

# 国家级气象资料存储检索系统的建立\*

沈文海 赵芳 高华云 高峰 马强

(中国气象局气象信息中心,北京 100081)

## 摘要

对国家级气象资料存储检索系统的需求进行了分析,进而对系统进行了功能设计,并在技术调研和分析的基础上确定了功能实现的技术路线;在数据管理方面采用数据库和文件系统相结合的方式;在存储管理方面采用分级存储管理模式,并解决了数据库的分级存储管理技术;在检索应用方面提供了程序调用式检索方式,并通过采用中间件技术实现了基于B/S结构的程序调用式检索;通过选用HPSS技术解决了由异种机型构成的高性能计算机群存储池问题;通过分别建立实时数据库和综合数据库解决了气象资料的规范化存储和实时业务使用需求之间的矛盾;在应用软件开发方面严格按照有关规范进行。

关键词:国家级 气象资料 存储检索系统 系统需求 功能设计

## 引言

近20年来,气象事业迅猛发展,新的观测手段不断出现,气象研究领域不断拓宽,预报准确率不断提高,伴随着对气象资料的服务要求也在大幅提升;与此相对应的气象资料服务手段和能力却相对较弱,国家一级的气象资料存储检索系统至今尚未建立,这在相当程度上制约了气象业务、科研的发展。

在此背景下,依托于“短期气候预测业务系统工程建设项目(一期)”,中国气象局气象信息中心自2002年起,承担国家级气象资料存储检索系统(下简称NMARS)的项目建设任务,经过调研、详细设计等各项工作,目前该项目已进入具体实施阶段。

NMARS是一个规模巨大且十分复杂的系统,在国内尚无可供借鉴的现成的成功案例,国外的类似案例也不多,这给系统的建设增加了难度;因此NMARS的设计工作——特别是正确的建设思路,是极其重要的。

## 1 需求概述和建设目标

国家一级的气象业务、科研工作对气象资料的服务需求主要包括“资料使用”和“资料存储”两个方面。

### 1.1 使用需求

包括对资料的种类、时空范围、检索时效和检索方法等方面的要求;不同种类的用户

\* 本文由“短期气候预测业务系统工程建设项目”资助。

2004-01-15收到,2004-05-16收到修改稿。

在需求方面亦有所不同;归纳起来,可将用户分为“实时业务”、“准实时业务”以及“科研、服务”等三类<sup>[1]</sup>。

实时业务用户因工作内容所决定,其使用资料的种类和时空范围相对固定,但检索时效要求高,检索方法要求规范严格。由于目前各实时业务系统的自动化程度均较高,系统运行平台相对固定,因此其检索方式也要求规范化、程序化和自动化,以保证整个业务系统运行的客观、准确和实时。

准实时业务用户与实时业务用户相仿,使用资料的种类和时空范围亦相对固定,检索时效亦有一定要求(但较实时业务用户略低),运行平台相对较多,在检索方法方面的要求与实时业务用户相比更为广泛。

科研、服务用户由于其各自工作的侧重点不同,对资料的种类和时空范围要求也各不相同;此外,由于其工作的性质所决定,对检索的时效要求不高(表1)。

表1 气象业务、科研对资料的需求

	资料种类、范围	检索方式	检索时效
实时业务	相对固定	程序化、自动化	秒级
准实时业务	相对固定	程序化、自动化、自由检索	数秒级
科研、服务	不固定	自由检索	分钟级

## 1.2 存储需求

包括观测资料存储、产品(含数值模式产品)存储、高性能计算机群用户空间需求以及系统安全备份等四个方面。

气象观测资料种类繁多,除传统沿用的地面、高空观测资料外,由于观测手段的不断增加,气象卫星、天气雷达、飞机观测、海洋船舶以及大气成分等资料均陆续纳入到气象资料的使用范畴。这些新增资料种类多,增量大,在使用方式上彼此也有较大差别。这一切要求 NMARS 具有良好的存储手段来保存所有气象资料(包括按标准规范要求对资料进行归档)。

