

李湘,王甫棣,姜立鹏,等. WIS 的实现技术研究及应用[J]. 气象,2011,37(10):1301-1308.

WIS 的实现技术研究及应用^{*}

李 湘 王甫棣 姜立鹏 祝 婷

国家气象信息中心,北京 100081

提 要: WIS 是世界气象组织(WMO)正在实施和建立的新一代信息系统,是未来支撑 WMO 所有计划以及其他国际组织有关计划的数据传输和服务平台。WIS 由国家中心、数据收集或产品中心和全球信息系统中心(GISC)等三级功能中心组成,其中,GISC 是 WIS 的核心功能中心,承担 WMO 及有关国际组织各项计划的实时数据和产品的收集分发,提供全部 WIS 数据的访问接入。中国气象局作为世界气象组织中有重要影响力的会员,建成北京 GISC,成为 WIS 中支撑责任区内及区域间数据交换的核心功能中心之一是中国气象局的既定目标。本文介绍了 WIS 的功能结构,分析了元数据同步、数据发现、数据同步等 WIS/GISC 关键功能的实现技术及其在北京 GISC 系统设计和开发中的应用。

关键词: WMO 信息系统(WIS),全球信息系统中心(GISC),元数据同步,数据发现,数据同步

Research and Application of the Implementation Technologies for WMO Information System

LI Xiang WANG Fudi JIANG Lipeng ZHU Ting

National Meteorological Information Centre, Beijing 100081

Abstract: WIS (WMO Information System) is an integrated information system and developed for meeting the requirements of all WMO Programmes, affiliated international organization and programmes with respect to information exchange, management and access. Its main functional components are National Centres, Data Collection or Product Centres, and Global Information System Centres (GISC). As an important influential member of WMO, CMA's goal is to serve as one of the GISCs which are core communication centers in WIS and responsible for collecting and distributing the information in their areas of responsibility, and exchanging the information meant for routine global dissemination with other GISCs. The paper gives an introduction to the functional architecture of WIS and the key technologies for implementing the services of metadata synchronization, data discovery and data synchronization, also presents their applications in the design and development of GISC Beijing.

Key words: WMO Information System (WIS), Global Information System Center (GISC), metadata synchronization, data discovery, data synchronization

引 言

世界气象组织(WMO)通过世界天气监视网计划(WWW)、世界气候计划(WCP)、大气研究和环

境计划(AREP)等 10 大科技计划和相关项目开展工作,其所有计划都包含收集数据、制作产品、提供数据分发和共享服务等内容。如果各项计划独立开发信息系统,则可能导致系统间缺乏兼容和效率,限制 WMO 各项计划间的信息交换,阻碍信息为更多

* 公益性行业(气象)科研专项(GYHY200906042)资助

2011 年 2 月 23 日收稿; 2011 年 7 月 26 日收修定稿

第一作者:李湘,主要从事气象通信业务管理和软件开发。Email: lixiang@cma.gov.cn

的用户所使用和共享,也会产生信息系统重复建设和高成本等弊端。为了解决这些问题,WMO 基本系统委员会(CBS)提出了 WMO 信息系统(WIS)的概念,目标是建立一个综合、通用的全球信息系统,为 WMO 和其他国际组织的有关计划的信息交换和共享提供支撑,为各国水文气象部门以外的政府部门及其他用户提供服务^[1]。

2007 年第十五届世界气象大会批准了 WIS 计划,标志着 WIS 进入实施阶段,它将依托目前支撑 WWW 计划数据传输和服务的全球通信系统(GTS)进行实施和过渡^[2]。GTS 是 WMO 现有信息系统中最为成功的部分,但是它基于点对点通信线路和专用传输规范进行数据交换,在支持 WMO 其他计划的数据传输以及为更多用户提供数据服务方面存在着局限性。根据 WIS 实施计划,GTS 将通过技术改进和能力升级发展为 WIS 的核心通信网络,现有 GTS 中心也需要通过能力建设和评估认证,升级为相应的 WIS 功能中心。

全球信息中心(GISC)是 WIS 的核心功能中心,承担全球交换资料的收集和分发,提供对 WIS 全部数据的发现和访问服务。中国气象局(CMA)目前是 GTS 主干通信网的亚洲区域通信枢纽(RTH),建成北京 GISC 是 CMA 的既定目标,也是巩固和提升 CMA 在 WMO 通信网络及信息系统中的地位和影响力的重要举措。同时,通过承担 GISC 职责,还将提升中国气象局对全球实时数据和产品的发现和获取能力,拓展各类国外资料的收集途径,从而为我国气象预报业务和科研提供更为丰富、更高时效的数据支撑和服务^[3-5]。

