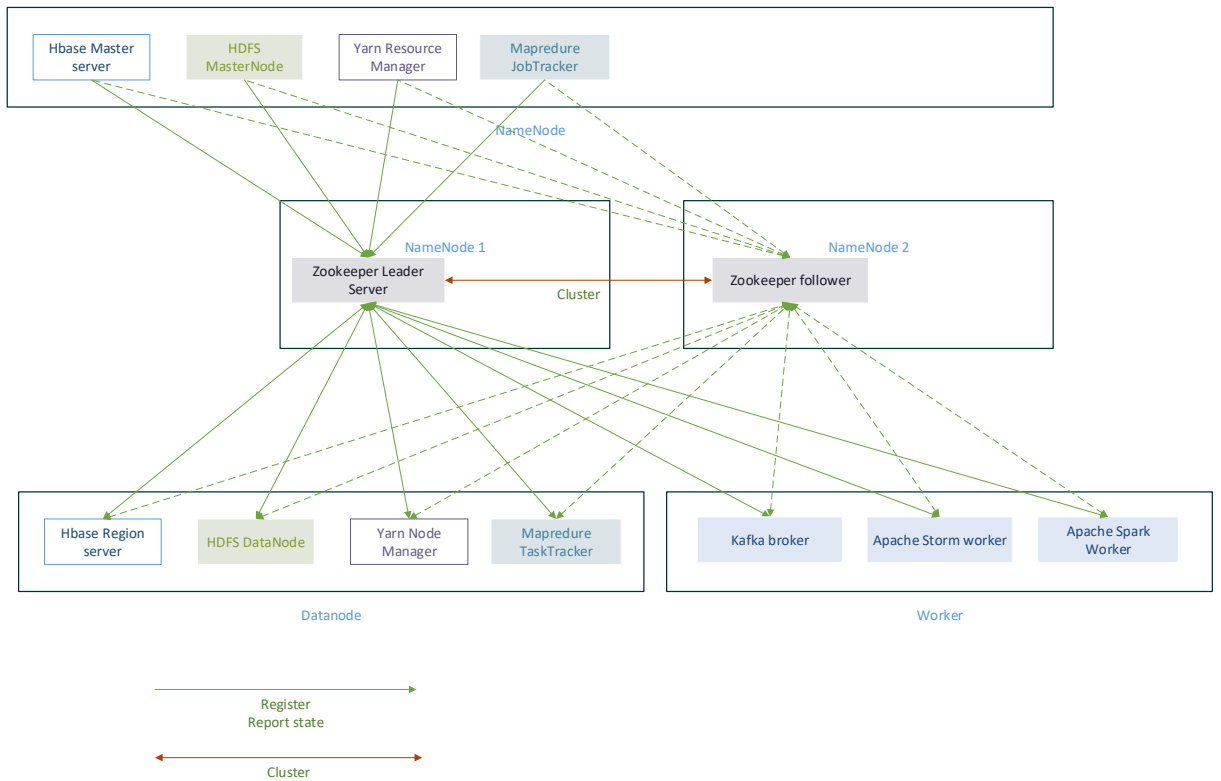
			NameNode 1 NameNode 1 NameNode 1 Datanode 1 Datanode 2 Datanode 3 Datanode 4 Datanode 5	
	API	ES REST		HTTP

### II.2.5.7. Kết nối Zookeeper

#### - Sơ đồ kết nối máy chủ Zookeeper



#### - Sơ đồ kết nối ứng dụng Zookeeper



Zookeeper hoạt động chế độ Cluster. Ở điều kiện hoạt động bình thường một node sẽ được hoạt động ở chế độ Leader, nhận các đăng ký và báo cáo trạng thái của client. Các node khác sẽ hoạt động ở chế độ follower.

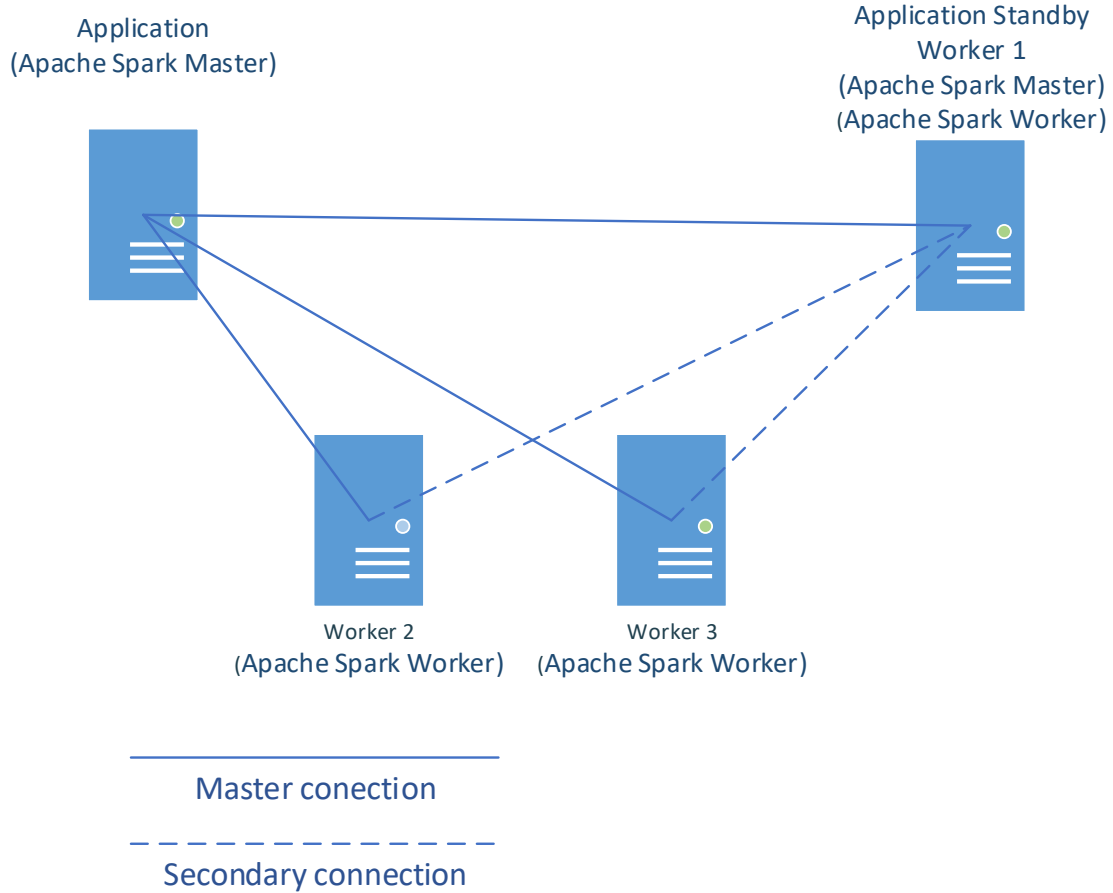
Khi một node ngừng hoạt động thì một node còn lại sẽ được bầu làm Leader.

Zookeeper mở kế nối cho các Client kết nối đến bao gồm các ứng dụng: Hbase, HDFS, Yarn, Mapredure, Kafka, Apache storm, Apache Spark.

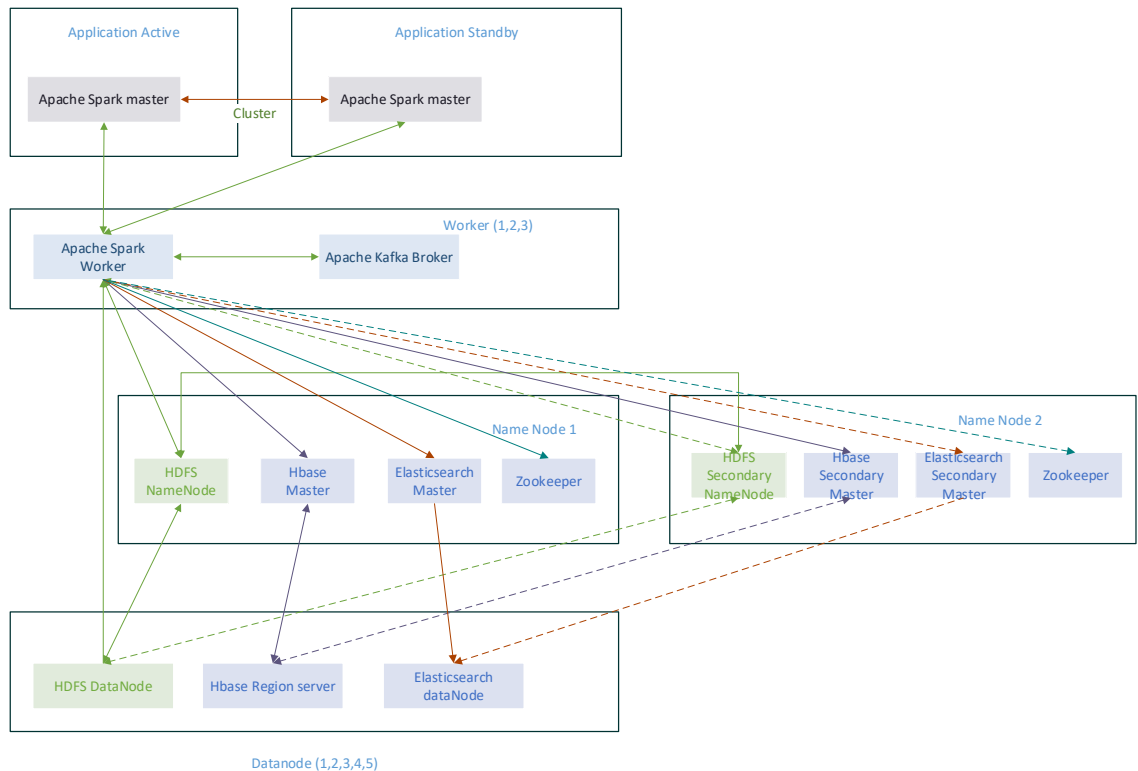
Tên ứng dụng	Dịch vụ	Port	Tên máy	Giao thức
Zookeeper	Zookeeper Client connect		Name Node 1	TCP
	Zookeeper admin connect		Name Node 1	TCP