随着气象服务领域的不断拓展,气象服务产品不断增加,此外各种天气、气候以及台风、城市污染等数值模式每日也产生出众多的为各种预报业务提供参考依据的模式产品,这些产品均需要长期保存。

从调研的结果看,需要保存的气象资料的总量及增量见表2<sup>[1]</sup>,这些资料全部需要存储在 NMARS 系统中,并可随时检索使用。

表2 需要保存的气象资料的总量及增量

	历史存量	年增量	备注
常规观测资料	282 GB	37.2 GB	不含未信息化的资料
雷达资料		6.7 TB	均为经压缩后的数据,其中基数据:5.6 TB,产品:1.1 TB
卫星产品资料		2.1 TB	2.1 GB/d
数值预报模式产品资料	7.1 TB	8.9 TB	含国内外数值预报模式的分析和预报产品
特种观测资料	40.2 GB	126.1 GB	不包括未来的沙尘暴观测资料
气象服务产品		2 GB	
总计	7.4 TB	17.8 TB	不包含气候环境资料和需数字化的纸质资料部分

作为北京高性能计算机应用中心的依托单位,中国气象局气象信息中心拥有由各种品牌大型计算机组成的高性能计算机群,这些计算机除为众多气象数值模式提供稳定可靠的运行平台外,还向社会提供高性能计算机资源。长期以来,存储资源的短缺一直制约着高性能计算机群效率的充分发挥。因此高性能计算机群存储资源的彻底解决也是 NMARS 建设中的一项重要内容<sup>[2]</sup>。

一个能长期稳定运行的存储系统,其系统以及数据的安全是不可缺少的必要条件。除在设备配置方面须充分考虑外,系统及数据的备份和恢复手段也是 NMARS 必须具备的<sup>[1]</sup>。

### 1.3 建设目标

在未来可预见的时间范围内,基本满足国家级气象业务、科研工作对气象资料的使用需求和存储需求,并为省级部门内部用户提供一定程度的气象资料服务。

为实现上述建设目标,NMARS 必须具备下列功能。

① 资料收集、处理功能:能够对通过通信系统、Internet 等各种途径获取的气象资料进行汇集和处理,以生成格式规范、质量可靠的数据库资料。

② 资料存储管理功能:能够针对各类资料自身的特点以及用户对资料的使用需求和惯例,对各类资料进行有针对性的,以满足用户对资料的使用需求,同时遵从规范、标准、高效等原则为主旨的有效管理。

③ 资料检索功能:提供满足用户对各类资料的基本检索要求,同时遵从规范、标准、高效原则的检索方式。系统必须满足用户对所有存储资料的在线检索需求。

④ 监控功能:能对全系统进行全面、规范和实时的状态监视和记录,并具备一定的故障自动恢复功能。

⑤ 用户管理功能:能对各级别用户进行较严格、合理的身份和行为管理,以保证系统安全和数据安全。

⑥ 备份和分级存储功能:具备对整个系统及数据的备份/恢复功能(包括:全备份、增量备份和差分备份)。具备对非常用资料实施分级存储策略,以提高一级存储设备使用效率的功能。

## 2 国外气象行业存储系统建设应用情况概述

由于历史原因,国外一些发达国家的气象部门在处理气象资料存储管理及应用方面的问题时各自采取了适合本部门具体情况策略,归纳起来有以下特点:

### (1) 系统构架:分级海量存储和管理

① 硬件平台:无论是美国还是欧洲中心气象部门的存储检索系统,均采用以磁盘和自动磁带库构成的分级海量存储硬件平台,对资料管理使用多级管理存储方式,即在线磁盘阵列、近线自动带库、离线磁带。从 NCEP 和 NCDC 的新一代系统架构看,海量存储系统的硬件平台将朝着大容量磁盘(超过 TB)和大容量带库(超过 PB)方向发展,并通过光纤通道交换机构成海量存储局域网(SAN)。