为此,中国气象局设立了科研和业务专项,以建立北京全球信息中心为目标,开展了 WIS 实现技术研究及 GISC 基础功能软件开发。目前已初步建立了北京 GISC 系统,并通过 WMO 评估。本文在介绍 WMO 信息系统功能结构的基础上,分析了元数据同步、数据发现、数据同步等 GISC 关键功能的实现技术,介绍了相关技术在北京 GISC 系统设计和开发中的应用。

1 WIS 的结构和功能

1.1 WIS 概念

WIS 是 WMO 新一代的数据交换和共享系统,

是综合、通用的信息传输和服务平台(图 1)。就功能而言,它是基于元数据组织和管理数据、具备数据收集和分发功能、提供数据发现和访问服务的综合服务系统。就架构而言,它是遵循统一的元数据标准和数据交换规范的通用信息系统,支持 WMO 各项计划的数据传输和共享,支持世界气候研究计划(WCRP)、全球气象观测系统(GCOS)、全球海洋观测系统(GOOS)等 WMO 与其他国际组织联合发起的研究计划和项目的数据传输和共享,并且可以为 WMO 各会员水文气象部门以外的政府部门和科研机构等直接用户提供服务。

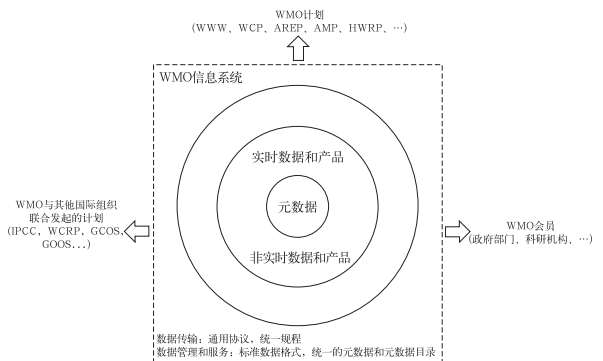


图 1 WIS 的概念

Fig. 1 WIS concept

WIS 的目标是基于统一的数据交换规则,实现实时数据高时效收集和分发,基于标准的元数据规范,整合全球天气、气候、水文及地球环境等领域的大量数据资源,提供便捷的数据共享服务。与 GTS 等 WMO 现有信息系统相比,WIS 的进步和突破主要体现在数据传输能力和数据服务能力两个方面。

在传输能力方面,首先 WIS 突破了 WMO 现有信息系统的通信网络结构,通信连接从点对点的专线连接发展为多点对多点的网络系统和点对多点的广播系统相结合,同时,引入 Internet 作为专用通信系统和网络的补充;其次,依托网状通信连接和广播系统,简化数据汇集流程,改进数据交换机制,提高数据分发能力;另外,依托 Internet 广泛的覆盖性,有效拓展数据传输范围,从 WMO 会员通信中心之间的数据传输和交换,延伸为面向更多组织和机构、甚至直接用户的数据传输服务。

在数据服务能力方面,WIS 基于元数据整合了 WMO 会员及相关国际组织的分布数据资源,不仅能够提供实时数据的传输和交换,还能提供实时和非实时数据的发现和访问服务。因此,与以实时数

据交换为主要功能的 WMO 现有信息系统相比, WIS 是兼具数据传输和共享服务功能的信息系统, 服务数据从实时数据延伸到非实时数据, 服务用户从 WMO 会员的水文气象部门拓展到水文气象部门以外的政府部门、科研机构及相关国际组织等。

1.2 WIS 的结构和布局

WIS 由 WMO 会员国家和地区的众多功能中心和连接各功能中心的通信网络组成(如图 2), 是层次化、分布式的全球信息系统。按照承担职责由少到多、服务区域由小到大, WIS 功能中心依次分为国家中心(NC)、数据收集或产品中心(DCPC)和全球信息系统中心(GISC)等三级^[1]。其中, NC 是

收集国内或地区内的观测数据和产品、向所在区域 GISC/DCPC 汇交数据和和相关元数据、以及面向其国内用户提供数据服务的国家级中心, 是组成 WIS 的基本功能单元; DCPC 是承担某项计划的产品制作、加工、处理的产品中心, 或是承担区域资料收集、交换和存档服务的数据收集中心, 负责所承担计划或所在区域的观测数据和产品收集、向所在区域 GISC 汇交区域内提供全球交换的数据和相关元数据、以及面向区域内用户提供数据服务; GISC 是 WIS 的核心功能中心, 收集和存储全部全球交换数据和元数据, 并提供对 WMO 所有计划及其他国际组织有关计划的全部数据和产品的访问接入。