#### II.2.5.8. Kết nối Apache Spark

- Sơ đồ kết nối máy chủ



- *Sơ đồ kết nối ứng dụng Apache Spark*



Apache Spark hoạt động dưới chế độ Cluster với 2 Master và 3 worker. Một Master node sẽ hoạt động ở chế độ Active và các Master node còn lại hoạt động ở chế độ Passive (Standby).

Khi một Master node ngừng hoạt động thì Master node Passive sẽ được chuyển sang chế độ Active giúp cho hệ thống hoạt động bình thường.

Khi một worker node ngừng hoạt động thì các Task/Job sẽ được điều hướng sang worker node khác hoạt động thay thế.

Khi worker node hoạt động trở lại thì các Task/Job sẽ được Master tiếp tục điều phối tới.

Các Apache Spark worker có thể kết nối tới các thành phần khác như Apache Kafka queue, HDFS, Hbase, Elasticsearch để xử lý dữ liệu.

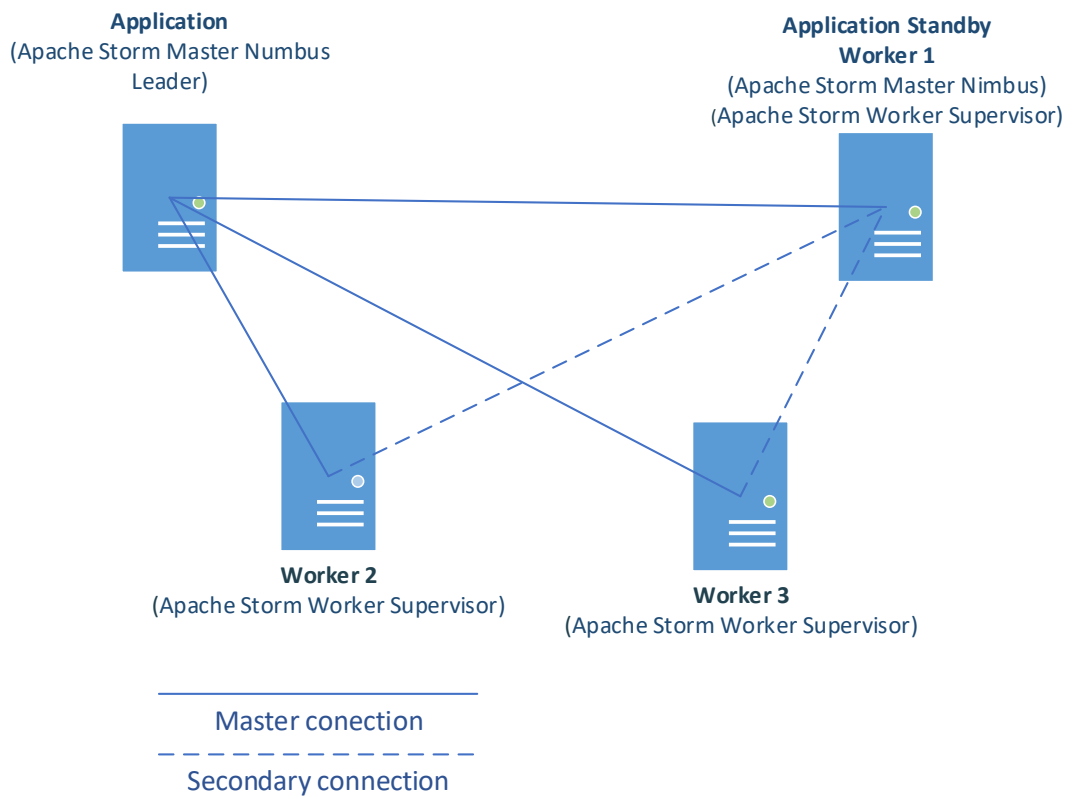
Khi thêm node mới thì cần cấu hình lại SPARK\_MASTER\_HOST trong file spark-env.sh trên node mới. Sau đó khởi động lại spark trên node đó thì nó sẽ trở thành một node trong cụm Spark

Tên ứng dụng	Dịch vụ	Port	Tên máy	Giáo thức
Apache Spark	Apache Spark cluster Manager		Application	TCP

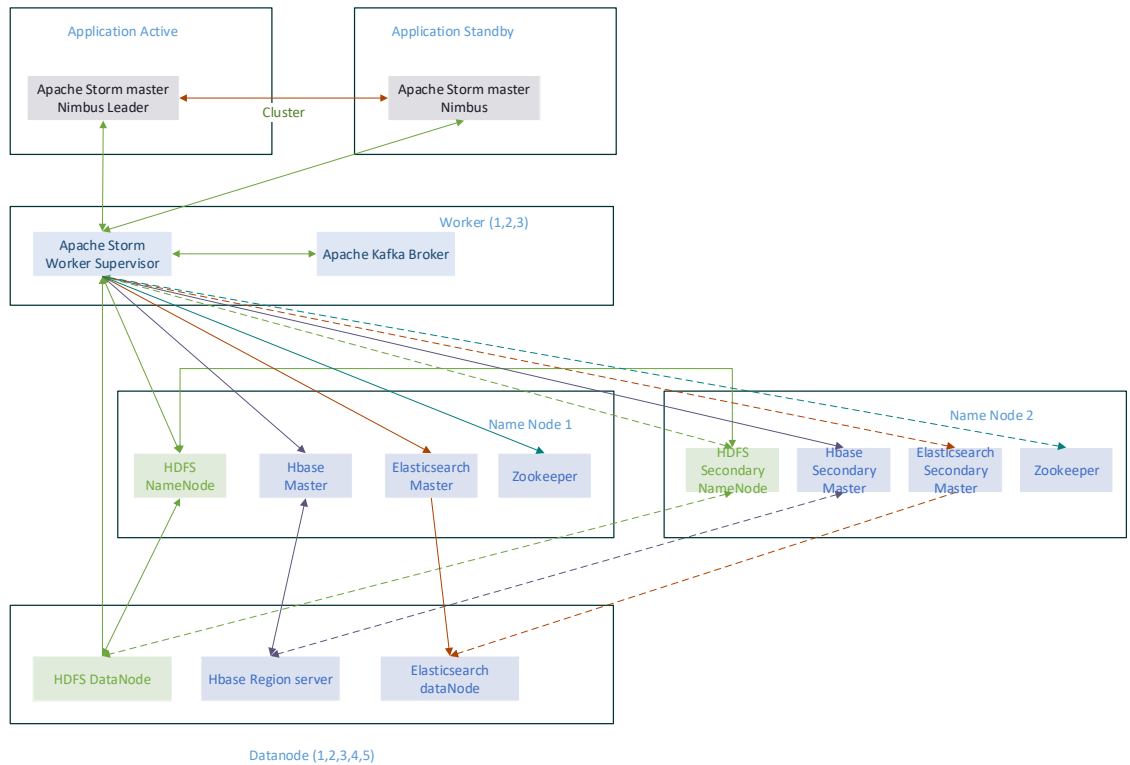
	Apache Master web UI	Spark		Application	P	HTT
	Apache worker web UI	Spark		Application	P	HTT
	Apache Worker	Spark		er	Work	TCP

### II.2.5.9. Kết nối Apache Storm

- Sơ đồ kết nối máy chủ Apache Storm



- Sơ đồ kết nối ứng dụng Apache Storm



Apache Storm hoạt động dưới chế độ Cluster. Với 2 Storm Master node và 3 Storm worker node.

Ở chế độ bình thường 1 Master node sẽ hoạt động chế độ Leader và các master node khác sẽ hoạt động ở chế độ Not Leader.

