# 国境界河数据中心建设方案

### 总体设计方案

方案的总体目标是建立一个面向全国国境界河数据的管理系统。系统收集来自不同数据采集终端的数据，进行统一的存储、处理和访问。如图1所示，系统整体由数据存储层、数据处理层、数据接入层和数据应用层等四个部分组成。

图1 国境界河数据建设总体方案

数据接入层为用户提供丰富的数据接入方法，包括人工录入数据、卫星传输数据以及各个边境的音视频终端采集数据和各种传感器采集数据等，对结构化数据、半结构化数据和非结构化数据等的多种异构数据提供支持。在数据存储层建立全国分级数据存储中心，对来自全国的异构国境界河数据进行分级存储，满足用户对各种不同结构数据的存储及访问需求。数据处理层建立统一的数据处理中心，为用户提供包括批量数据、实时数据等大数据处理能力，对常规业务数据统计、日常报表制作、数据挖掘算法等提供有力的支撑和保障。数据应用层根据具体的业务需求，开发各种不同的应用，为用户提供丰富的功能，并提供多种可视化方案，包括面积&尺寸可视化、颜色可视化、图形可视化地域空间可视化、概念可视化等。系统提供安全访问控制策略对数据提供分级权限访问控制，保证数据的安全访问。

### 详细设计

#### 2.1基础设施建设

传统模式下，用户建设一套IT系统不仅仅需要购买硬件等基础设施，还需要购买软件的许可证，需要专门的运维人员来进行维护。当规模扩大时还需要继续升级各种软硬件设施以满足需要，这会导致运维成本增加、效率降低。由此云计算应用而生。云计算的最终目标是将计算、服务、应用作为一种公共设施提供给公众，让用户从“购买产品”转变为“购买服务”。许多用户都将自己的数据信息进行云迁移。但是面对种类繁多的云服务，需要合理规划，才能顺利实现基础设施的建设。众所周知，按照商业模式的不同，云计算可以被分为三大类：公有云、私有云和混合云，这三种模式构成了云基础设施构建和消费的基础。

* 公有云：公有云是面向大众提供计算资源的服务，公有云并非用户所有，由第三方提供云计算资源，用户通过互联网来使用资源。公有云的提供商有Amazon、微软、阿里云、腾讯云等。
* 私有云：私有云是企业传统数据中心的延伸和优化，是为一个客户单独使用而构建的，因而能对数据、安全性和服务质量的提供最有效的控制。私有云可以部署在企业数据中心的防火墙内，也可以将他们部署在一个安全的主机托管场所，私有云的核心属性是专有资源。
* 混合云：混合云融合了公有云和私有云，是近年来云计算的主要模式和发展方向。出于安全考虑，用户更愿意将数据存放在私有云中，但是同时又希望可以获得公有云的计算资源，在这种情况下混合云越来越多的被采用，它将公有云和私有云进行混合和匹配，以获得最佳的效果，这种个性化的解决方案，到达了既省钱又安全的目的。

国境界河数据中心使用混合云架构进行基础设施的建设，将核心数据存在私有云中，对外提供的服务运行在公有云中，达到节省成本、提高效率的目的。公有云可以直接购买供应商的资源，所以此处主要论述私有云的建设。

openstack现已成为主流私有云建设方案。Openstack是一整套开源软件项目的综合，它允许企业或服务提供者建立、运行自己的计算和存储设施。Openstack包括7个核心组件：Compute（计算）、Object Storage（对象存储）、Identity（身份认证）、Dashboard（仪表盘）、Block Storage（块存储）、Network（网络）、Image Service（镜像服务），他们之间的关系如图2所示。

* Compute：代号为Nova，最核心的组件，管理计算资源。
* Identity：代号为Keystone，提供身份认证和授权的组件。
* Dashboard：代号为Horizon，为所有的Openstack服务提供了一个模块化的用户界面，使用这个界面，可以完成绝大多数操作，例如启动实例、分配ip，设置访问控制等。
* Image Service：代号为Glance，镜像的存储服务。
* Network：代号为Neutron，网络管理组件，实现了软件定义网络。
* Object Storage：代号为Swift，对象存储组件。
* Block Storage：代号为Cinder，提供块存储服务，实现云硬盘的热插拔功能。