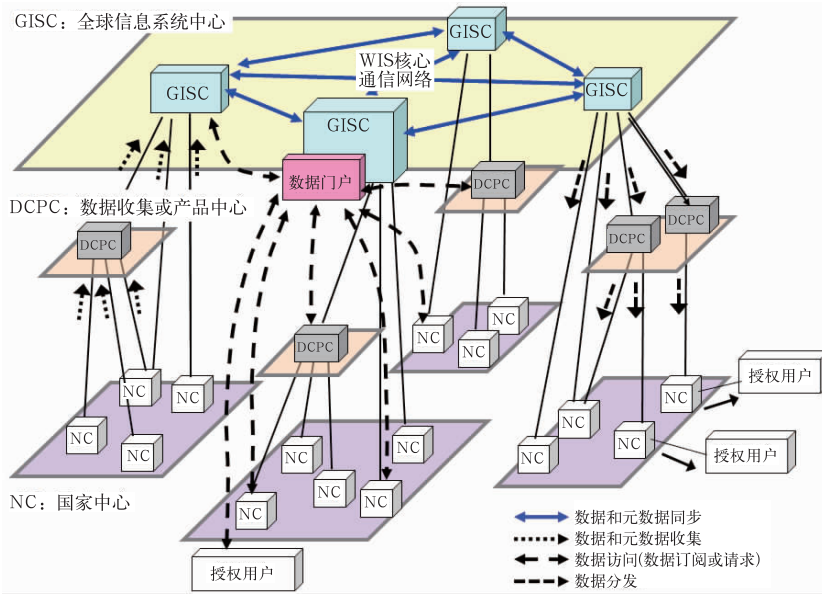


图 2 WIS 的结构和功能

Fig. 2 Functional structure of WIS

WIS 中, NC、DCPC、GISC 等名称仅仅用于描述 WIS 的相应功能, 并不代表实际的机构实体, 一个 WMO 会员或一个国际组织, 可以同时承担多项 WIS 功能, 建立和运行多个 WIS 功能中心^[1]。其中, NC、DCPC 的数量没有规定和限制, WMO 会员的国家气象中心、以及承担 WMO 和其他国际组织的有关计划和项目的国家中心都可以申请成为 WIS 国家中心; WMO 和其它国际组织的产品中心或数据收集中心都可以申请成为 DCPC, 例如: WMO 各会员的国家数值预报中心, 欧洲中期天气预报中心(ECMWF), 欧洲气象卫星组织(EUMETSAT)等。与 NC 和 DCPC 不同, 根据 WIS 规划, 一

个 WMO 区协(Regional Association)可以实施和运行一个或多个 GISC, 但是考虑实施的技术复杂性和运行成本, 全球范围内 GISC 的总数应不超过 10 个, GISC 总责任区应覆盖全球。

1.3 GISC 的关键功能

GISC 是支撑 WIS 数据交换和服务的核心功能中心。它收集责任区内提供全球交换的数据和产品, 与其他 GISC 交换全球数据, 向责任区内的 DCPC 和 NC 分发全球交换数据; 同时, 收集和存储 WIS 提供服务的全部数据和产品的元数据, 维护元数据目录, 提供数据发现服务; 另外, GISC 还在线

保存 24 小时内的全球交换数据,并为责任区 NC 和其他用户提供数据检索和访问服务。

GISC 系统的基本功能包括数据收集、数据分发、以及以元数据为支撑的数据服务等,功能结构如图 3。GISC 传输和服务的数据具有多学科、多类型、海量、以及形式多样、格式复杂、存储分布等特点。因此,利用元数据整合分布、异构的数据资源,依托元数据同步(收集和提供)等元数据互操作技术提供与数据存储位置无关的数据发现和访问服务,以及依托快速数据交换(收集和分发)机制支持关键业务数据高时效传输等,是 GISC 的关键功能,也是 WMO 现有信息系统不具备的功能。

2 GISC 关键功能的实现技术

2.1 元数据标准

元数据是描述数据的数据,随着网络技术的发展,元数据已经从最初对信息对象的描述和抽象扩展成贯穿于数据管理、数据交换、数据发现和用户使用等网络信息组织利用全过程的重要工具和方法。WIS 中,利用统一的元数据标准——WMO 核心元数据描述数据和产品,管理和整合分布存储在各 WIS 中心的异构数据资源,提供数据发现和访问服务。

