

Design and Implementing of Meteorological Data Management System

Jia Xiong¹ Zhang Weimin¹ Yu Yi^{1,2} Qin Yaolei¹

1. Computer Science, University of National Defense Technology, Changsha, China

2. NO.94865 Troops of Chinese People's Liberation Army, Hangzhou, China

jiaxiong@nudt.edu.cn wmzhang104@163.com raul09@163.com qinyin111@163.com

Abstract—Based on characteristic of meteorological data, meteorological data management system(MDMS) is designed and implemented. It is divided to client-side and server-side parts. The server-side which is responsible for meteorological data archived retrieved uses two-tier design, including logical server and physical server. The client-side provides based on portal framework Gridsphere provide web interface for users and researchers to use this system in grid environment. This system can manage heterogeneous meteorological data and improve the efficiency of meteorological data archived and retrieved. The system has been successfully applied to Chinese Ensemble Prediction Grid Project (CEPGP),and CEPGP works efficiently to demonstrate the feasibility of the system.

Keywords—Meteorological data, MDMS, Gridsphere portal framework

I. INTRODUCTION

Meteorology is an information-intensive industry, and meteorological data includes ground, aerial, agriculture, marine, satellite, radar and other weather observations (probe) test records, and the various meteorological data sets, a variety of statistics and numerical analysis of climate data, weather service products made by processing these records. For researchers, the real-time data and historical data of meteorological data is the basis for meteorological research; for meteorological users, the various types of meteorological products are weather service products users obtain. Meteorological data show the following characteristics [1]:

A. Massive Information

Meteorological data includes daily observations, and the different height of the field data. As time goes, historical data will be explosive growth; meanwhile, continuously introduce new forecasting model, different research methods, and data types increasing result in the amount of data dramatically.

B. Wide Varieties

Meteorological data includes not only all weather observations (probe) test data, but also all kinds of weather statistics and numerical analysis of data, so data types are very large, such as the introduction 4DVar analysis will introduce new data types. Data types with the development of weather forecasting technology continue to grow.

According to the characteristics of meteorological data, Meteorological data archiving and rapid retrieval is necessary. As the types of meteorological data will be growing, and consequently the meteorological data

management system based on the object-oriented metadata database is implemented. At the same time, for high-performance distributed environment, we implement Meteorological data management portal system based on portal system called Gridsphere. The system can provide the portal interface that manages heterogeneous meteorological data for the users in different regions. It is integrated into the application of ensemble prediction grid portal system, and effectively improves the application of meteorological data.

II. BACKGROUND

A. Gridsphere[2,3] Portal Framework Based On Portlet

Portlet is a Java-based web component, managed by the portlet container.

Portlet is responsible for handling user requests, and dynamically generated content to the user. Portlet can be used as pluggable user interface components. Portlet specification is JSR 168 that the communication standard between portlet and portlet container, and a portlet comply with JSR 168[4] specification can be easily deployed to a different portal framework. Gridsphere, a part of the project GridLab, is a kind of portal framework based on portlet that supports users add custom functions by plug-in components, integrating portlet container, providing portlets environment, portlet life cycle management, and a number of common core portlets, Such as localization portlet, portal layout portlet, user management portlet, login portlet, and other group management portlet. Therefore, Gridsphere is highly scalable, and easily integrates with existing portal application sub-functions, thereby speeding up the development process of the new portal.

B. Ensemble Prediction Application Grid Portal System

Ensemble prediction application grid portal system is grid portal management system of Chinese Ensemble Prediction Grid Project (CEPGP), providing the Web interface that enables users to easily use CEPGP services, effectively shielding heterogeneity between the grid middleware and resource nodes, Fig.1 architecture of ensemble prediction application grid portal system allowing the user through the traditional web browser to access and use the heterogeneous resources. It provides security services, data management services, operations management services, resource management services, executive management services and cooperative development services. Ensemble prediction application grid portal system architecture is shown in Fig. 1.

III. METEOROLOGICAL DATA MANAGEMENT SYSTEM

Meteorological data management service is based on the underlying Meteorological Data Management System (MDMS) to provide services to users. MDMS based on Client/Server architecture to achieve, is divided into client and server. Client-side is the meteorological data management service portlet in portal system, providing web interface for user to archive and retrieve meteorological data, visualizing the operating results. Server-side is MDMS, dividing into logical and physical servers to archive and retrieve meteorological data, and based on metadata to locate, organize and manage meteorological data.

A. MDMS Server-side Architecture

MDMS Server-side is responsible for meteorological data archive storage and retrieves detection, whose architecture is shown in Fig.2.

Server-side deployed grid server includes logical server and physical server.

