ISBN: 978-981-18-7950-0
2023 the 13th International Workshop on Computer Science and Engineering (WCSE 2023)
doi: 10.18178/wcse.2023.06.009

Overall Architecture Design for the Construction of Surveying and Mapping Geographic Big Data Center

Daibing Lan, Yinggang Guo+ and Jinlu Gao

Northwest Institute of Nuclear Technology, Xi'an 710024, China

Abstract. In response to the problems in storage, processing, sharing, integration, fusion application, standardized management, and other aspects of surveying and mapping geographic information, the overall architecture of the surveying and mapping geographic information big data center has been designed. The "four layers and two systems" include infrastructure layer, data resource layer, platform support layer, application service layer, standard specification system, and protective operation system, providing reference for the construction of provincial and municipal surveying and mapping geographic information big data centers.

Keywords: surveying and mapping geographic information, big data center, overall architecture

1. Introduction

A geographic information data center is an organization that centrally stores, manages, and distributes geographic information data. It is an important infrastructure for achieving national spatial planning and resource and environmental management. The construction and management of geographic information data centers need to comprehensively consider the collection, storage, processing, utilization, and sharing requirements of geographic information data, involving multiple links such as data collection, data fusion, data quality control, data standardization, data sharing, and data security protection. Surveying and mapping geographic information have the characteristics of wide coverage and high data accuracy [1]. With the rapid development of artificial intelligence and big data, the content and form of geographic information products have developed into three-dimensional, dynamic, ubiquitous, and intelligent. The types and quantities of surveying and mapping geographic information data represented by large-scale terrain and geomorphic data, high-resolution image data, oblique photography, and 3D models are gradually increasing [2-4]. The problems of insufficient storage space, insufficient data processing capabilities, low resource sharing level, low data integration level, weak fusion application level, lack of standardized management, and dispersed and independent data are becoming increasingly prominent, seriously restricting the application effect of surveying and mapping geographic information data [5, 6]. Therefore, taking "data aggregation, deep integration, strong sharing, promoting intelligence, building platforms, and wide application" as the overall goal, this article and utilizes new technologies such as big data [7, 8], cloud computing, and spatial geographic information to design the overall architecture of the surveying and mapping geographic information big data center, call it the ' the "four layers and two systems". In order to provide reference for the construction of provincial and municipal surveying and mapping geographic information big data centers.

2. Overall Architecture Design

The overall architecture of the surveying and mapping geographic information big data center system has been designed as "four layers and two systems". The "four layers" refer to the infrastructure layer, data resource layer, platform support layer, and application service layer, while the "two systems" refer to the standard specification system and protective operation system. The overall architecture is composed of five layers: basic support, internetworking, resource elements, service supply, and service generation. Various service capabilities are generated according to the goal of digitalization, networking, service-oriented, and intelligence. The overall system architecture design is shown in Fig. 1.

+ Corresponding author.
E-mail address: fariel_gyg@163.com.