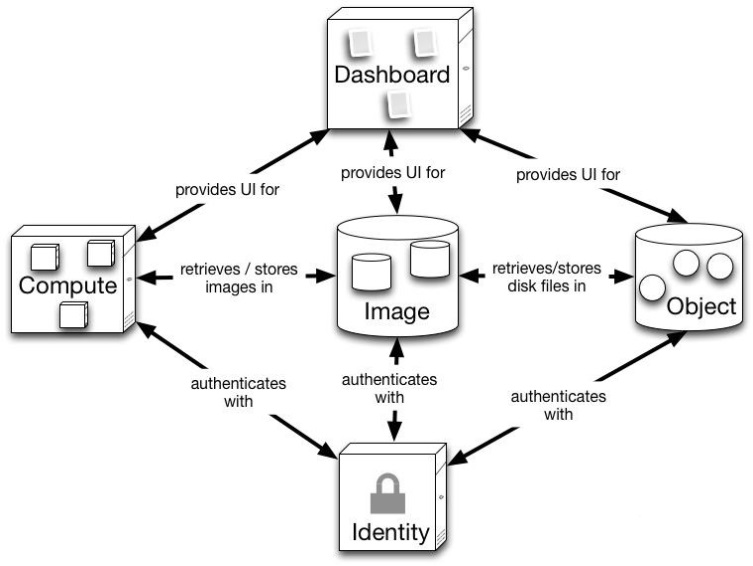


图2 openstack各个组件的关系

#### 2.2数据接入

数据接入系统作用是将多源异构数据接入到系统中。数据接入系统支持离线采集、实时采集、手工录入等功能。离线数据采集将大量预知信息（地理信息、设备信息、经验信息等）纳入管理系统中，为其他功能的实现提供数据支持。实时数据采集将已经发生的事件实时地传输到系统中，是整个系统中最重要的数据来源。同时，该系统支持后期手工录入功能，提供web界面录入和批量API录入的功能。数据接入中涉及的具体技术有网关、ETL、数据采集中间件、录入接口等。

**2.2.1 网关**

根据监测对象的不同，接入数据形态各异，例如有来自传感器的文本数据，也有来自无人机的音视频数据等等，这些多源异构数据在进入管理系统之前，需要网关来进行过滤控制转发。

**2.2.2 ETL**

ETL是将数据经过抽取、清洗、转换之后加载到数据仓库的过程，目的是将分散、零乱、标准不统一的数据整合到一起，是构建数据仓库重要的环节，系统从数据源抽取出所需要的数据，经过数据清洗，最终按照预先定义好的数据仓库模型，将数据加载到数据仓库中。

**2.2.3 数据采集中间件**

考虑系统对短时间内数据量剧增的应对方案，采用中间件的形式来减缓接入服务器的压力，提高数据整合的实时性。常见的开源中间件有Flume、Kafka。Flume是一个分布式、可靠、高可用的海量数据聚合系统，支持在系统中定制各类数据发送方，用于收集数据，同时提供对数据进行简单处理、写入各种数据接收方（可定制）的能力。Kafka是一种分布式的，基于发布/订阅的消息系统，具有高吞吐量、内置分区、支持数据副本和容错的特性，适合在大规模消息处理场景中使用。可以参考使用如图3架构：



图3 数据采集中间件示意图

**2.2.4 录入接口**

该系统提供录入接口，方便运维人员对该系统进行维护升级。录入接口分为web页面和api，web页面支持少量数据的方便录入，当数据量较大时，可以根据api编码录入。

#### 2.3数据存储

数据存储部分需要对全国范围内的国境界河的数据进行存储。系统中存在人工录入数据、卫星传输数据以及各个边境的音视频终端采集数据和各种传感器采集数据等多种不同的数据源，由于来自不同数据源的数据可能具有不同的数据结构，数据存储部分需要对多种数据源的数据提供支持。由于行业的特殊性，国境界河的数据存储中心需要包括地理信息库、水文数据库、实时水雨情库、工程信息库、人员信息库、社会经济信息库、基础地形库、动态影像库、历史数据库、方法库、监测信息数据库等。庞大的数据量要求数据中心需要具有海量数据的存储功能。综合考虑不同数据的存储和使用特点，建立以关系型数据库、nosql数据库和分布式文件存储为基础综合存储系统。整个存储系统由一个全国数据中心、多个省级数据中心和若干以行政区为单位的存储节点组成。为系统提高效、安全、分布式管理的功能特点，确保系统的稳定性。如图4所示。



图4 存储中心示意图

**2.3.1 数据类型**

系统中的数据主要包括人工录入的界河数据、历史经验数据、设备数据，卫星采集传输数据，音视频终端采集数据以及各种传感器采集数据。按照数据的组织形式，将数据分为结构化数据、半结构化数据和非结构化数据。如图5所示。