Logical server interact to logical metadata database to confirm the semantic information of meteorological data, such as confirming the data fields data or forecast products. Physical server interact to physical metadata database to confirm the physical information of meteorological data, such as the grid node and storage address meteorological data stored. Fig.2 also shows the process of retrieving meteorological data. Concrete steps described below:

1) User sends the retrieval request to CEPGP server through client-side. CEPGP server forwards the request to MDMS logical server.

2) Logical server forwards the request to metadata database. Metadata database returns data reference list to logical server by querying logical metadata. Logical sends data reference list to physical server.

3) Physical server sends the data reference list to metadata database. Metadata database returns the grid node and storage address the meteorological data stored by querying physical metadata.

4) According to the storage address, physical server retrieves meteorological data queried and returns results to the user.

Server-side uses hierarchical design approach. Given a huge amount of meteorological data, and strong scalability, introduce object-oriented design of the metadata. As metadata has object polymorphism, Extending meteorological data and storage devices is very easy. So system scalability and efficiency of the archived/retrieved are increased.

B. Metadata[5] and Data Organization

The metadata database and meteorological data database are deployed on MDMS server-side and each grid node. Metadata describes the meteorological data. MDMS provides metadata capture and broadcast publish, and grid server node is responsible for triggering each node to update the metadata data regularly. Each node will broadcast the update result to the grid management system to ensure that the meteorological data of each node can be updated in real time. It should be noted that logical metadata is only displayed on

grid server node, and the physical metadata is deployed in each node. Logical meta-data describes the types of meteorological data, such as field data, observation data and provides the logical description of meteorological data. Physical metadata describes the physical storage location of meteorological data, such as the nodes, files, and offsets meteorological data archived.

Meteorological data database for storing the archived meteorological data are deployed on each grid node. Mainly three ways to organize data on meteorological data archived:

1) NetCDF (Network Common Data Format)

NetCDF is used to store conventional observations, which is very flexible to transfer mass array-oriented data and is the format of file of many data acquisition software to widely apply to terrestrial, marine and atmospheric science.

2) GRIB (GRId In Binary)

GRIB data format is used to store the numerical prediction products, which is able to represent multi-dimensional data and has a modular structure. It supports multiple compression formats, especially spectral data and image data. It uses the IEEE standard floating-point representation.

3) Directory/file Format

Directory/files format can also store forecast products, usually filed based on the time of meteorological data.

IV. METEOROLOGICAL DATA PORTAL MANAGEMENT SYSTEM

Meteorological Data Portal Management System is the client-side of the whole management system, which is the realized portlet based on Gridsphere portal framework. It provides web interface for on-demand data service, such as downloading meteorological data, rapidly retrieving and archiving meteorological data by remote batch processing, so it shields effectively off the difference between the web system and the underlying meteorological data.

A. MDMS Application Service Package

Because the user request received by the logical server of MDMS is a sequence of different instructions which are packaged as a script to describe every field data or observation data stored in the meteorological database. The syntax format of script is shown in Table I.

TABLE I. THE GRAMMAR FORMAT FOR USER REQUEST

Syntax Format	Meaning
verb,	verb: retrieve compute list read write
keyword1	keyword: a known MDMS variable, e.g. date and time
=value1,	keyword, field keyword, observation
keyword2	keyword.
=value2,	value: value assigned to the
... = ...,	keyword, e.g. Analysis or temperature
keywordN	
=valueN	

So, we use portlet service technology to package the application deployed on logical server into uniform portlet services, realize the separation of MDMS and business logic called by the upper, facilitate client-side to call the portlets, and reduce the system's coupling. Users only need to customize different verbs and keywords through the browser

to create user instructions to call MDMS logical server application. MDMS Logical server archived /retrieval application packaging process is as follows:

- According to the characteristics of the user instructions received by the logical server of MDMS, such as verb, keyword, define the related portletservice interfaces.
- Deploy portlet services by define the profile of the portletservices.
- By Gridsphere unified factory class, use factory pattern to create the application instance of the service, to realize the remote calling and managing the packaging services.

B. Meteorological Data Management Service

Meteorological Data Management Service, a part of Meteorological Data Portal Management System, provides meteorological data archived service, retrieved service and visualization service.

1) Meteorological Data Archival/Retrieval Service

As meteorological data is organized in three ways, meteorological data management service using the factory pattern to provide and implement a uniform access interface that improves the data format of scalability and access transparency. First, client-side calls the portlet service packaging the logical service application to send user request command to logical server. Then logical server forwards the request to logical metadata database. Finally, according to the result logical metadata database returns, local server calls different data archival/retrieval method. Fig.3 shows the sequence diagram of retrieving meteorological data.

2) Meteorological Data Visualization Service

This service takes advantage of Metview application to visualize the retrieve results of meteorological data, then sends visualization products to user, such as prediction image with different time and different domain.