WMO 核心元数据是面向应用的、描述数据内容特征及数据传输和服务特征的数据,属于 ISO 19115 的子集合。它引用了 ISO 19115 标准中全部 14 个实体的设置,保留了 ISO 19115 中大部分描述空间特征的元素,规定了数据描述中必选的核心元素和可选元素(见表 1),同时,它也依据 WIS 数据描述需求对 ISO 19115 中元素值域作了扩展,例如:为了描述气象观测站点分布不规则的特点,在数据集空间分辨率元素的值域中增加了 irregularPoints 选项等。

WMO 核心元数据采用 XML 格式表示,遵循 ISO 19139 XML schema 规范,是组织和管理 WIS

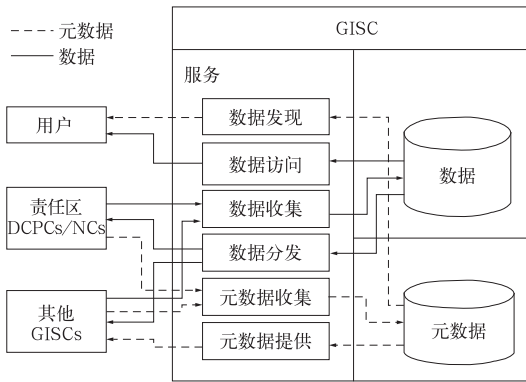


图 3 GISC 系统功能结构

Fig. 3 Schematic logical structure of GISC

表 1 WMO 核心元数据

Table 1 WMO core profile

必选	可选
Title(数据集名称)	Spatial Resolution(空间分辨率)
Identifier(数据集标识)	Data Reference System(参考系)
Reference Date(数据集制作日期)	Geographic Extent(地理范围)
Data Organisation(数据责任机构名称)	Geographic Name(地理描述名称)
Data Organisation role(数据责任方角色)	Vertical Extent(垂直跨度)
Abstract(摘要)	On-line Reserouce(在线资源)
Keywords(关键词)	Off-line Media(离线介质)
Topic Category(数据集类别)	Format Version(数据格式版本)
Data Point of Contact(数据责任方 联系方式)	Individual Name(责任人姓名)
Metadata Language(元数据语言)	Metadata File Identifier(元数据标识)
Metadata Character set(元数据字符集)	Metadata Standard Name(元数据标准)
Metadata Data Stamp(元数据制作时间)	Metadata Standard Version(元数据标准版本)
Metadata Organisation(元数据责任机构名称)	Data Lineage or Quality(数据历史/质量)
Metadata Organisation Role(元数据责任方角色)	Update Frequency(数据更新频次)
Date or Period(数据集时间或时间范围)	Bounding Box(东西南北边界)
Processing Level(数据集处理程度)	Bounding Polygon(边界范围)
Format Name(数据格式)	Access Rights or Restrictions(访问权限和限制)
Metadata Point of Contact(元数据责任方联系方式)	Position(责任人单位)

数据资源、提供的数据发现和检索服务的基础信息,也是用户了解、获取和使用 WIS 数据和产品的参考信息。

2.2 元数据同步技术

WIS 是典型的分布、异构信息系统,它由分布在全球数以百计的节点组成,为了提供与数据存储位置无关的数据发现、访问和检索服务,GISC 必须能够收集并存储 WIS 提供服务的全部元数据,并实现 GISC 之间以及 GISC 与其他功能中心之间的元数据同步。

OAI-PMH 协议的应用,是实现上述目标的有效解决方案。

OAI-PMH (Open Archive Initiative Protocol for Metadata Harvesting) 是由开放文档先导组织 (OAI) 提出的、在分布式网络环境中获取元数据信息的标准化协议^[6-7]。协议的核心是定义了数据提供者 (Data Provider) 和服务提供者 (Service Provider) 两个角色以及扮演角色的条件,其中,数据提供者负责元数据的生成、发布、管理和组织,拥有一个或多个仓储 (Repository),并向数据服务者提供元数据发布服务;数据服务者是元数据的收割方,通过向数据提供者发出请求,接收返回的元数据。