Khi Master node leader ngừng hoạt động thì một master node khác sẽ tự động được bầu làm Master node leader và hệ thống hoạt động bình thường.

Khi một storm worker ngừng hoạt động thì dữ liệu/Task sẽ được phân phối tới các worker khác.

Khi worker hoạt động trở lại thì dữ liệu/Task tiếp tục được phân phối tới và trở lại hoạt động bình thường.

Các Storm worker có thể kết nối tới các thành phần khác như Apache Kafka queue, HDFS, Hbase, Elasticsearch để xử lý dữ liệu.

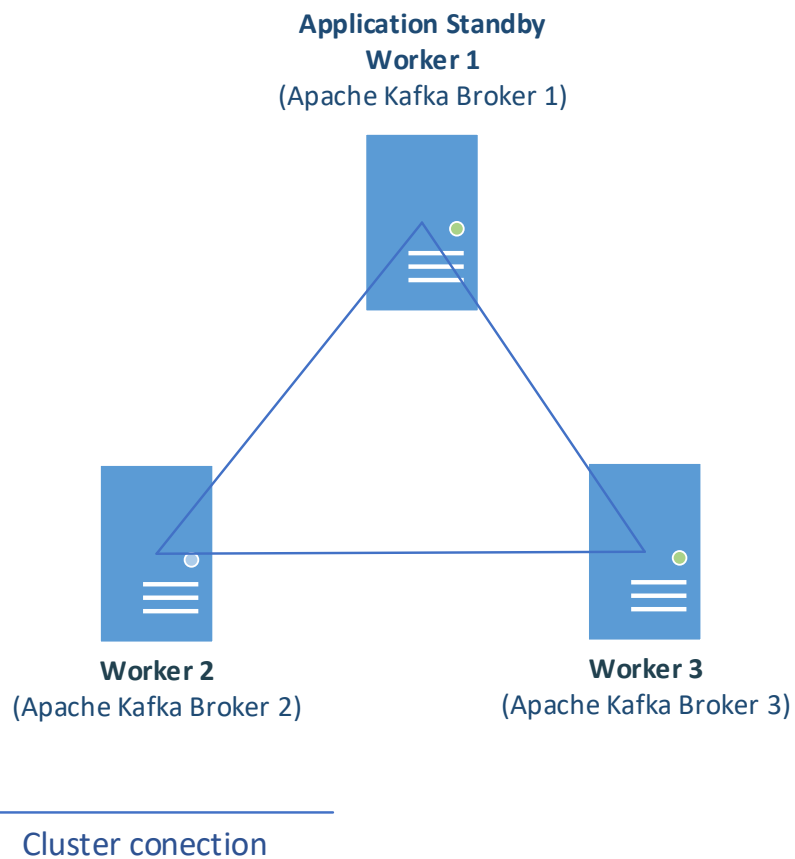
Khi thêm node mới cần cài Storm worker và sửa file storm.yaml trên node đó giống với node hiện tại và reload lại worker

Tên ứng dụng	Dịch vụ	Port	Tên máy	Giao thức
Apache Storm	Apache Storm Master node Nimbus leader		Application	TCP

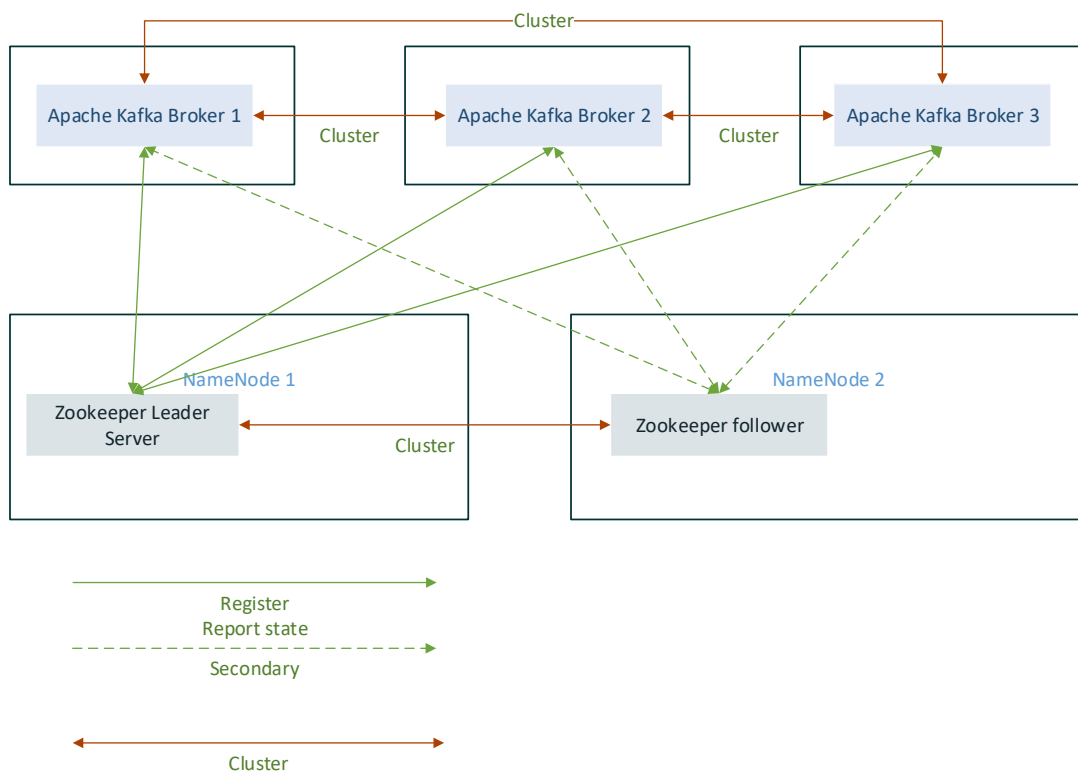
	Apache Master node Nimbus	Storm		Application standby	TC P
	Apache worker Supervisor 1	Storm		Worker 1	TC P
	Apache worker Supervisor 2	Storm		Worker 2	TC P
	Apache worker Supervisor 3	Storm		Worker 3	TC P
	Apache WebUI	Storm		Application	H TTP

### II.2.5.10. Kết nối Apache Kafka

#### - Sơ đồ kết nối máy chủ Apache Kafka



#### - Sơ đồ kết nối ứng dụng Apache Kafka



Apache Kafka queue hoạt động dưới chế độ cluster bao gồm 3 node broker 1, broker 2, broker 3. Các broker kết nối trực tiếp với nhau và dữ liệu được phân bổ phân tán tới các partition primary và replica trên các broker. Khi một broker ngưng hoạt động thì dữ liệu ở các partition replica của broker khác sẽ được sử dụng thay thế đảm bảo hệ thống hoạt động ổn định và dữ liệu tin cậy.

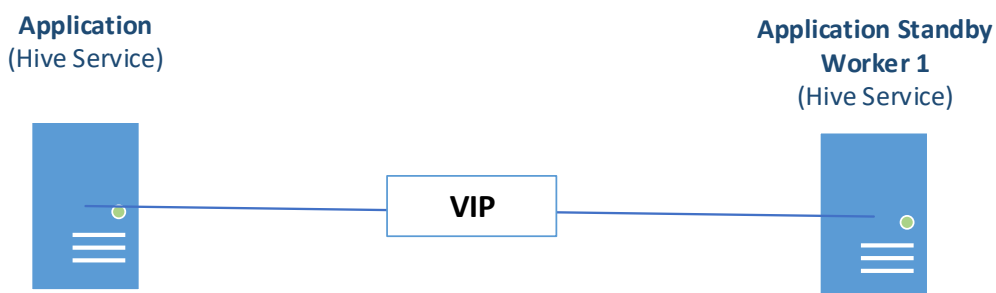
Apache Kafka kết nối tới Zookeeper để quản lý cấu hình các node.

Khi thêm node mới, cần cài đặt Apache Kafka. Sau đó cấu hình broker.id cho node trong file server.properties của Kafka. Tiếp theo cần cấu hình lại file server.properties trong các node khác trong cụm với trường advertised.listeners là IP address của node mới. Cuối cùng, start lại node mới thêm.