② 软件平台:对数据的组织采用结构化和非结构化两种格式,并结合使用 HSM 技术和商用数据库技术进行管理。

由于国外气象存储检索系统建设较早,以及各自的需求不同,因此这些系统大都采用 HSM 软件技术基于文件(即非结构化数据格式)进行管理。随着商用数据库技术的日益成熟,目前也开始结合商用数据库技术进行管理。如 NCDC 已经使用商用数据库对常用的、面向主题的小量资料进行管理,并采用 SYBASE 作为进行观测资料分析的数据管理平台,同时正尝试使用 ORACLE 作为使用频度高的常用历史资料管理工具,特别是作为 WWW 资料服务的数据库管理系统。加拿大的 NAS 存储系统使用的是 ORACLE 数据库<sup>[3]</sup>。

### (2) 因特网等先进技术在资料应用手段中的应用

美国气象部门对外服务广泛使用 Internet 技术。NCDC 基本用户界面为浏览器,不论是内部的业务操作、数据处理还是外部用户检索。目前,NCDC 有超过 30% 的服务通过 Internet 网络完成。到 2005 年,这一比例将达到 70% 以上<sup>[4]</sup>。

## 3 系统的功能设计

根据需求调研结果,NMARS 应具备资料收集处理、资料存储管理、资料检索、系统监控、用户管理和备份/分级存储管理等多项功能<sup>[5]</sup>;根据多年来在实际工作中获取的经验,并认真参照国外部分发达国家气象行业中资料存储的有关经验,对上述功能做如下设计。

### 3.1 资料的收集与处理

(1) 资料的收集功能:通过国际/国内通信系统的实时资料收集、通过 Internet 的准实时资料收集、通过骨干网的产品收集、通过介质转储的非实时资料收集、通过交换或汇交获得的科学试验和考察资料、通过购买、交换获得的国内外各种数据集和有关地球环境资料等,以及非信息化资料的数字化处理。

(2) 数据处理功能:包括资料的预处理、数据格式标准化处理、资料的质量控制及资料的统计加工等部分。

### 3.2 资料存储管理

资料存储管理功能包括:数据存储管理策略的制订、数据管理方法的选择、数据保存时间(在线、近线存储设备上)的确定、资料的追加、更新和清除以及资料归档等方面。

#### (1) 数据存储管理策略

综合考虑用户的需求和系统设备配置的性价比,NMARS 系统决定采用多级存储方式进行存储管理:采用大容量磁盘阵列的在线存储,用来保存实时性强、使用频率高的常用观测资料和产品以及系统管理信息、元数据等;采用自动磁带库系统作为近线存储,用来保存在线存储设备上超过保存期限或使用频率相对较低的资料;采用脱机磁带等离线存储形式,用于数据的长期存档等。

#### (2) 数据的管理方式

根据资料种类的不同和数据应用方式的区别,NMARS 系统采用数据库管理和文件系统管理相结合的存储管理方式,并将二者有机地结合起来,实现对系统整体数据的集中统一管理。

数据库管理方式:对日常天气预报和气候诊断与预测等业务所需的实时性强、要求响应速度快、查询方式复杂而数据量相对小的实时气象资料采用结构化数据库管理,以有利

于实时业务对数据的应用需求。

文件管理方式:对那些数据量大、使用频度低、查询方式相对简单的各种产品数据、科学实验数据、中间结果数据等采用文件方式管理,而其相关的索引信息和元数据则纳入到数据库管理中,以便于数据的查询检索和元数据的统一管理。