OAI-PMH 是建立在 HTTP 协议基础上的应用协议,它使用 HTTP GET/POST 方法发出请求,使用 HTTP 的状态码来表示请求/响应中的错误,同时 OAI-PMH 对 HTTP 协议进行了扩展,规定了获取提供者信息和获取元数据记录的 6 个命令动词,任何仓储只要支持 OAI-PMH 规定的 6 个命令动词的使用、支持 HTTP GET/POST 请求,就可成为数据提供者,而服务提供者则收割来自数据提供者的 XML 格式编码的元数据并在此基础上提供增值服务。

2.3 元数据访问技术

除 GISC 基于本地元数据目录提供数据发现服务外,分布存储元数据目录的检索和访问是适用于 NC 或 DCPC 依托责任区 GISC 元数据目录为其用户提供数据发现服务的实现技术。利用 SRW (Search/Retrieve for the Web)/SRU (Search/Retrieval via URL) 协议建立分布数据资源检索服务,实现对 WIS 各功能中心分布存储元数据的透明访问,是 WIS 提供与数据存储位置无关的数据发现和

访问服务的可用实现技术。

SRW (Search/Retrieve for the Web)/SRU (Search/Retrieval via URL) 是基于网络和 Web 服务的分布信息资源检索协议^[8],支持从简单的 JavaScript 到 XSLT 等的访问。其中,SRW 遵循 SOAP 通信协议,依据 Web Service 规范封装查询请求,采用 SOAP 消息交换机制和 XML 编码完成信息传递;SRU 的实现较 SRW 简单,它采用 HTTP/GET 方式在 URL 中发送请求,应答端采用 SOAP 定义的语法格式返回查询结果,支持 SOAP 消息中的内容序列,但不采用 SOAP 标签。SRW/SRU 通过提供通用的框架结构,实现了 Web 查询的标准化及查询结果的结构化,屏蔽了复杂多样的底层资源的组织管理及服务逻辑操作,具有良好的开放性和可扩展性。

2.4 数据同步技术

数据同步是基于网状通信连接和数据增量收集的数据传输机制,用于网络连接各节点之间关键业务数据的快速收集和交换。现有 GTS 系统中,数据交换采用收集、转发机制,数据需要由国家中心经区域内 RTH 向 GTS 主干网中心汇集,再经由主干网中心向其他区域转发,传输环节多、时效低,并且存在数据重复收集(通过不同路由收到相同数据)。在 WIS 中,GISC 之间建立了网状连接,通过数据同步,GISC 可以同时获取其他 GISC 新增收集和发布的数据,有效减少了数据传输环节和重复数据传输,包括全球交换数据在内的关键业务数据和产品的传输时效将得到明显提高。数据同步可以采用 HTTP、FTP 等传输协议实现。

3 北京 GISC 系统设计与实现

3.1 总体设计

北京 GISC 系统是承担 WIS 全球信息中心功能的业务系统,以元数据为核心组织数据资源和服务资源,通过统一的数据交换和传输服务接口,提供对本地数据和远程数据(国外中心存储数据)的发现、检索和访问服务,总体技术架构(见图 4)划分为运行环境层、数据资源层、服务层和应用层,同时,依据和遵循 WMO 核心元数据标准等 WIS 标准规范实施北京 GISC 系统建设。

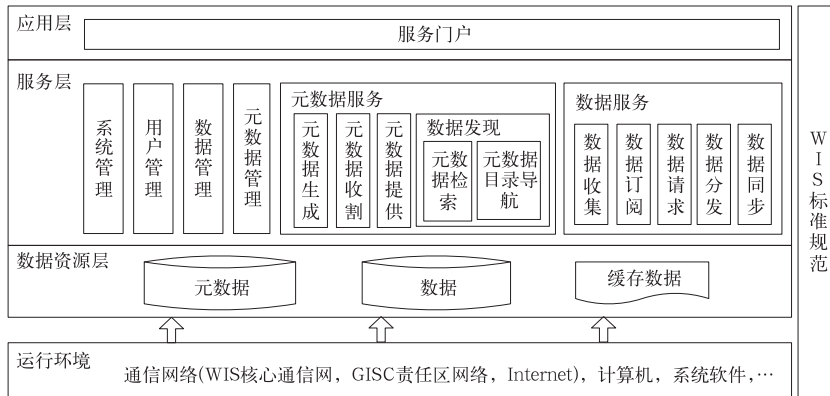


图 4 北京 GISC 系统总体技术架构

Fig. 4 Logical structure for the data exchange and data sharing system at GISC Beijing