Tên ứng dụng	Dịch vụ	Port	Tên máy	Giao thức
Kafka	Kafka Brocker 1		Worker 1	TCP
	Kafka Brocker 2		Worker 2	TCP
	Kafka Brocker 3		Worker 3	TCP

### II.2.5.11. Kết nối Apache Hive

#### - Sơ đồ kết nối máy chủ Apache Hive

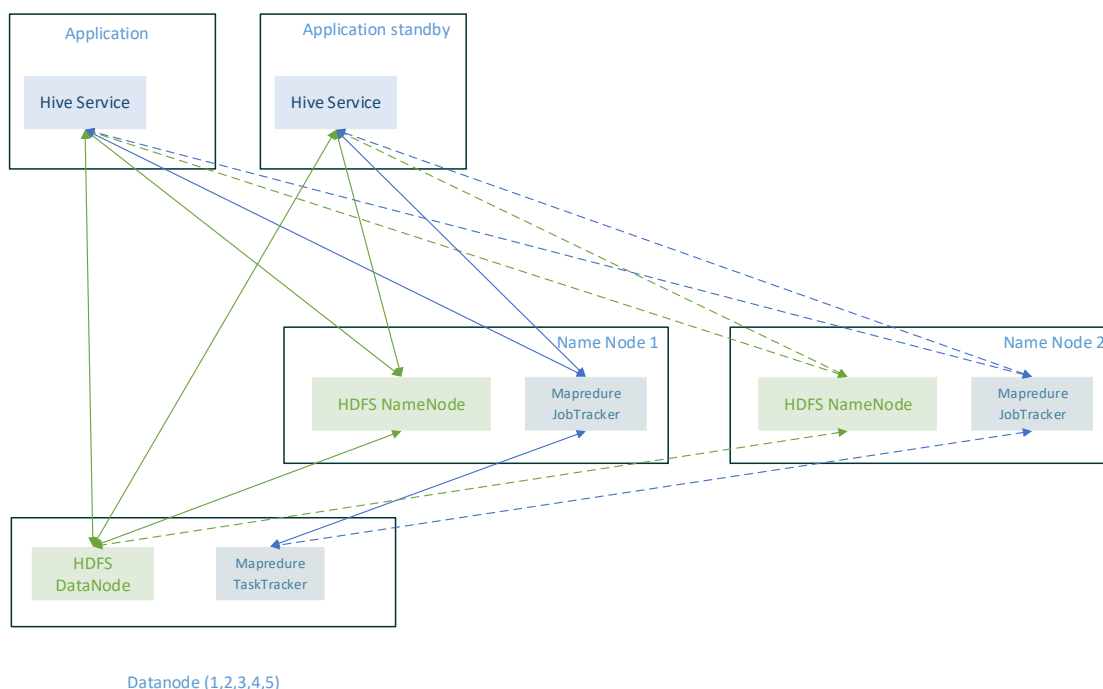


Hai server Application và Application Standby kết nối heartbeat với nhau qua ứng dụng KeepAlive để duy trì VIP cho một server.

Ở điều kiện bình thường server Application sẽ được gán VIP và mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application.

Khi Server Application ngừng hoạt động thì VIP sẽ tự động được gán sang Application Standby, lúc này mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application standby cho tới khi Application server hoạt động ổn định trở lại.

#### - Sơ đồ kết nối ứng dụng Apache Hive



Apache Hive hoạt động độc lập trên 2 server Application và Application Standby.

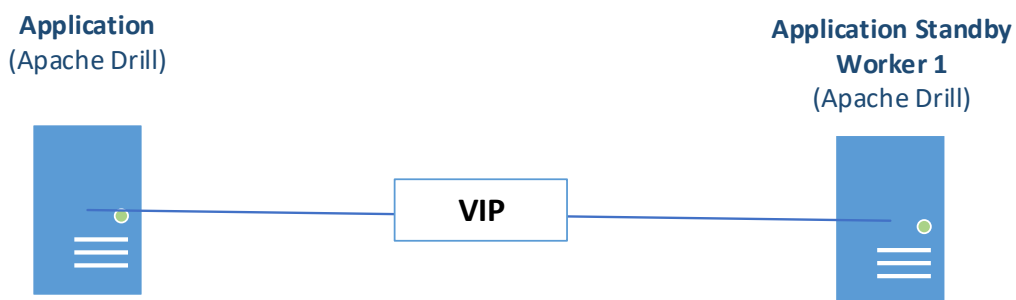
Khi một trong hai server gặp sự cố thì server còn lại sẽ được điều hướng request tới thông qua VIP đảm bảo không bị gián đoạn dịch vụ.

Hive Service được mở các kết nối tới các ứng dụng HDFS, Mapreduve để phục vụ truy vấn dữ liệu của Hadoop.

Tên ứng dụng	Dịch vụ	Port	Tên máy	Giao thức
<b>Hive</b>	hiveserver thrift port		Application Application Standby	TCP
	hiveserver thrift http port		Application Application Standby	TCP
	hiveserver webui port		Application Application Standby	HTTP
	Hive Metastore db		Application Application Standby	TCP

### II.2.5.12. Kết nối Apache Drill

#### - Sơ đồ kết nối máy chủ Apache Drill



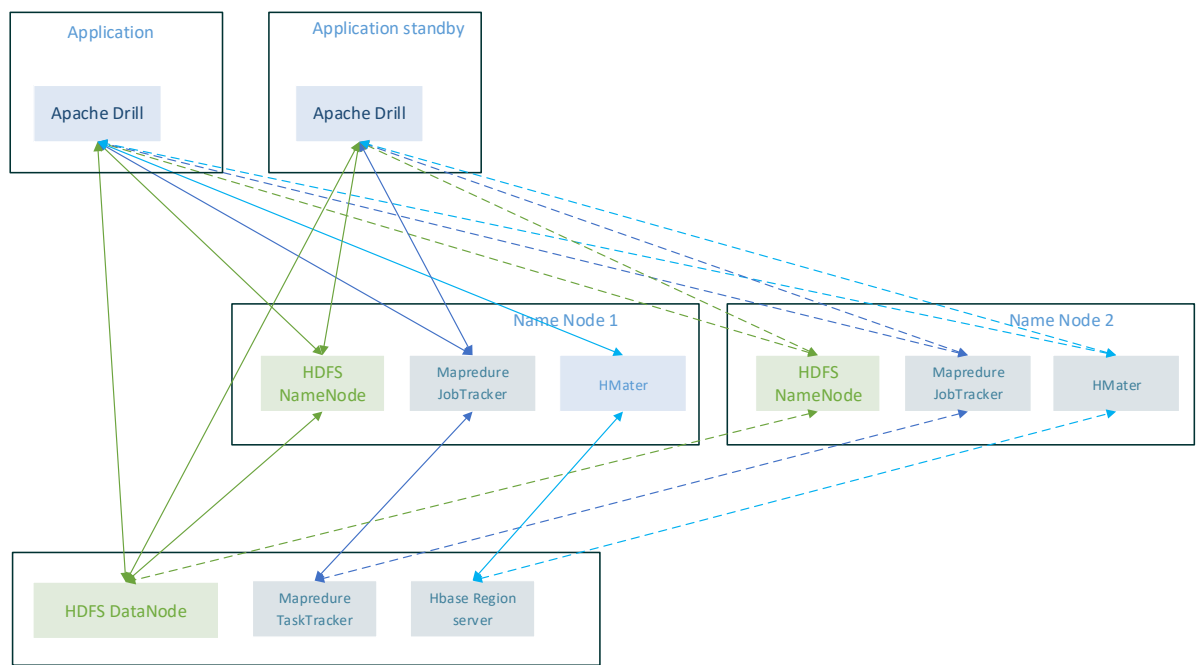
Hai server Application và Application Standby kết nối heartbeat với nhau qua ứng dụng KeepAlive để duy trì VIP cho một server.

Ở điều kiện bình thường server Application sẽ được gán VIP và mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application.

Khi Server Application ngừng hoạt động thì VIP sẽ tự động được gán sang Application Standby, lúc này mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application standby cho tới khi Application server hoạt động ổn định trở lại.

Khi thêm node mới thì Apache Drill sẽ được tự cài đặt

**- Sơ đồ kết nối ứng dụng Apache Drill**



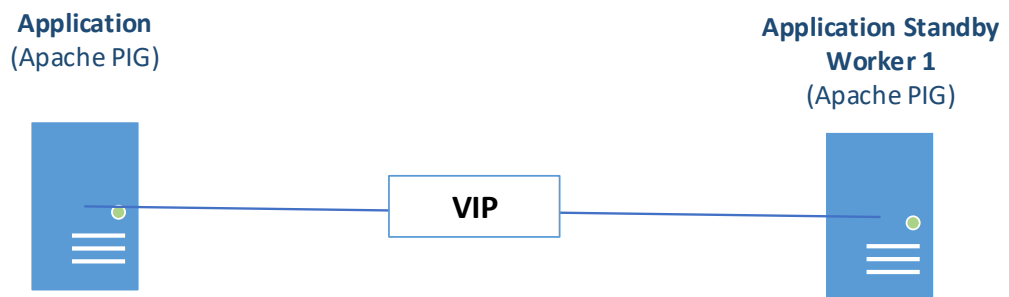
Datanode (1,2,3,4,5)

Tên ứng dụng	Dịch vụ	Port	Tên máy	Giao thức
Apache Drill	Drill Web UI		Application Application Standby	HTTP
	Drill exec.rpc.user.server.port		Application	TCP