C. User Requests Scheduling

MDMS implements the scheduling of user requests working with the role of control functions provided by Gridsphere. Gridsphere can configure different levels of user rights. The user requests has higher rights are first executed. When the requests are send by the same level users, MDMS provides three scheduling options. First, process user request according to the order of user request coming. This

scheduling option will make more urgent the user request delay, but we can enhance the user level to improve. Second, sort user requests according to the recommend urgent token carried. This program applies to the situation that most requests are not urgent. Third, implement user requests pre-processing function to predict user request coming in a period of time. Then sort the requests by the finished time. In such circumstances that user requests are rare, this method will be inefficient, but we can shorten the time period to resolve this problem.

V. METEOROLOGICAL DATA PORTAL MANAGEMENT SYSTEM WEB INTERFACE

Meteorological data portal management system has been applied to CEPGP. Its operation interface is shown in Fig. 4.

VI. CONCLUSION

Meteorological data management system (MDMS) is designed and implemented, which is based on Client/Server architecture. The server-side is divided to logical server and physical parts, through metadata to locate, organize, archive, retrieve and manage meteorological data. The client-side is based on Gridsphere portal framework using JSR 168 portlet technology, which is able to adapt to the grid environment. The function of this system is provided as web services, so the users and researchers in different places can share the meteorological data. The system has been a part of CEPGP. CEPGP works efficiently to demonstrate the feasibility of the system.

REFERENCES

- [1] Taylor, J.,;Buizza, R. Using Weather Ensemble Prediction in Energy Demand Forecasting[J]. International Journal forecast,2003,19:57-70
- [2] Gridsphere <http://www.gridsphere.org/>
- [3] Jason Novotny;Michael Russell;Oliver Wehrens.GridSphere:A Portal Framework for Building Collaborations[EB/OL].<http://www.gridsphere.org/documents>,2004.
- [4] JSR168 <http://www.jcp.org/en/jsr/detail?id=168>
- [5] Lorcan Dempsey;Rachel Heery. Specification for resource description methods Part 1: A review of metadata: a survey of current resource description formats.

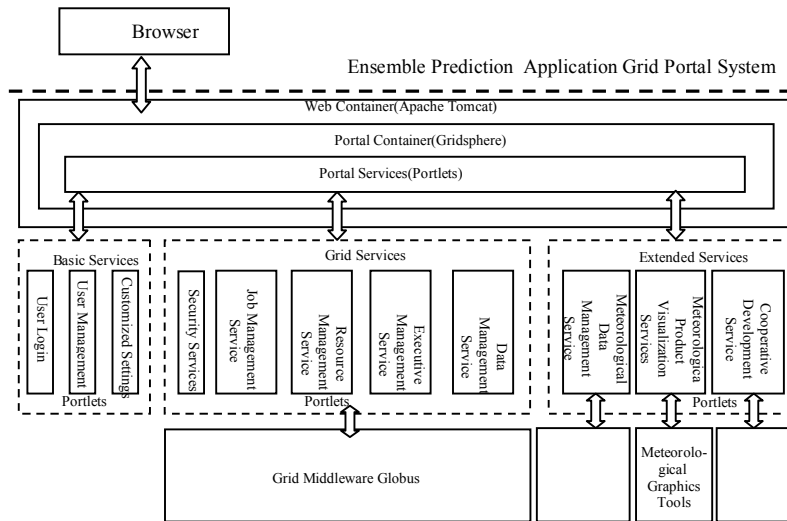


Figure 1. Ensemble prediction application grid portal system architecture

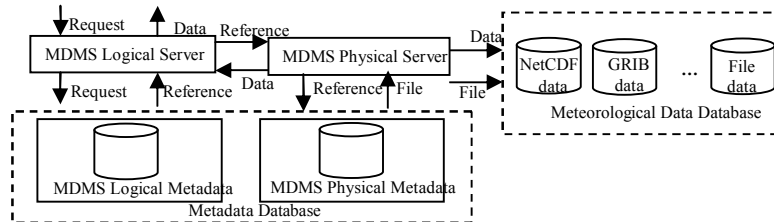


Figure 2. Architecture of MDMS Server-side

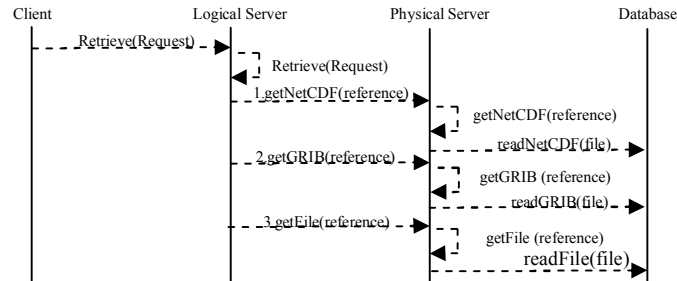


Figure 3. Sequence Diagram of Retrieving Meteorological Data



Figure 4. Meteorological Data portal Management System Web Interface