运行环境层是支撑北京 GISC 系统运行的网络 and 系统平台,包括:连接北京 GISC 的 WIS 核心通信网络、责任区网络以及局域网络等网络环境,部署和运行北京 GISC 应用系统的计算机平台和存储设备,以及相关的操作系统和数据库系统等。数据资源层提供数据存储方式,是北京 GISC 系统的服务主体,包括:存储本地数据和产品的数据库、存储 24 小时全球交换数据的文件系统、以及存储元数据信息的元数据库等。服务层是业务逻辑实现层,提供北京 GISC 系统功能模块的划分和实现,主要包括系统管理、用户管理、数据管理、元数据管理、元数据服务和数据服务等功能部件。应用层是面向用户的服务展示层,通过服务门户提供北京 GISC 系统各种功能的操作界面,展现各种数据和相关元数据信息。

北京 GISC 系统基于 J2EE 平台设计,采用以 Web 服务为主的应用集成技术开发,主要特点包括:

(1) 基于统一元数据标准和元数据互操作技术实现了分布存储于各 WIS 功能中心的异构数据资源的整合。

(2) 基于元数据目录提供开放的数据发现服务。北京 GISC 门户通过元数据目录导航和元数据检索提供对全部 WIS 数据的发现和访问服务。

(3) 基于分布元数据检索服务实现对责任区内 NC、DCPC 提供 WIS 数据发现服务的支撑。责任区 NC、DCPC 可以利用分布检索访问北京 GISC 元数据目录,为其国内或责任区内用户提供 WIS 全部数据的发现和访问服务。

(4) 利用数据同步实现全球关键业务数据的收集,收集数据可以通过数据订阅、数据请求等方式提供访问,通过 FTP、HTTP 以及 SMTP 等协议提供

数据传输服务。

3.2 元数据生成

北京 GISC 现有元数据模板依据 WMO 核心元数据标准(版本 1.1)设计,主要包含元数据文件标识“fileIdentifier”、元数据创建时间“dateStamp”、元数据遵循的标准“metadataStandName”、以及数据描述信息“identificationInfo”等 15 个子元素,可以描述产品和数据的名称、时间、地理位置等属性、以及数据格式和数据获取方式、地址等数据服务信息,是北京 GISC 提供数据发现和访问服务的基础。

目前,北京 GISC 已生成并存储了约 11 万条元数据记录和对应的元数据文件,包括全球交换数据、CMA 数值预报产品、FY-2 卫星产品以及 TIGGE 产品的元数据。这些元数据文件由元数据生成软件根据数据和产品的元数据信息、元数据模板以及元数据文件命名规则自动生成。其中,全球交换数据的元数据文件依据 GTS 全球交换数据目录(Vol. C1)和全球站点信息表(Vol A)等元数据信息生成;CMA 数值预报产品、FY-2 卫星产品、TIGGE 产品,则需要首先按照规定格式制作元数据信息文件,再由元数据生成软件依据元数据信息文件生成。

3.3 基于 OAI-PMH 协议和 FTP 协议的元数据同步

3.3.1 GISC 元数据同步

北京 GISC 的元数据同步服务基于 OAI-PMH 协议实现,由元数据收割、元数据仓储和元数据提供等服务组成。元数据收割服务通过执行 ListSets、ListMetadataFormats、ListIdentifiers 和 GetRecord

等命令动词以及解析响应结果实现,提供自动收割和手工收割两种方式,支持增量收割;元数据仓储提供本地元数据和收割元数据的统一存储和管理,由元数据目录和元数据文件存储组成,元数据目录采用数据库管理,存储元数据标识、标题、摘要等基本元数据信息和元数据文件保存路径,并通过保存路径与元数据文件关联;元数据提供服务依据权限管理和控制规则发布元数据仓储中的元数据文件,通过解析和响应 Identify、ListSets、ListMetadataFormats、ListIdentifiers、ListRecords、GetRecord 等命令动词实现。

3.3.2 DCPC 和 NC 的元数据收集

对于收集 DCPC 和 NC 的元数据,北京 GISC 除支持 OAI-PMH 协议外,还支持 FTP 协议(见图 5)。元数据收集涉及对新增元数据、元数据更新信息、以及元数据删除信息的解析和处理。例如:当产品中心取消某一数据提供时,产品中心需要发布和提供元数据删除信息,以便所属 GISC 应能够及时从其元数据目录中心删去对应的元数据记录,并将删除信息同步或发布到其他 WIS 中心,使删除的元数据从整个 WIS 服务系统中消失。与 OAI-PMH 不同,FTP 协议本身没有提供对元数据变更信息解析和处理的支持,而 WIS 也尚未制定统一协议。因此,采用 FTP 传输和提供元数据的中心,需协商建立元数据提供和收集的双边协议规范。