### (3) 数据保存时间

NMARS 需要合理地规定各类数据在不同存储设备中的保存期限,以满足用户对检索资料的时效和时间序列的要求。

### (4) 资料的追加、更新与清除

实时资料的更新由运行的程序自动完成,其数据处理程序始终随着资料格式和内容的变化而更新。

对于非实时资料的追加,有自动和人工两种方式。对由实时转入非实时的数据采用自动方式,对那些分散的、通过加工收集的零散数据则采用人工投入命令方式追加入库。

为了保证有足够的磁盘空间以存储最新的气象数据,系统必须制定定时清除库中超出保存周期的数据的策略,清除策略依据不同类型数据保存周期的不同而有所区别。除了程序定时自动清除外,系统还提供交互式人工指定清除功能。

### (5) 资料的归档

具有永久保留和使用价值的资料必须按照一定的数据格式和规范进行归档,归档资料的存储介质一般是磁带或光盘。资料归档时系统自动形成归档目录,资料标签和档案标签,以元数据的形式存放在磁盘中。

存储系统的归档按照资料类别分别进行。归档保存的时间按照中国气象局颁布的《气象记录档案管理规定》执行。资料归档的目的是为了再次使用,因此在设计资料归档方案时,必须考虑归档资料的应用问题,以保证在需要时能够进行存档资料的复制、检索与应用。

## 3.3 数据检索功能

数据的检索对用户来说是透明的,即无论用户索要的资料存放在何处、何种设备上,用户都能通过同一界面获得数据。

NMARS 除提供有效的数据管理和基本的公用数据检索接口外,还应尽可能为用户提供更好的数据服务,包括特定用户的数据需求和格式转换、较丰富的数据显示与表达方式等。

NMARS 提供两种用户检索方式:程序调用式检索和基于 Web 方式的人机交互式检索。程序调用方式主要满足各业务系统从数据库中提取资料的过程自动化,该方式有基于“客户端/服务器结构”(C/S 结构)和基于“客户端/应用服务器/数据库服务器”(即 C/W/S 结构)两种体系结构的检索应用。

基于 Web 方式的人机交互式检索提供部门内外用户通过浏览器方式对数据库进行的交互式联机检索或数据下载,为保证数据的安全性和规范管理,基于浏览器检索的 Web 服务建立了较完善的用户管理系统,实现用户网上登记、身份确认、在线资料下载成本支付等功能。

## 3.4 监控功能

数据库监控是指对数据库整体运行状态进行实时监控,并具备对一些异常情况进行

处理的能力。数据库监控包括系统监控和应用监控。

#### (1) 系统监控

设备运行状况监控:NMARS 的设备包括高性能服务器、存储设备、存域网设备、存域网管理服务器、以及有关的网络及接口设备等;对这些设备的监控的内容包括服务器 CPU 的使用率、运行进程的数量、存储设备的空间使用率、网络端口的流量、设备本身运行的状况等。

系统监控信息的显示:NMARS 有专门的系统监控平台,在监控平台上有比较完善的监控信息显示功能,及报警功能。系统监控信息平台除主要提供给系统管理员使用,还可提供给授权的特殊用户使用。因此该平台具有多用户共享功能。

#### (2) 应用监控

数据处理流程监视:数据流监视的对象主要是数据收集、加工处理、分发及检索应用等环节,系统在各个重要运行环节设置观察点,并能随时获取该点的运行状态参数,当某一个环节出现异常时,能够及时发出警告信息通知系统管理人员。

数据存储状态监控:对存储检索系统中资料的使用状况(包括系统中数据迁移的情况、数据存储类型及总量、数据应用情况等)信息进行采集,并进而对这些信息进行相应的统计分析,根据分析结果对相应的策略(如数据迁移策略)进行必要的调整。

对这些状态进行有效监控将有助于从宏观上对系统进行掌握和控制,从而保证系统运行的有效性。

用户检索服务监视:包括对检索系统软件运行状态的监视;对用户的身份及行为(包括访问时间、访问次数、访问数据内容及数据量等)进行记录和监视。

### 3.5 用户管理

#### (1) 系统用户的分类与分级

用户按类别分为“内部”和“外部”两类。内部用户是指中国气象局系统的各业务和科研单位,包括国家级和省级业务科研用户。外部用户是指内部用户之外的非商业性社会用户,外部用户的分类依据“气象资料共享实施细则”,分为四级:一般用户、会员用户、重要用户和特殊用户。