			Application Standby	
	Drill exec.rpc.bit.server.port		Application Application Standby	TCP
	Drill exec.rpc.bit.server.port + 1		Application Application Standby	TCP

### II.2.5.13. Kết nối Apache PIG

#### - Sơ đồ kết nối máy chủ Apache PIG

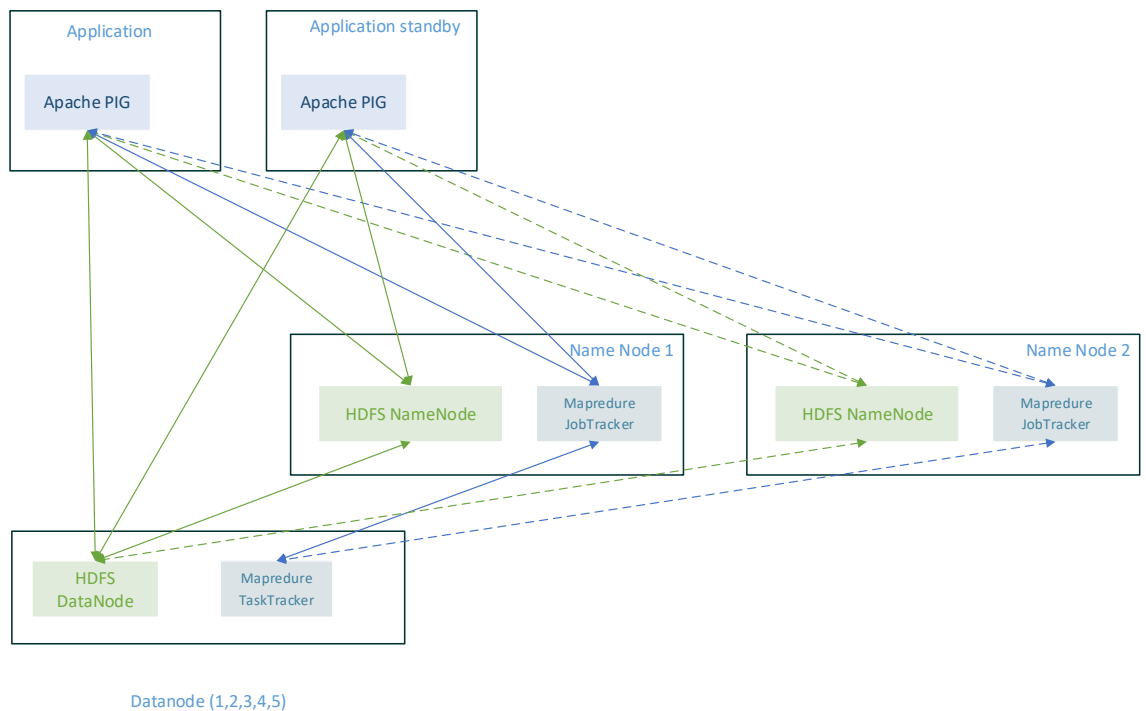


Hai server Application và Application Standby kết nối heartbeat với nhau qua ứng dụng KeepAlive để duy trì VIP cho một server.

Ở điều kiện bình thường server Application sẽ được gán VIP và mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application.

Khi Server Application ngừng hoạt động thì VIP sẽ tự động được gán sang Application Standby, lúc này mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application standby cho tới khi Application server hoạt động ổn định trở lại.

#### - Sơ đồ kết nối ứng dụng Apache PIG



Ứng dụng Apache PIG được cài đặt trên 2 máy chủ Application và Application Standby.

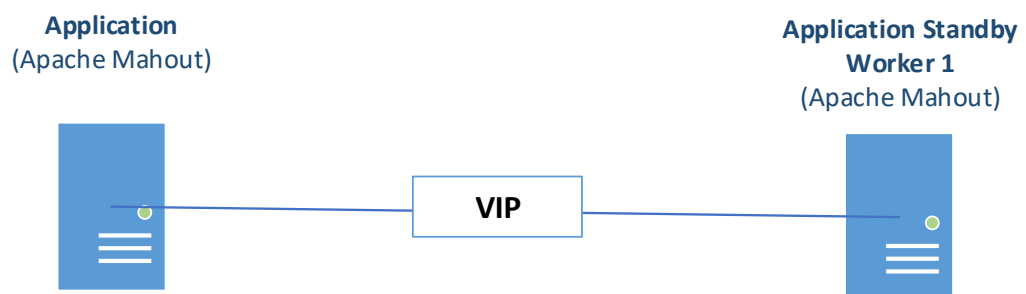
Apache PIG hoạt động độc lập trên 2 server Application và Application Standby.

Khi một trong hai server gặp sự cố thì server còn lại sẽ được điều hướng request tới thông qua VIP đảm bảo không bị gián đoạn dịch vụ.

Mahout được mở các kết nối tới các ứng dụng bên trong như HDFS, Mapredure để xử lý các tác vụ và dữ liệu trong hệ thống Hadoop, Hbase.

#### II.2.5.14. Kết nối Apache Mahout

##### - Sơ đồ kết nối máy chủ Apache Mahout

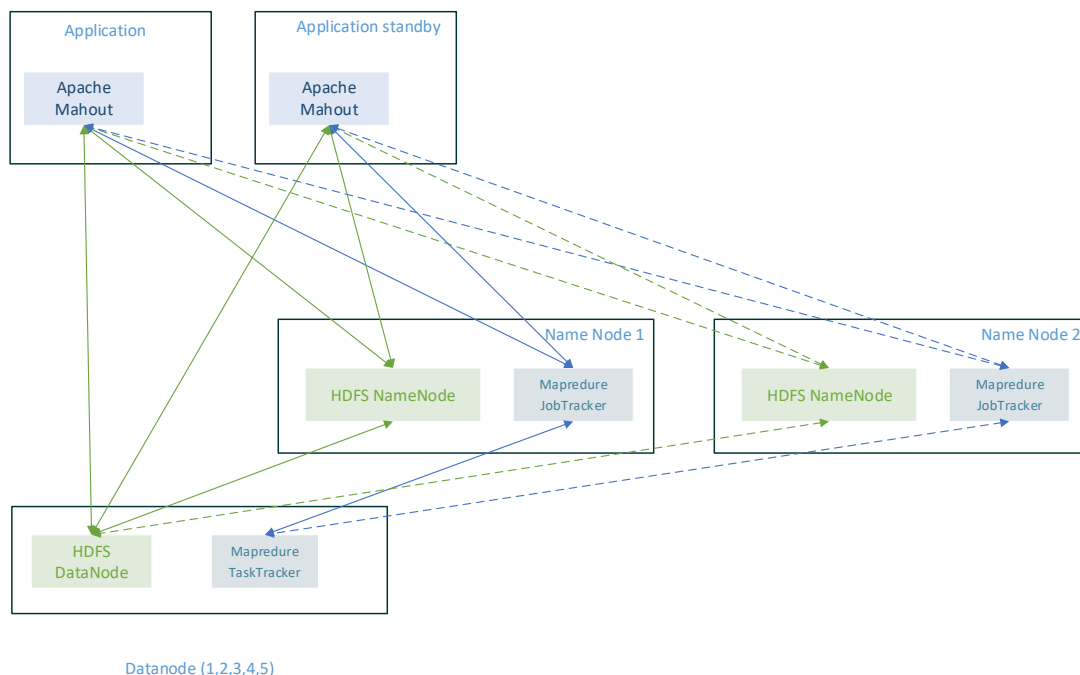


Hai server Application và Application Standby kết nối heartbeat với nhau qua ứng dụng KeepAlive để duy trì VIP cho một server.

Ở điều kiện bình thường server Application sẽ được gán VIP và mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application.

Khi Server Application ngừng hoạt động thì VIP sẽ tự động được gán sang Application Standby, lúc này mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application standby cho tới khi Application server hoạt động ổn định trở lại.

- **Sơ đồ kết nối ứng dụng Apache Mahout**



Ứng dụng Apache Mahout được cài đặt trên 2 máy chủ Application và Application Standby.

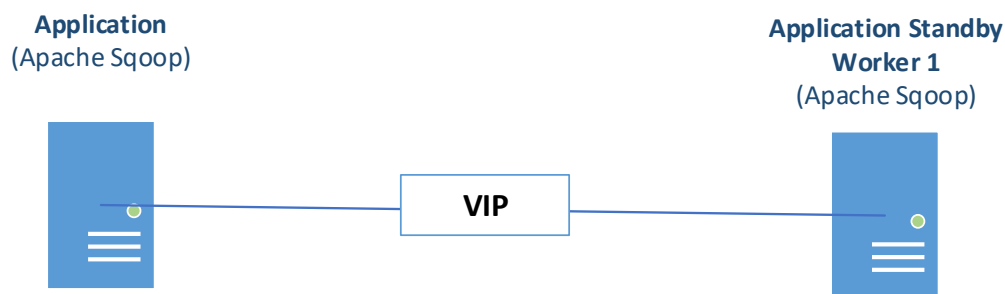
Apache Mahout hoạt động độc lập trên 2 server Application và Application Standby.