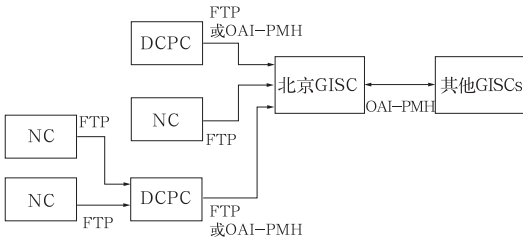


图 5 北京 GISC 元数据收集和同步的实现

Fig. 5 Implementation of metadata harvesting and metadata synchronization at GISC Beijing

北京 GISC 设计并实施了 FTP 元数据提供和收割协议,核心内容如下。

(1) 元数据文件命名:遵循 WIS/GTS 传输文件命名规则,由数据标识符指示码 pflag、数据标识符 productidentifier、数据中心标识符 C_CCCC、元数据文件生成时间 YYYYMMDDhhmmss、数据中心自定义描述 additionalPart、以及文件格式标识符“xml”和文件压缩方式等部分组成,格式如下:

《pflag》_《productidentifier》_ C_《CCCC》_《YYYYMMDDhhmmss》[_《additionalPart》].xml. [《compression》]

文件名中,除元数据文件生成时间、文件格式和文件压缩方式等字段外,其他字段编码格式和取值应与相应的数据文件相同。

(2) 新增元数据发布:数据中心或产品中心发布新增数据或产品时,应同时提供相应的元数据文件。

(3) 元数据更新信息的发布、收集和解析:元数据变更时,数据中心或产品中心应重新生成并提供相应的元数据文件,元数据生成时间(MD_Metadata/dateStamp)应依据元数据变更时间赋值。元数据收集中心通过增量收割收集元数据更新信息,依据元数据生成时间判别更新信息的有效性,根据判别结果更新元数据目录和元数据仓储中的相应记录。

(4) 元数据删除信息的发布、收集和解析:数据中心或产品中心停止提供某个或某些数据或产品时,应通过在相应元数据文件的文件名 additionalPart 字段中附加固定代码“DEL”发布元数据删除信息,即提供文件名为《pflag》_《productidentifier》_ C_《CCCC》_《YYYYMMDDhhmmss》[_《additionalPart》-DEL].xml. [《compression》]的元数据文件。元数据收集中心通过增量收割收集和解析元数据删除信息文件,从元数据目录中删除相应的元数据记录,并在元数据仓储中发布相应的删除信息。

3.4 基于 SRU 的元数据分布访问

为了满足责任区 NC 或 DCPC 的访问需求,北京 GISC 系统开发部署了 SRU 服务和基于 SRU 的分布式检索功能。

北京 GISC 通过由 SRU 转换器、检索器及

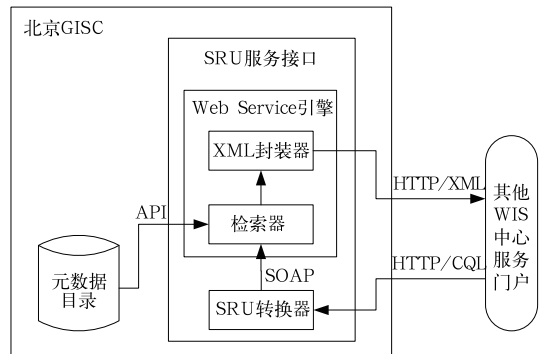


图 6 北京 GISC SRU 服务的实现

Fig. 6 Implementation of SRU service at GISC Beijing

XML 封装器组成的 SRU 服务接口提供 SRU 服务(见图 6)。其中,SRU 转换器负责接收用户通过其他 WIS 中心服务门户提交的 URL 请求,并转换为 SOAP 消息格式;检索器接收经转换器生成的 SOAP 消息,执行消息解析,生成并提交符合本地元数据目录检索接口的查询请求,获取检索结果;XML 封装器负责对结果信息进行结构化编码,生成检索结果 XML 文档,并返回提交请求的 WIS 中心。

另外,在北京 GISC 基于 SRU 的分布式检索功能的实现中,开发了解析和展示 XML 格式检索结果的模板,可以自定义风格展示 SRU 检索结果。

3.5 基于 ATOM 协议和 FTP 协议的数据同步

数据同步是基于网状通信连接和数据增量收集的数据传输机制,用于实现 GISC 之间全球交换数据的收集和交换。数据同步可以采用不同协议和技术实现。北京 GISC 系统开发并测试了基于 ATOM 协议的数据同步和基于 FTP 协议的数据同步。