#### (2) 用户的权限

用户的权限包括使用资料的权限和检索资料的优先级权限两个方面。对所有内部用户和外部用户而言,对系统存储的资料都只有获取权限,而无修改资料和添加资料的权限。对内部用户而言,所有 NMARS 存储的公开资料都可以检索获得。如果各级用户需要得到非公开资料,则必须履行一定的行政审批手续。

根据“气象资料共享实施细则”的规定,不同级别的外部用户分别享有在网上获得不同资料的权利和义务。具体见国家气象中心颁发的“气象资料共享实施细则(试行)”。关于不同级别的外部用户可以获得哪些不同的资料,应参考“气象资料分类分级标准”。

### 3.6 数据备份与分级存储

#### (1) 备份和恢复

系统备份和恢复:包括对系统软件及数据库管理软件的备份。系统软件的备份定期进行。除了定期备份外,当系统有重大变更时也应及时进行备份。所备份的系统软件必

须具有易恢复性。

数据的备份和恢复:数据备份的机制包括数据源备份、日志备份、库备份、文件备份等。数据恢复包括数据库恢复、文件恢复、二级存储至一级存储的恢复、三级存储至二级存储的恢复等。

由于 NMARS 系统的复杂性,不同的数据库或数据集的备份需要采用不同的方式才能满足不同业务的需求。

#### (2) 数据的分级存储管理

数据分级存储管理是指数据的主体根据分级存储管理策略,分别在价格及使用性能较高的存储设备和价格及使用效能相对较低的存储设备上存储。其目的是提高价格及性能较高的存储设备的使用效率。数据迁移依照一定的策略进行。相应的策略一般有:使用频度低限策略、存储空间上限策略等。

数据库资料的分级存储:数据库资料的分级存储主要指对所管理的满足迁移策略的非结构化数据(即文件库)资料的向二级存储设备上迁移。

由于 NMARS 要求对所有资料满足在线检索,因此结构化数据必须存储在数据库库体中,随着日积月累,库体将逐渐膨胀,故当库体接近存储设备容量极限时,对数据库本身也要进行分级存储管理。

高性能存储系统的分级存储:NMARS 将建立面向中国气象局气象信息中心所有主要计算机设备(包括用户)的公用的高性能存储系统,该系统为业界所有主流品牌计算机提供公用存储资源及相应的服务。该系统将采用分级式存储管理方式,为用户提供形式上无限的存储空间。同时该系统由大容量磁盘阵列、自动磁带库、脱机磁带以及高性能存储系统软件(含分级存储管理功能)构成。

综上所述,NMARS 系统在功能设计方面主要包括以下内容:通过按照规范要求自行开发的资料预处理软件实现资料的收集和预处理。通过应用商用数据库实现对以数据为检索单位的资料(以观测资料为主)的数据管理;通过文件系统实现对以记录为检索单位的资料(以数值模式产品、雷达资料、卫星遥感资料及服务产品等非结构化数据资料为主)的数据管理,并由商用数据库对其元数据——进而对其检索路径——进行统一管理。向用户提供的人机交互式检索采用基于浏览器的技术方案,程序调用式检索分别提供采用基于 C/S 结构和基于 B/S 结构的两种技术方案。通过在业界主流平台上按照规范要求自行开发的监视系统应用软件实现对 NMARS 全系统的全程无空隙监视及预警。通过对资料及用户级别的划分和使用策略的制订,以及通过商用数据库、操作系统和应用软件对策略的实现,来达到对用户身份及行为的管理。通过备份/恢复功能实现 NMARS 的系统和数据安全;通过分级存储管理实现用户空间的大幅扩展和存储资源的合理有效利用。