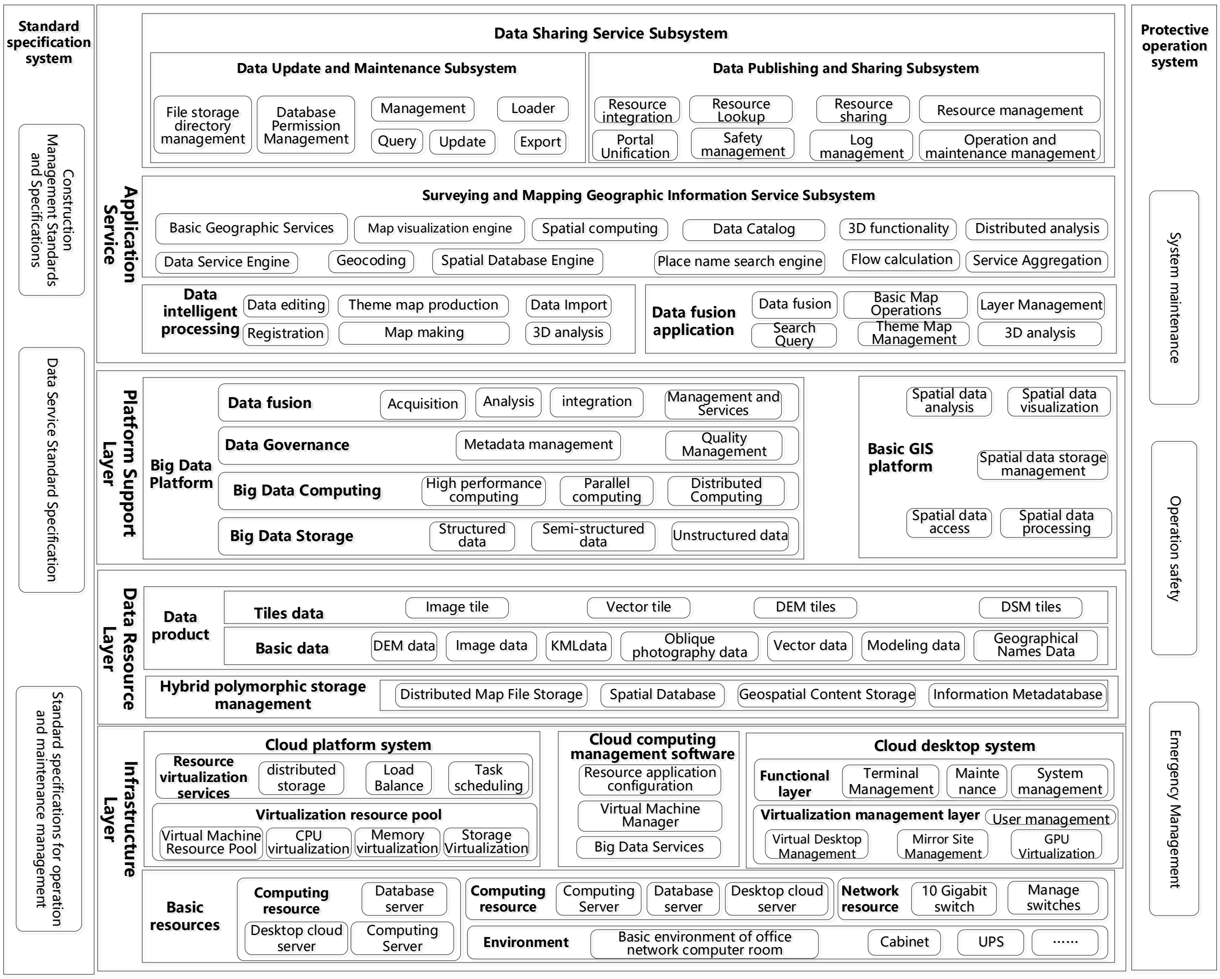

Fig. 1: Overall architecture of surveying and mapping geographic information big data center.

2.1. Infrastructure Layer

The infrastructure layer refers to the underlying resource support of the surveying and mapping geographic information big data center. It can provide software and hardware, operating systems, virtualization environments, databases, and theoretical regulations for the deployment and operation of the system, and provide a supporting environment for data resource exchange, big data analysis, and geographic information platform. It is necessary to plan and construct information infrastructure resources such as computing, storage, network, desktop cloud, and computer rooms, and use advanced technologies such as virtualization and desktop cloud to build a cloud platform system for surveying and mapping geographic information data processing and a cloud desktop system. This will achieve cloud management, on-demand allocation, dynamic scheduling, and orderly expansion of various resources. It will quickly respond to changes in business needs, and provide a necessary and solid foundation for the application of upper level surveying and mapping geographic information business.

2.2. Data Resource Layer

The data resource layer refers to the data resource support of the surveying and mapping geographic information big data center, including basic spatial data, thematic data, map cache data, raw data, and basic map services. All kinds of multi-source heterogeneous basic surveying and mapping geographic information data (including DEM data, image data, KML data, tilt photography data, vector data, modelling data, place name data, etc.), product data produced by data processing system (including image tile, vector tile, DEM tile, DSM tile and other tile data) are stored in distributed map files, spatial database, geospatial content storage Information element database and other data storage methods realize the mixed polymorphic storage management of data products. The data resource layer will realize the full process management of all kinds of surveying and mapping geographic information data from production, acquisition, processing, storage, transmission and utilization, build a unified data resource pool, and provide data resource support for upper business applications.