图5 数据格式

结构化数据，既行数据，可以将数据使用不同的字段进行描述，以行为单位进行存储。这种数据多存储在数据库里，可以使用二维表结构进行建模和表示。对于国境界河数据中心里存储的人员信息、设备信息、部分监测数据等可以使用结构化形式表示的数据，都可以使用结构化数据建模方法进行建模，统一存储到结构化数据库里。这种数据表示方式使数据的表示更加直观，查找和处理更加方便，应对简单数据的存储、处理和快速查找的需求。

半结构化数据，包括邮件、HTML、报表和资源库等。典型的应用案例包括邮件系统、档案系统和部分资源库系统。半结构化数据并不符合关系型数据库或其他数据库表的形式关联起来的模型结构，但是包含相关标记，用来分隔予以元素以及对记录和字段进行分层。这种数据表示结构也被成为自描述结构。常见的半结构数据表示格式包括xml和json格式。国境界河数据中心存储部分，无法使用结构化建模方法进行建模的部分实体数据、检测数据以及设备信息数据，使用xml和json数据格式进行存储和表示。半结构化数据表示方法给予数据表示更大的灵活性，更加自由的表示更多的属性信息。由于这种信息是自描述的，方便以后进行扩展。

非结构化数据，就是没有固定结构的数据。通过卫星传输的国境界河图片、影像信息，终端采集的音频、图像、视频信息都属于非结构化数据。系统中的部分国土GIS信息也属于数据非结构化数据。对于非结构化数据，无法使用有效的数据模型进行建模存储，一般直接使用二进制的方式进行存储。考虑这种信息会涉及全国范围内的大部分区域，并且随着时间推移数据量会不断增加，使用分布式存储方式对非结构化数据进行存储。这种存储方式具有存储量大、方便数据扩容和共享的特点。

**2.3.2 分级存储**

为了高效地对全国的国境界河数据进行存储管理，建设了包含全国数据中心、省级数据中心和行政区数据存储节点的三级数据存储系统。在每一级存储中心充分考虑该级存储所存数据的结构特点，综合考虑结构化数据的结构化特点、半结构化数据对数据表示的可扩展性、非结构化数据存储规模特点，建立以关系型数据库、nosql数据库和分布式存储为基础的存储中心。各级存储中心提供完善的容灾备份策略，根据不同的使用场景选择合适的容灾备份等级，对整个系统的高可用性提供有力的支撑。各级存储中心通过网络上报统计数据，逐级汇总，保证数据的高效存储和访问。

全国存储中心定期收集省级存储中心的统计数据，对全国的国境界河数据进行存储。接收并存储卫星传输的全国国境界河监控数据。

省级存储中心负责定时对下级各个存储节点的数据进行汇总，并且要实时监控下级存储中心的运行状态。使用省级存储中心的数据，可以了解本省内国境界河的整体状态。

由各级行政区存储节点负责本行政区内国境界河数据的存储。这一部分负责存储本行政区内的所有人员信息、设备信息以及本区内的所有监测数据。定时对检测数据进行汇总并传输到上一级存储中心。通过对行政区级别的存储节点数据进行查看，了解该行政区国境界河的整体状态。

#### 2.4数据处理



图4 数据处理

数据处理层对数据存储层所存储的人员信息、设备信息、终端监测信息和卫星传输信息进行处理，为数据访问层和数据展示层提供相应的数据计算服务。数据处理层提供批量数据计算和实时数据计算两种数据计算模型，对不同的业务模型提供支撑。

**2.4.1 批量处理**

批量数据处理工作主要用在应对复杂业务以及海量数据处理时，无需人工干预，仅需定期读入批量数据，然后完成相应的业务处理并进行归档。全国的国境界河数据具有规模大、数据结构复杂、数据增长速度快的特点。因此，对于这些数据的计算需要使用大数据处理模型。对于数据规模大、实时性要求不高的业务场景，适合使用批处理任务模型。这些主要的业务场景主要包括对于大范围内的国境界河信息的综合统计分析、对于某个地区的历史数据的统计分析以及各种定时数据分析任务。此类业务通常具有以下几个典型的特点：

1. 自动执行。每个任务都是根据提前设定的工作步骤自动完成。
2. 数据量大。此类任务所涉及的数据量少则百万，多则数千万甚至上亿。
3. 定时执行。根据设定的任务周期执行相应的任务。例如，每天执行、每周执行或者每月执行。