## 4 系统的总体结构

系统在逻辑上由三个数据库组成,分别是实时数据库、综合数据库和对外共享数据库。其中,实时数据库存储满足天气、气候基本业务系统运行所需的各种资料,内容、保存时段及数据量估算见表 3。

表 3 实时数据库功能

资料种类	保存时间	数据量(GB)
常规观测资料(地面、高空)	2个月	60.0
农气资料	1年	6.0
全国辐射资料	2个月	6.0
全球海洋观测资料(船舶、浮标)	2个月	6.0
数值模式产品资料	分析资料:2个月 预报资料:7天	510.0
天气雷达资料	7天	285.0
气象卫星产品资料	2个月	21.0
服务产品子库	1年	2.6
气象灾害资料(台风要素资料、沙尘暴要素资料等)	1年	
其他资料	2个月	50.0
总计		946.6

综合数据库存储 NMARS 所保存的所有资料,内部由 7 个子库构成,分别管理以下资料<sup>[5]</sup>:

子库 1(管理信息子库):元数据、管理信息、监控系统信息;

子库 2(常规观测资料子库):地面资料、高空资料、辐射资料、农气资料、海洋资料;

子库 3(数值模式产品子库):国内外数值预报模式产品及常用的长时间序列分析资料产品(如:欧洲气象中心再分析资料、美国 NCEP/NCAR 再分析资料)等;

子库 4:雷达资料、卫星产品资料、服务产品、气象灾害资料、图形/图像资料;

子库 5(数据集资料子库):历史气候代用资料、科学考察及试验资料、大气成分资料;

子库 6(实时数据库要素资料延时子库);

子库 7(报文资料延时子库)。

对外共享数据库由 3 个子库组成,分别为:

元数据子库:数据基本信息、数据参照信息和数据完整性信息;

管理信息子库:数据处理信息、用户信息、数据存储信息等;

资料子库:常规气象观测资料、农业气象资料、专题气象资料、大气环境资料、气象科学试验资料、数值预报模式分析产品资料、气象卫星资料、省级地面一般站资料等。

实时、综合、对外共享数据库各自所具有的功能见表 4。

表 4 实时、综合、对外共享数据库的功能

序号	功能名称	实时数据库	综合数据库	对外共享库
1	实时资料收集、处理	✓		
2	非实时资料收集、处理		✓	
3	实时资料分发	✓		
4	程序调用式资料检索	✓	✓	
5	人机交互式资料检索	✓	✓	✓
6	结构化数据管理	✓	✓	✓
7	非结构化数据管理	✓	✓	包含数值预报产品
8	资料定时清除	✓		
9	系统、数据备份/恢复	✓	✓	✓
10	数据迁移/回迁		✓	
11	数据归档		✓	
12	系统监视	✓	✓	✓
13	用户管理	✓	✓	✓



## 5 几个关键技术策略

### (1) 实时业务对系统检索时效的要求

NMARS 作为国家级的气象资料存储系统,其存储资料的范围包括所能收集到的所有气象资料,并同时提供所有这些资料的在线检索,这要求该系统具有强大的存储能力和容量。另一方面,NMARS 同时向国家级的气象业务、科研提供资料检索服务,故该系统同时应当具备良好的、可满足业务科研工作需求的检索性能。由于资料的增量巨大,要确保所有资料的在线检索,对资料的存储结构、格式和策略都需要加以一定的限制,而这些限制将在相当程度上影响这些资料检索效率的提高。

为避免两者之间的矛盾,NMARS 采取资料冗余的策略,即:单独建立一个资料内容、时间范围都相对有限的“实时数据库”,以专门服务于实时业务对资料的检索需求;该库针对实时业务对资料的检索特点和时效要求,对所管理的资料内部结构进行专门设计,以满足实时业务需求。此外,建立旨在存储所有资料的“综合数据库”,以满足资料大容量、长时间序列的存储需求<sup>[2,5]</sup>。