Khi một trong hai server gặp sự cố thì server còn lại sẽ được điều hướng request tới thông qua VIP đảm bảo không bị gián đoạn dịch vụ.

Mahout được mở các kết nối tới các ứng dụng bên trong như HDFS, Mapredure để xử lý các tác vụ và dữ liệu trong hệ thống Hadoop, Hbase.

**II.2.5.15. Kết nối Apache Sqoop**

- **Sơ đồ kết nối máy chủ Apache Sqoop**

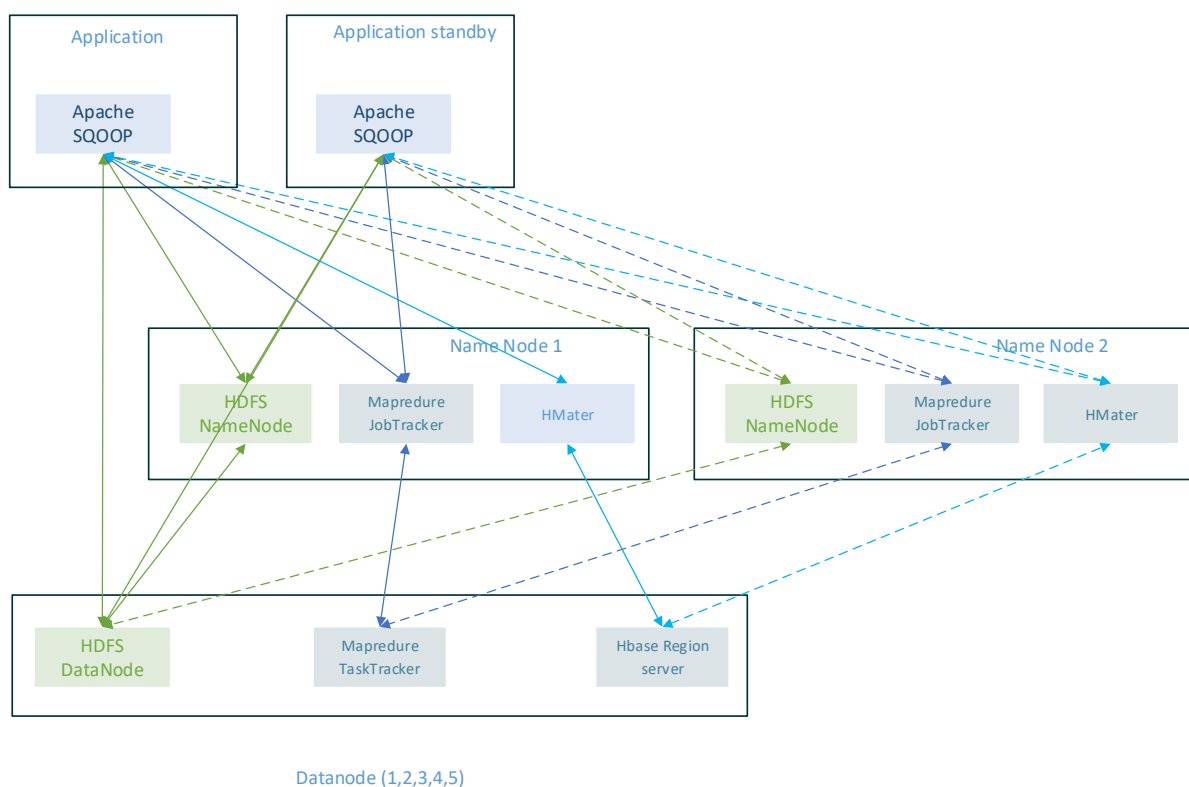


Hai server Application và Application Standby kết nối heartbeat với nhau qua ứng dụng KeepAlive để duy trì VIP cho một server.

Ở điều kiện bình thường server Application sẽ được gán VIP và mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application.

Khi Server Application ngừng hoạt động thì VIP sẽ tự động được gán sang Application Standby, lúc này mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application standby cho tới khi Application server hoạt động ổn định trở lại.

**- Sơ đồ kết nối ứng dụng Apache Sqoop**



Ứng dụng Apache Sqoop được cài đặt trên 2 máy chủ Application và Application Standby.

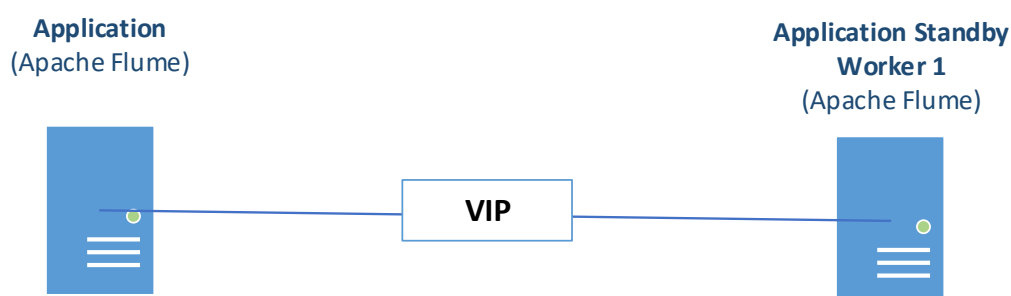
Apache Sqoop hoạt động độc lập trên 2 server Application và Application Standby.

Khi một trong hai server gặp sự cố thì server còn lại sẽ được điều hướng request tới thông qua VIP đảm bảo không bị gián đoạn dịch vụ.

Sqoop được mở các kết nối tới các hệ thống dữ liệu bên ngoài và các ứng dụng bên trong như HDFS, Mapredure, Hbase để phục vụ đưa dữ liệu vào trong hệ thống Hadoop, Hbase.

#### ***II.2.5.16. Kết nối Apache Flume***

##### ***- Sơ đồ kết nối máy chủ Apache Flume***

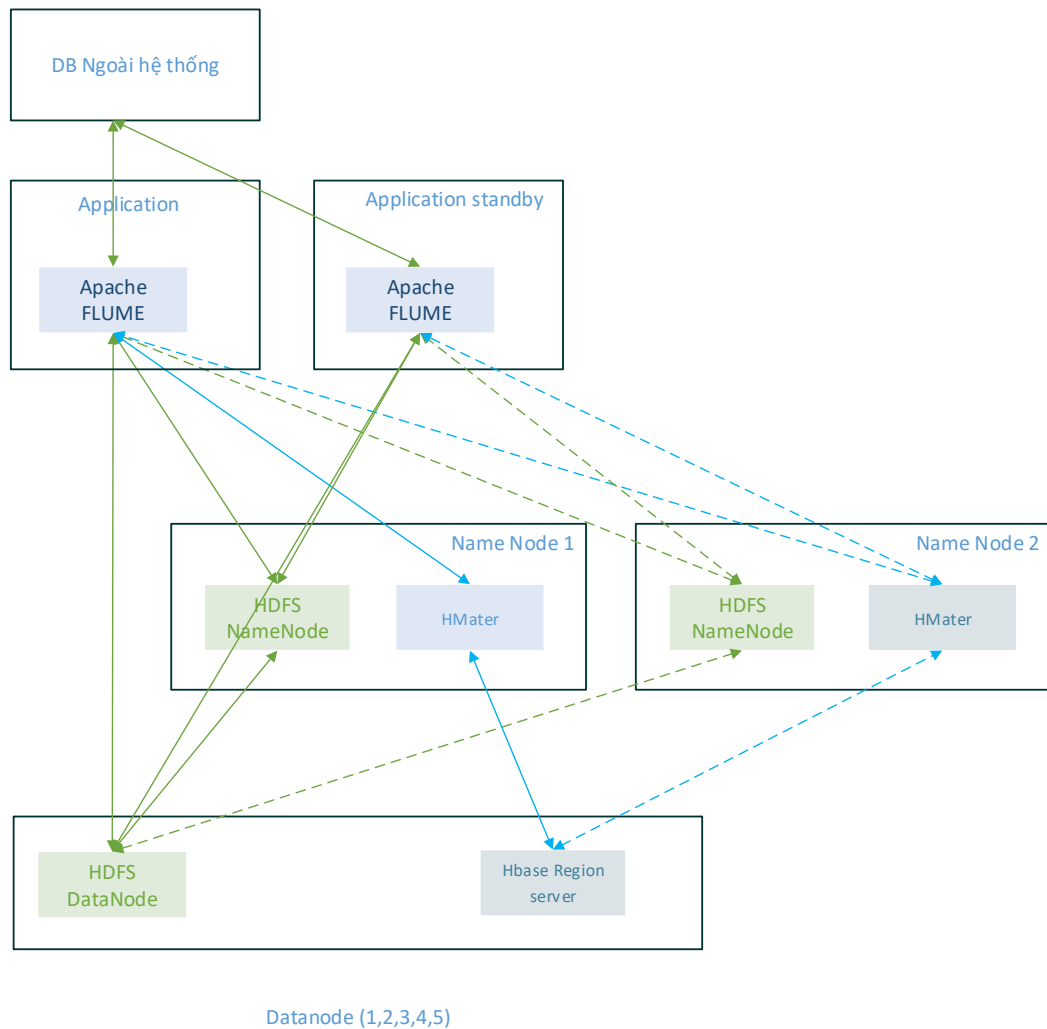


Hai server Application và Application Standby kết nối heartbeat với nhau qua ứng dụng KeepAlive để duy trì VIP cho một server.