2.3. Platform Support Layer

The platform support layer refers to the basic platform environment for the business application construction of the surveying and mapping geographic information big data center, which is composed of development components, basic base map services, thematic base map services, spatial analysis services, service resource management, monitoring statistics, GIS cluster management, big data platform and other modules. It mainly provides service support for the system application platform and related business systems, including big data platforms and basic GIS platforms.

The big data platform supports the hybrid storage framework of structured data, semi-structured data and unstructured data, as well as the hybrid computing capabilities of high-performance computing, parallel computing and distributed computing, providing basic support for upper distributed applications. Its core functions include metadata management, data quality management and other data governance function modules, data collection, data integration, data asset management and service, data analysis and other data fusion function modules.

The basic GIS platform software is the core platform for the construction of surveying and mapping geographic information big data centers, providing service support for various spatial geographic data applications in the big data centers. It fully supports cloud native architecture and has the characteristics of microservices, containerization, and automated layout technology. It relies on high-performance cross platform technology to support the construction and efficient operation of the entire platform.

2.4. Application Service Layer

The application service layer refers to various surveying and mapping geographic processing business systems customized and developed by the surveying and mapping geographic information big data center, mainly including four subsystems: data processing, data resource sharing services, data intelligent processing, and data fusion applications.

The data processing subsystem should mainly rely on high-performance geographic information technology for data storage, processing, analysis and calculation. To meet the requirements of processing surveying and mapping geographic information, publishing data services, sharing and exchanging data resources, and managing cloud GIS microservice environments, it is necessary to construct surveying and mapping geographic information data processing software, application server software, portal server software, and management server software. The basic GIS platform is the core platform of the surveying and mapping geographic information big data center, providing service support for various spatial geographic information data applications in the big data center. The selection of GIS platform directly affects the security, reliability, development, scalability, and response performance of the upper application system.

The resource sharing service system realizes the whole process of data engineering services from data collection, cleaning to data governance, and finally completes data storage and warehousing. It is necessary to modify and process the surveying and mapping geographic information data in accordance with data standard specifications, data processing specifications, data model specifications, etc., in order to achieve the application requirements. The data is cleaned and quality checked in the resource sharing service system, and stored in storage. On this basis, resource sharing services build unified data asset management, providing metadata management, data label management, data lifecycle management, and data catalog cataloging. After completing the data asset organization work, it is necessary to collect data elements based on application requirements and construct appropriate data models to achieve effective correlation of multiple types of data. The data will be provided in two ways: distribution and application integration.

Modify and process data according to data standard specifications, data processing specifications, data model specifications, etc. to achieve application requirements. The data is cleaned and quality checked in the resource sharing service system, and stored in storage. On this basis, resource sharing services build unified data asset management, providing metadata management, data label management, data lifecycle management, and data catalog cataloging. After completing the data asset organization work, it is necessary to collect data elements based on application requirements and construct appropriate data models to achieve effective correlation of multiple types of data. The data will be provided in two ways: distribution and application integration.

The information intelligent processing subsystem is based on components such as metadata collection and data collection, which summarize various existing surveying and mapping geographic information data into a database. Then, through data conversion, quality inspection, and other processes, it achieves standardized processing of multi-source surveying and mapping geographic information data, laying the foundation for further development and utilization.

Due to the dispersion of existing surveying and mapping geographic data resources, the lack of effective correlation in data resource construction and management applications, and significant differences in data structures, a mechanism for multi-source data fusion processing urgently needs to be formed. The fusion application subsystem should adopt fusion processing methods such as format conversion, content extraction, scale matching, and region stitching to meet the needs of different application scenarios. Integrate data from different sources according to unified standards, certain fusion rules, and models to provide a dataset for the fusion of spatiotemporal, scale, location, logical association, and other aspects of spatial information data. Through processing toolsets such as format conversion, projection conversion, cleaning synthesis, classification extraction, as well as tools such as grids, vectors, 3D scenes, DEM, tile generation, and correlation, provide service support for data fusion analysis, Ultimately, achieving effective utilization of data and improving its utilization value.