### (2) 系统安全对结构设计的影响

对外共享服务是 NMARS 的另一项重要功能,社会用户通过 Internet 对本系统进行访问,并获取相应的资料;由于社会用户成份复杂,且无法控制,网络安全成为实现该功能的最大障碍;为避免因网络安全造成整个系统的损失,除采取目前通常采用的网络安全策略外,本系统将对外共享功能从系统主体部分剥离开来,单独建立“对外共享数据库”,并将其移至内部防火墙以外,以避免其网络安全故障波及整个系统。

对于内部而言,同样存在网络安全问题;系统在结构设计方面,将应用服务器做为物理上单独存在的服务器分离出来,以增加系统的网络安全度<sup>[5]</sup>。

### (3) 库内数据管理方式规范化要求及在结构设计上的策略

NMARS 所管理的资料种类繁多,既有以常规地面、高空观测资料为代表的结构化数据,也有以数值预报产品、雷达资料等为代表的非结构化数据,有的资料检索频繁,使用方式复杂,也有的资料检索频度相对较低,使用方式相对简单;在库中如何有效地管理这些资料,使之既达到标准规范的存储管理要求,又尽可能地适合这些资料的存储和使用特点,是本系统要解决的另一个问题。系统在综合数据库的设计方面,采用子库形式,将不同种类、不同使用方式的资料归类管理,以期达到上述目的。

## 6 结 语

NMARS 系统作为全新的、十分庞大而复杂的系统,其整个设计过程本身就是一项艰巨的工作,其中凝结着众多设计人员高超的智慧和辛勤的汗水。限于篇幅,本文不可能对系统设计的各个方面逐一做详细介绍(特别是作为 NMARS 灵魂的应用软件的设计开发,以及 NMARS 所采用的关键技术等)。目前系统已进入全面实施阶段,各项工作正在逐步展开。可以预期,该系统将作为一个标志性建设,载入气象事业现代化的记录中,并

在未来的岁月中为我国气象事业做出其应有的贡献。

### 参考文献

- 1 沈文海,熊安元,高华云,等. 中国气象局国家级存储检索系统需求分析报告. 国家气象中心,2002.
- 2 沈文海,高华云,赵芳,等. 国家级气象资料存储检索系统系统设计书. 国家气象中心,2003.
- 3 荣维枝,马强,杨青,等. 中国气象局国家级存储检索系统技术调研报告. 国家气象中心,2002.
- 4 李集明,王国复. 国外大气科学领域数据中心发展概况. 国家气象中心,2001.
- 5 沈文海,赵芳,高峰,等. 国家级气象资料存储检索系统功能规格设计书. 国家气象中心,2003.

## THE CONSTRUCTION OF NATIONAL METEOROLOGICAL ARCHIVAL AND RETRIEVAL SYSTEM

Shen Wenhai Zhao Fang Gao Huayun Gao Feng Ma Qiang  
(*Meteorological Information Center, CMA, Beijing 100081*)

### Abstract

The system requirements of NMARS (National Meteorological Archival and Retrieval System) and the function designs for the system are suggested. Based on the technique investigation and analysis, the technique schemes for the implementations of those functions are also given: the RDBMS integrated with the external file system is used for data management; the HSM (Hierarchical Storage Management) architecture is adopted for storage management, including the storage management of RDBMS; the program retrieval function is provided for data retrieval. In terms of the middleware technology, the realization of the function is based on the B/S (Browser/Server) architecture; the HPSS (High Performance Storage System) solution is applied in the storage pool issue of the HPC (High Performance Computer) cluster system composed of several HPCs with different architecture; the RDB (Real-time Database) system and IDB (Integrated Database) system are designed to solve the contradictory between the standardization storage of the meteorological data and the performance requirements of the real-time operational system; the corresponding specifications are strictly followed in software development.

**Key words:** NMARS System requirements Function designs Technique schemes