ATOM 数据同步基于订阅机制实现,数据提供方首先要将气象数据文件上传到发布服务器,发布服务器基于 ATOM 连锁格式生成描述气象数据文件的种子,利用 ATOM 发布协议发布种子和气象数据文件,订阅服务器通过轮询获取 ATOM 种子,解析种子提取气象数据文件列表后,通过 HTTP 或 HTTPS 协议下载获取相应的气象数据文件^[9]。FTP 数据同步支持“拉取”和“推送”传输,并具备增量传输功能。

在中国气象局和德国气象局(DWD)已经开展的数据同步试验中,对基于 ATOM 的数据同步和基于 FTP 的数据同步进行了测试。试验使用相同测试数据集(182878 个文件,约 13 GB),针对不同的数据同步场景进行了多次测试。ATOM 测试场景包括:CMA 发布,DWD 订阅;DWD 发布,CMA 订阅。FTP 测试场景包括:CMA 推送,CMA 拉取,DWD 推送,DWD 拉取。测试结果表明,FTP 同步应用的传输性能和质量(文件丢失率)都好于 ATOM,而且 ATOM 同步机制相对复杂,不适于在高频次、大数据量的实时气象数据传输中应用。

根据中德数据同步试验结果,北京 GISC 确定采用 FTP 协议实施数据同步,并部署了 FTP 数据同步应用。

4 结 语

WIS 是综合、通用的全球信息系统,它的实施

将推动和促进全球天气、气候、水文资料的收集和共享服务。参与 WIS 实施,承担 GISC 职责,不仅可以提升中国气象局对全球资料的发现和获取能力,而且可以拓展各类国外资料的收集途径,为我国气象预报服务和科研提供更多、更高质量的数据支撑。目前,国家气象信息中心初步完成北京 GISC 系统建设,实现了与日本气象厅和德国气象局 WIS 系统的互联和元数据同步,并且已经第十六次世界气象大会批准成为首批全球信息系统中心。

目前,北京 GISC 系统能够提供全球交换资料、CMA 数值预报产品、FY-2 卫星产品、TIGGE 产品等实时和非实时资料的发现和访问服务;提供日本气象厅 MTSAT 卫星产品及 RSMC(Regional Specialized Meteorological Center)产品的发现和访问服务;以及提供德国气象局 GPCC(Global Precipitation Climatology Centre)、GCC(Global Collecting Centres for Marine Climatological Data)、RCC(WMO Regional Climate Centre)等中心的数据和产品的发现服务。以业务运行为目标,开展 GISC 服务备份机制和实现技术研究将是未来一段时间内 WIS 研发的重点任务。同时,随着 WIS 实施的推进,将有更多的 WIS 功能中心投入运行,北京 GISC 收集和提供服务的数据种类和数量也将进一步拓展和增加。

参考文献

- [1] WMO. WMO Information System(WIS)[EB/OL]. http://www.wmo.int/pages/prog/www/WIS/ref_docs_en.html, 2005.
- [2] WMO. WIS Project and Implementation Plan[EB/OL]. http://www.wmo.int/pages/prog/www/WIS/ref_docs_en.html, 2005.
- [3] 矫梅燕. 天气业务的现代化发展[J]. 气象, 2010, 36(7): 1-4.
- [4] 李湘. 气象通信系统发展与展望[J]. 气象, 2010, 36(7): 56-61.
- [5] 许健民, 杨军, 张志清, 等. 我国气象卫星的发展与应用[J]. 气象, 2010, 36(7): 94-100.
- [6] Carl Lagoze. The Open Archives Initiative Protocol for Metadata Harvesting[EB/OL]. <http://www.openarchives.org/OAI/openarchivesprotocol.html>, 2004.
- [7] 郭少友. 基于 OAI-PMH 的信息资源整合[J]. 大学图书馆学报, 2005, 23(3): 16-18.
- [8] 陈金莉, 王燕, 冯英, 等. 我国数字图书馆标准规范建设: Z39.50 协议应用指南[D]. 2004.
- [9] 王甫棣, 李湘. 基于 Atom 技术的气象数据交换[R]. 气象通信与信息技术委员会《提升数据服务能力, 引领气象信息系统全面发展》学术会议, 中国气象学会, 2009.