Ở điều kiện bình thường server Application sẽ được gán VIP và mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application.

Khi Server Application ngừng hoạt động thì VIP sẽ tự động được gán sang Application Standby, lúc này mọi truy cập của người dùng/ứng dụng khác thông qua VIP sẽ được điều phối sang server Application standby cho tới khi Application server hoạt động ổn định trở lại.

##### ***- Sơ đồ kết nối ứng dụng Apache Flume***



Ứng dụng Apache Flume được cài đặt trên 2 máy chủ Application và Application Standby.

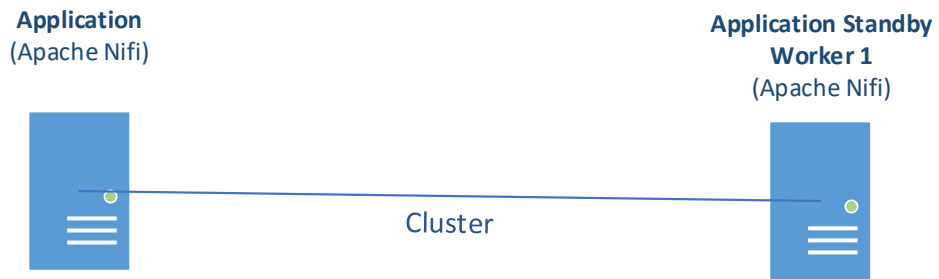
Apache Flume hoạt động độc lập trên 2 server Application và Application Standby.

Khi một trong hai server gặp sự cố thì server còn lại sẽ được điều hướng request tới thông qua VIP đảm bảo không bị gián đoạn dịch vụ.

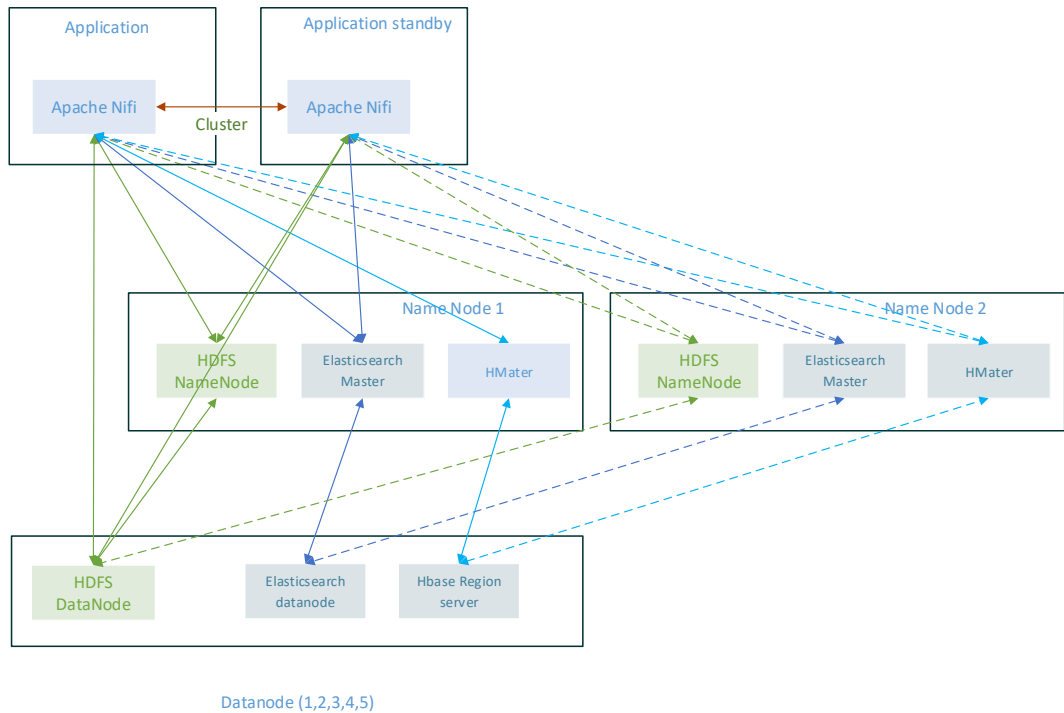
Flume được mở các kết nối tới các hệ thống dữ liệu bên ngoài và các ứng dụng bên trong như HDFS, Hbase để phục vụ đưa dữ liệu vào trong hệ thống Hadoop, Hbase.

#### II.2.5.17. Kết nối Apache Nifi

- Sơ đồ kết nối máy chủ Apache Nifi



- Sơ đồ kết nối ứng dụng Apache Nifi



Apache Nifi hoạt động dưới chế độ Cluster gồm 2 node trên 2 server Application và Application Standby.

Ở điều kiện bình thường ứng dụng trên server Application sẽ đóng vai trò là điều phối coordinator bsg ứng dụng trên server còn lại sẽ đóng vai trò follower.

Khi một node ngừng hoạt động thì ứng dụng ở node còn lại sẽ hoạt động với vai trò điều phối coordinator cho tới khi node được khôi phục.

Apache Nifi kết nối tới các hệ thống dữ liệu bên ngoài và các thành phần bên trong như Apache Kafka Queue, HDFS, Elasticsearch, Hbase để thực hiện đưa dữ liệu từ hệ thống ngoài vào các thành phần bên trong.

Tên ứng dụng	Dịch vụ	Port	Tên máy	Giao thức
	nifi.web.https.port			TCP

<b>Apache Nifi</b>	nifi.remote.input.socket.port			TCP
	nifi.cluster.node.protocol.port			TCP
	nifi.cluster.node.load.balance.port			TCP

## ***II.2.6. Yêu cầu chi tiết về việc đào tạo hướng dẫn sử dụng; triển khai, hỗ trợ, quản trị, vận hành sản phẩm***

### ***II.2.6.1. Yêu cầu chi tiết về việc đào tạo hướng dẫn sử dụng***

#### ***a. Yêu cầu chung về đào tạo***

Việc đào tạo hệ thống phần mềm phải đảm bảo tuân thủ quy trình, nội dung các công việc thực hiện theo đúng quy định của Nhà nước về quản lý đầu tư ứng dụng công nghệ thông tin sử dụng nguồn vốn ngân sách nước, quy trình đào tạo, hướng dẫn sử dụng hệ thống phần mềm đồng thời tuân thủ các bước theo tiêu chuẩn CMMI-LEVEL 3 hoặc tương đương.

#### ***b. Yêu cầu tổ chức lớp đào tạo***

##### ***- Lớp đào tạo hướng dẫn sử dụng cán bộ vận hành***

+ Mục tiêu đào tạo: Giúp học viên nắm được các chức năng sử dụng, khai thác hệ thống một cách hiệu quả.

+ Hình thức đào tạo: Đào tạo trực tiếp tại Trung tâm thông tin và dữ liệu KTTV - Tổng cục Khí tượng thủy văn (Trung tâm thông tin và dữ liệu KTTV bố trí phòng học)

+ Quy mô khóa đào tạo: Gồm 01 lớp.

+ **Học viên: 10 người.**

+ Yêu cầu về giảng viên: 01 giảng viên; 01 trợ giảng.

+ Thời gian đào tạo: **02 ngày/ 01 lớp.**

+ Nội dung đào tạo:

- ✓ Giới thiệu nghiệp vụ chung của hệ thống phần mềm cài đặt.

- ✓ Hướng dẫn sử dụng các chức năng sử dụng của chương trình.
- ✓ Hướng dẫn cách thức khai thác, thống kê số liệu báo cáo đầu ra của hệ thống.
- ✓ Thực hành theo từng vai trò người dùng.
- ✓ Giải đáp các thắc mắc.

**- Lớp đào tạo cán bộ quản trị, cài đặt hệ thống**

+ *Mục tiêu đào tạo:* Giúp học viên nắm được các thao tác cài đặt và kỹ năng quản trị hệ thống, quản trị bảo mật dữ liệu, vận hành để có khả năng hỗ trợ khai thác vận hành hệ thống một cách hiệu quả.

+ *Hình thức đào tạo:* Đào tạo trực tiếp tại Trung tâm thông tin và dữ liệu KTTV - Tổng cục Khí tượng thủy văn (Trung tâm thông tin và dữ liệu KTTV bố trí phòng học)

+ Quy mô khóa đào tạo: Gồm 01 lớp.