在国境界河数据中心通过基于map-reduce模型的hadoop大数据分析平台对于对海量数据的批处理分析提供了强有力的支持。Hadoop是一个开发和运行处理大规模数据的软件平台，是apache的一个使用java语言实现的开源软件框架，在由大量计算机组成的分布式集群中对海量数据进行分布式存储。Hadoop使用数据存储层的分布式存系统完成计算过程。支持关系型数据库、nosql数据库和分布式存储系统的读写操作。

**2.4.2 实时处理**

实时数据处理是指计算机对现场数据在其发生的实际时间进行收集和处理的过程。实时数据处理一般都是对海量数据进行的，一般要求处理时间为秒级。在对国境界河进行监控的过程中，实时数据是监测系统中最重要的组成部分。实时数据时一种带有时态性的数据。与普通的静止数据最大的区别在于实时数据带有严格的时间限制，一旦处于有效时间之外，数据价值将大大下降。但是随着时间推移，实时数据将会变成历史数据并存下来作为备用。例如，可以作为批处理任务的数据输入。整体流程如图5所示。为了充分利用国境界河的实时监测数据，系统提供了强大的实时计算能力对监测数据进行实时计算，实时反映被监测对象的状态。

图 5 数据实时处理

实时计算主要分为两个部分：数据的实时入库和数据的实时计算。数据实时入库由数据实时采集及数据实时接入提高支撑和保证。数据处理层通过提供storm和spark等实时数据处理平台对数据的实时计算提供支持。Storm是专门用来处理实时数据的计算平台。也被称作实时处理领域的Hadoop。Strom在设计思路中充分考虑到大规模可扩展能力、利用一套“故障快速、自动重启”方案为处理提供了容错性支持，从而有力的保证了每个元组都得到切实的处理。Spark是一个围绕速度、易用性和复杂性分析构建的大数据处理框架。Spark Streaming是Spark常用的组件之一，对离散化的流式数据处理提供了强有力的支持。

**2.4.3数据分析**

国境界河数据中心提供为国境界河的研究工作提供了丰富的数据源。其中，包括通过卫星传输的图像数据、通过监控终端传输的音视频数据以及各种传感器数据。系统提供了基于机器学习和深度学习的数据融合技术、图像处理技术、音视频处理技术，动态构建数据挖掘模型，为国境界河的监测提供科学指导，为河流流量监测、水质监测、水位监测、空气质量监测、国土质量监测和洪涝干旱灾害预警等提供重要的技术手段，实现国境界河管理向科学管理、动态管理、精细管理和智能管理的转变。

图 6 算法模型

如图6所示，使用数据融合技术对系统中存储的文本数据、音频数据、图像数据、视频数据进行多维度、多语义的数据融合。充分利用不同结构数据的表示特点，以挖掘潜在信息价值为目标将多种数据源中的相关数据提取、融合、梳理、整合成一个优势互补的分析数据集。这个分析数据集是个独立、灵活的实体，可随着数据的变化而重装、调整和更新。

目前多源的数据融合方法按照已有的工作可以分成三大类。第一类是阶段性的方法，先用一种数据，再用一种数据。第二个是基于特征拼接的方法。我们所熟知的深度学习方法，还有传统的特征串联加上一些正则化的方法，都是属于这里面的分支。第三类是基于语义信息融合的方法。这里包括了多视角、基于概率学模型的方法、基于相似度方法、以及迁移学习的方法。这种方法是根据人的思维方式设计的类人思考方法，所以是基于语义信息的方法。

为了充分挖掘数据集中潜在的信息价值，使用基于机器学习和深度学习的方法对融合后的数据集进行数据分析和挖掘，建立全方位、智能化、细粒度的国境界河自然特征监测模型和风险评估预测模型。国境界河信息与地理信息系统有着密不可分的联系，将地理信息系统作为国境界河数据分析过程中的一种辅助资源和手段，可以提高数据分析结果的准确度，提高信息处理效率。将数据分析与地理信息系统相结合，为国境界河信息管理的标准化、网络化、空间化提供了有效的工具。在对国境界河进行管理的过程中，综合运用上述数据模型分析各项管理战略和开发方案的影响，同时结合当地的自然和社会条件，最终拟定国境界河资源管理的最优方案。

传统的机器学习方法包括决策树算法、SVM算法、朴素贝叶斯算法、聚类算法和回归算法等。对于这些经典模型的有效使用，可以提高数据分析的准确率。目前基于人工神经网络的深度学习方法在文本分析、图像分析、音频分析和视频分析等方面表现出较高的准确率，在业内有广泛的应用。将深度学习算法应用到国境界河监测数据模型的构建过程，提高了数据分析模型的准确率。根据数据分析模型的具体业务场景，将数据分析模型分为离线模型和线上模型。同时，系统提供用户与数据分析模型的智能交互反馈机制，使用用户能够及时、持续的对模型进行改进，提升数据分析模型的准确率。