2.5. Standard Specification System

Standard specifications are an important guarantee for the orderly and sustainable operation of big data centers, as well as an important basis for achieving data integration, database establishment, data application, and rational utilization of resources. Geographic information big data center should carry out the construction of a data standard system, including basic standard specifications, data standard specifications, service standards, and management standards. By referring to the national geographic spatial information framework, e-government related standards, the design guidelines for the national geographic information public service platform, and other relevant standards such as basic geographic spatial information and geographic framework data, a standardized system for the geographic information big data center system need to be formulated, and the construction and management standards, data service standards, and operation and maintenance management standards for the surveying and mapping geographic big data center also need to be formulated. The construction of a standardized system can help standardize the collection, processing, management, publication, and application of geospatial information data.

2.6. Protective Operation System

On the one hand, with the increasing amount of data in surveying and mapping geographic big data center, the complexity of data application models, and the diversity of application types, network security will face more severe risks. On the other hand, the large-scale application of new technologies such as cloud computing, big data, the Internet of Things, and artificial intelligence faces many new security risks, which also pose new challenges to network security protection [9, 10]. According to the overall requirements of surveying and mapping geographic management for data security, closely focusing on the practical needs and development needs of data management security, in response to a series of security risks and threats faced by future business development, technological development, and capability development, security protection should be carried out from aspects such as security physical environment protection, security communication network protection, security zone boundary protection, and security computing environment protection. Specifically, protective measures such as centralized encrypted storage, firewalls, antivirus, confidentiality systems, network behaviour auditing, and intrusion detection are adopted to achieve security and compliance purposes.

3. Conclusions

This article focuses on the problems in the application of surveying and mapping geographic information data, and designs a "four layer and two system" overall architecture of surveying and mapping geographic information big data center. The construction of the big data center can promote the improvement of the digitalization, networking, service-oriented, and intelligent level of surveying and mapping geographic data management, and improve the service and comprehensive application efficiency of surveying and mapping geographic information data. The next step will be to design various functional modules and business processes within the overall architecture system, and research will be conducted on key technologies such as virtualization parallel computing, massive data fast computing and storage access, multi-source heterogeneous business data conversion, and heterogeneous hybrid cloud computing.

4. References

[1] Fritsch D. Three-Dimensional Geographic Information Systems - Status and Prospects[C]// *International Society for Photogrammetry and Remote Sensing Congress*. 2009.

[2] Xu H, Badawi R, Fan X, et al. Research for 3D visualization of Digital City based on SketchUp and ArcGIS[J]. *Proceedings of SPIE - The International Society for Optical Engineering*, 2009, 28(3): 52-54.

[3] Fonseca F, Davis C, Camara G. Bridging Ontologies and Conceptual Schemas in Geographic Information Integration[J]. Geoinformatica, 2003, 7(4): 355-378.

[4] Esch T, Asamer H, Soukup T, et al. Monitoring Urbanization with Big Data from Space - The Urban Thematic Exploitation Platform[C]// *Conference on Big Data from Space - BiDS'17*. 2017.

[5] Ndebele M, Mazhindu A N. Design and Implementation of an Android GIS Data Collection Application for Students with Non-Programming Experience Using the MIT App Inventor: A Case Study on a Tick Data Collection APP[J]. *Intelligent Information Management*, 2023, 15(1):13.

[6] Xu S, Liu X. Design and Study on Management Tools of Land Data Center for Integration of Urban and Rural Areas[J]. *Journal of Geographic Information System*, 2009, 1(1):42-47.

[7] Zhu X. Feasibility of the Provincial Surveying and Mapping Geographic Information Big Data Center Construction in Zhejiang Province[J]. *Geospatial Information*, 2015, 13(3):77-78.

[8] Qiao C. Big Data and Its Influence on Geospatial Industry[J]. *Bulletin of Surveying and Mapping*, 2013(1):107-109.

[9] Huang Y. In the internet of everything, security is the Foundation- Challenges and Countermeasures for National Information Security in the Era of Big Data[J]. *China New Telecommunications*, 2020, 22(21): 151-152.

[10] Li H. National Information Security Issues in the Era of Big Data[J]. *Journal of Sichuan Police College*, 2016, 28(4): 87-91.