+ Học viên: **05 người.**

+ Yêu cầu về giảng viên: **01 giảng viên.**

+ Thời gian đào tạo: 03 ngày/ 01 lớp.

+ Nội dung đào tạo:

- Hướng dẫn cài đặt và cấu hình hệ thống.
- Hướng dẫn quản trị hệ thống.
- Hướng dẫn quản trị người dùng, nhóm người dùng.
- Hướng dẫn theo dõi giám sát hệ thống.
- Thực hành trên hệ thống
- Giải đáp các thắc mắc.

**II.2.6.2. Yêu cầu về triển khai, hỗ trợ, quản trị, vận hành sản phẩm**

Việc triển khai cài đặt hệ thống phần mềm phải đảm bảo tuân thủ quy trình, nội dung các công việc thực hiện theo đúng quy định của Nhà nước về quản lý đầu tư ứng dụng công nghệ thông tin sử dụng nguồn vốn ngân sách nước, quy trình triển khai cài đặt hệ thống phần mềm đồng thời tuân thủ các bước theo tiêu chuẩn CMMI-LEVEL 3 hoặc tương đương.

#### *II.2.6.2.1 Yêu cầu về triển khai, hỗ trợ*

Tuân thủ quy định về công tác triển khai, giám sát công tác triển khai và nghiệm thu dự án đầu tư ứng dụng công nghệ thông tin sử dụng nguồn vốn ngân sách nhà nước bao gồm các hoạt động cơ bản trong công tác triển khai, giám sát công tác triển khai:

- Thực hiện công tác triển khai và lập nhật ký công tác triển khai.
- Thực hiện giám sát công tác triển khai và lập nhật ký giám sát công tác triển khai.

Các công việc cần thực hiện để đảm bảo chất lượng, tiến độ triển khai gồm:

- Lập bảng tiến độ thực hiện chi tiết trước khi triển khai
- Lập hệ thống quản lý chất lượng phù hợp với quy mô triển khai
- Bố trí nhân lực, thiết bị triển khai
- Cài đặt ứng dụng phần mềm và vận hành thử hệ thống hạ tầng kỹ thuật, thiết bị sau cài đặt
- Thực hiện đào tạo hướng dẫn sử dụng, đào tạo cho cán bộ quản trị; hỗ trợ, quản trị, vận hành và các công việc triển khai khác theo đúng hợp đồng và thiết kế chi tiết
- Xử lý, khắc phục các sai sót, lỗi phát sinh trong quá trình triển khai (nếu có)
- Báo cáo chủ đầu tư về tiến độ, chất lượng, khối lượng theo quy định của hợp đồng
- Phối hợp với chủ đầu tư và các đơn vị liên quan trong quá trình giám sát công tác triển khai, kiểm thử hoặc vận hành thử, nghiệm thu, bàn giao

#### *II.2.6.2.2 Yêu cầu duy trì, vận hành sản phẩm*

Tuân thủ quy định về duy trì, vận hành hệ thống thông tin ngành tài nguyên và môi trường đối với dự án đầu tư ứng dụng công nghệ thông tin sử dụng nguồn vốn ngân sách nhà nước.

Thực hiện duy trì, vận hành phần mềm hệ thống tuân thủ theo các bước như sau:

- Kiểm tra, giám sát hệ thống
  - + Kiểm tra các cổng kết nối của phần mềm, dịch vụ.

- + Kiểm tra các service của phần mềm, dịch vụ trên hệ điều hành máy chủ dịch vụ.
- + Kiểm tra các tính năng của phần mềm, dịch vụ.
- + Kiểm tra nhật ký logs hoạt động của phần mềm, dịch vụ.
- Ghi nhận sự cố
  - + Ghi nhận sự cố.
  - + Xác minh sự cố.
  - + Cập nhật danh mục sự cố.
- Phân tích sự cố
  - + Phân loại, đối chiếu danh mục sự cố.
  - + Phân tích các nguyên nhân có thể gây ra sự cố.
  - + Đề xuất giải pháp khắc phục sự cố.
- Khắc phục sự cố
  - + Nghiên cứu giải pháp được đề xuất.
  - + Thực hiện giải pháp khắc phục.
  - + Kiểm tra hệ thống sau khi thực hiện giải pháp khắc phục.
  - + Cập nhật danh mục sự cố.
- Báo cáo, thống kê, nhật ký
  - + Tổng hợp, xây dựng báo cáo trong quá trình duy trì vận hành phần mềm hệ thống.
- Cập nhật
  - + Lập kế hoạch, thông báo cho các bộ phận liên quan.
  - + Thực hiện sao lưu các dữ liệu cần thiết.
  - + Tiến hành cập nhật dịch vụ.
  - + Kiểm tra vận hành sau cập nhật.
- Sao lưu
  - + Lập kế hoạch phương án sao lưu.
  - + Kiểm tra, xác định các nội dung cần sao lưu.
  - + Thực hiện sao lưu.
  - + Kiểm tra tính toàn vẹn, đầy đủ của các bản sao lưu.
- Phục hồi
  - + Lập kế hoạch phương án phục hồi khôi phục dữ liệu.
  - + Kiểm tra hệ thống.

- + Thực hiện phục hồi.
- + Kiểm tra hoạt động của dịch vụ sau khi thực hiện phục hồi.
- Quản lý thông tin, cấu hình
  - + Lập kế hoạch thực hiện.
  - + Thực hiện cấu hình, thay đổi, cập nhật dữ liệu.
  - + Kiểm tra hoạt động phần mềm sau cấu hình, thay đổi.

## **II.2.7. Kế hoạch triển khai**

### **II.2.7.1. Dự kiến các mốc thời gian, tiến độ thực hiện**

Căn cứ vào kế hoạch triển khai chung của dự án “Xây dựng, hoàn thiện hệ thống thông tin, cơ sở dữ liệu tài nguyên và môi trường” trong đó kế hoạch thực hiện như sau:

#### **- Năm 2023-2024 (quý III/2023 – quý II/2024):**

- + Xây dựng Xây dựng HTTT đất đai quốc gia (MPLIS); hạ tầng số, kết nối, an toàn thông tin cho CSDL đất đai quốc gia tập trung, thống nhất.

- + Xây dựng các nền tảng, hạ tầng quản trị, thu thập dữ liệu tài nguyên và môi trường.

#### **- Năm 2024:**

- + Tiếp tục: Xây dựng các nền tảng, hạ tầng quản trị, thu thập dữ liệu tài nguyên và môi trường.

- + Xây dựng, hoàn thiện hành lang pháp lý phục vụ thu nhận, tạo lập, quản lý, kết nối, chia sẻ toàn diện nguồn tài nguyên số về tài nguyên và môi trường: Ưu tiên tập trung xây dựng Mô hình CSDL và quy định kỹ thuật về nội dung và cấu trúc dữ liệu của các từng lĩnh vực.

- + Xây dựng, hoàn thiện các CSDL, HTTT quốc gia, quan trọng ngành TN&MT đến năm 2025.

- + Hoàn thiện nền tảng tích hợp, chia sẻ, kết nối dữ liệu tài nguyên và môi trường.

- + Đào tạo và truyền thông (ưu tiên thực hiện các hạng mục đào tạo về giải pháp và công nghệ).

#### **- Năm 2025: Thực hiện các hạng mục còn lại, tổng kết dự án.**

Hạng mục “**Xây dựng, triển khai mô hình cơ sở dữ liệu ứng dụng công nghệ Big Data cho Trung tâm thông tin và dữ liệu tại Tổng cục KTTV**” nằm trong

nhóm hạng mục “Xây dựng các nền tảng, hạ tầng quản trị, thu thập dữ liệu Tài nguyên và Môi trường” do đó sẽ tuân thủ theo kế hoạch dự kiến chung và chi tiết dự kiến các mốc thời gian và tiến độ như sau:

- Công tác lập thiết kế chi tiết và dự toán đầu tư: Quý III/2023
- Tổ chức lựa chọn nhà thầu: Quý IV/2023
- Triển khai thực hiện hợp đồng: Quý IV/2023 đến quý I/2024
- Hoàn thành nghiệm thu và bàn giao sản phẩm: Quý I/2024

Bảng tiến độ tổ chức thực hiện cụ thể như sau:

Công việc	Thời gian thực hiện dự kiến (ngày)						
	1	2	3	4	5	6	7
Lập thiết kế chi tiết	30						
Thẩm định và phê duyệt TKTC		15					
Tổ chức lựa chọn nhà thầu			45				
Phê duyệt kết quả lựa chọn nhà thầu				10			
Triển khai thực hiện hợp đồng					90		
Hoàn thành nghiệm thu						10	
Đưa hệ thống vào vận hành chính thức						2	
Quyết toán hạng mục							20
<b>Tổng thời gian thực hiện</b>	<b>220 ngày</b>						