#### 2.5数据应用

在数据接入、存储和处理的基础之上，开发数据应用层，充分利用采集信息和处理产生的信息来支撑数据应用。各个数据应用采用微服务的方式构建，运行于容器之上，根据访问量实现应用的横向扩展、动态扩容等功能。涉及的关键技术有容器、微服务、日志分析等，简述如下。

**2.5.1 容器**

继虚拟化技术出现后，容器技术逐渐成为对云计算领域具有深远影响的变革技术，容器技术的发展和应用，为各行业应用云计算提供了新思路，同时容器技术也对云计算的交付方式、效率、Paas平台的构建等方面产生深远的影响，具体体现在简化部署、快速启动、服务组合、易于迁移等方面。

**2.5.2 微服务**

微服务是一种架构风格，一个大型复杂软件应用由一个或多个微服务构成，系统中的各个微服务可被独立部署，各个微服务之间是松耦合的。每个微服务仅关注于一件任务并很好地完成该任务，每个微服务代表着一个小的业务能力。

微服务的思路是将单一的巨大应用拆分为众多松耦合的微小服务，通常是按照业务功能来分解的，每一个服务虽然小但却实现完整的功能，可以单独构建和部署，单个服务的修改和部署不影响其他正在运行的服务，提供语言无关的API接口供其他模块调用，这种风格与传统的面向服务架构SOA比较相似，经过多年的发展，SOAP、Web Services、ESB等技术的出现使SOA得以实现，众多厂商也制定了相关标准，两者最重要的区别在于SOA使用复杂的ESB集成为单一应用，而微服务是轻量级的，松耦合，可以独立部署。微服务可由不同的团队独立开发，互不影响，而且微服务架构对持续交付是一种巨大的推动力，允许在频繁发布不同服务的同时保证系统其他部分的可用性和稳定性。

该系统数据应用层会根据不同时期的需求，不断的添加新的功能，若使用传统的一体式架构，则更新和修复该应用会变得越来越困难，微服务架构结合容器技术正好解决该痛点。

**2.5.3 缓存**

在高并发应用中，把业务数据持久化后，应用从持久化设备中读取数据并在前端显示，随着用户量、数据量增大，访问的集中，会出现持久化设备负担过重（典型的就是数据库），影响应用响应速度，应用延迟严重等重大问题。这时候就需要一种缓存机制提高并发读取速度的性能，memcached能在大中型系统中提供优秀的缓存服务。Memcached是高性能的分布式内存服务器，其分布式缓存访问模型如图7所示：



图7 分布式缓存访问模型

应用程序通过Memcached客户端访问Memcached服务器集群，Memcached客户端主要由一组API、Memcached服务器集群路由算法、Memcached服务器集群列表及通信模块构成。

**2.5.4 日志系统**

对于一个大型系统，收集记录分析日志十分重要，对于整个系统的稳定运行、业务应用的改进提升都有积极的意义。从业务的角度，可以把日志分为这几类：

* 业务日志：用户的操作日志，主要来监控业务逻辑的执行情况
* 错误日志：系统的错误信息
* 摘要日志：系统操作上下文的摘要信息
* 统计日志：汇总的统计信息

由于日志信息繁杂，从中提取出有用信息变得十分困难，搭建分布式的日志系统是解决这个问题的途径，可以使用ELK等成熟的日志框架对日志进行聚合分析等操作，如图8所示：



图8 日志分析系统解决方案

图中，各个子模块中都集成Logstash的Shipper，收集日志后发送到中间件，然后传输到Logstash集群中，通过ElasticSearch进行索引，并提供web页面和接口进行查询。

**2.5.5 访问控制系统**

访问控制系统是针对越权使用资源的防御措施。防止对资源进行未授权的访问，从而使系统在合法范围内使用。

访问控制的主要类型有：

* 物理访问控制
* 网络访问控制
* 操作系统访问控制
* 数据库访问控制
* 应用系统访问控制

国境界河监测系统是分层分级的，针对不同的层级有不同的用户，不同用户的权限不同，因此主要采用网络访问控制和应用系统访问控制。网络访问控制主要包括入网访问控制策略、操作权限控制策略、目录安全控制策略、属性安全控制策略、网络非法入侵检测等；应用系统访问控制是指对于不同权限的用户，系统呈现和提供的服务完全相同。