

4. 基于大数据技术的宁波市居民电子健康档案建设

宁波市卫生部门承担着全市的公共卫生、疾病防治、计生服务、各级医院监管等多方位管理任务，我们以不断提升医疗服务、增强群众获得感为出发点和落脚点，通过努力，已初步实现所有公立医疗卫生机构的信息互联，人口健康信息平台已将下属各部门的数据收集整理基本完毕，包括居民的基本健康信息、各级医院的病历数据、疾控部门的传染病和慢性病数据等。日益递增的庞大医疗健康数据迫切需要一个健康大数据平台进行储存和运作，同时建立大数据下的电子健康档案体系，并在此基础上实现对数据进行多位关联分析、挖掘数据的潜在价值，充分发挥健康大数据的作用和功能。2016年9月，原宁波市卫计委、市发改委、市智慧办联合印发了《宁波市智慧健康保障体系二期建设规划》，将医疗健康大数据建设纳入了重点工程之一。

2016、2017年，宁波市卫生部门先后成立了“一中心两所”，即健康医疗大数据研究中心、公共卫生大数据研究所和医疗大数据研究所作为机构保障，同时加快了大数据人才培养，以及与教学、科研和大数据研究机构的合作。我们积极探索，提出了大数据平台整体解决方案，利用新的大数据技术，构建健康大数据平台，来承载不断完善增长的居民电子健康档案数据。最终实现数据统一融合、关联共享、分析研究等目的，并且以极高的运行效率、比原有系统快一百倍左右的运行速度来保证此项目的顺利运行。



图 4.1 宁波市卫生信息中心 宁波市健康医疗大数据研究中心

4.1 建设内容概况

(1) 大数据基础平台

大数据基础平台的目的是为所有数据提供一个容器，为其他项目提供硬件支撑。以后所有数据都将存储在这个服务器集群之中，也方便了数据管理和后续开发。要求是：运行速度足够支持全市数据的大数据分析；扩展性好，能在短时间内无需大量资金人力即可完成数据扩容；能利用原有硬件，进一步节约成本。

(2) 数据库整合

数据库整合的目的是将各部门数据库统一导入大数据平台的 Hadoop 集群分布式数据库中，并为下一步继续数据分析展示挖掘做好准备。要求是所有数据已经形成一个分布式稀疏表，此表能够正常运行在上面提到的大数据基础平台之中。不同数据来源的数据能够正常兼容不会出现逻辑错误。

(3) 大数据交换平台建设

在大数据基础平台和数据库整合的基础上建设大数据交换平台。要求是对整体数据的精确查询应为毫秒级、模糊查询应为秒级；平台要有完备的数据安全防护措施、不同部门之间调用数据需要有权限并有记录；支持数据自动备份，备份份数可以自定义，并且能区分冷热数据；提供对外开放的数据接口。

(4) 多维度数据可视化分析

多维度可视化分析同样以大数据基础平台和数据库整合的内容为基础，目的是将整合好的数据图表化。同时将不同部门的数据逻辑关联起来，比如以自然人为单位将疾控数据和医院就医数据关联起来，然后从多角度多维度进行分析，从而发现卫生工作中的问题。要求是需要支持任意维度分析，数据调取时间应为分钟级，数据可以做到实时分析，更改分析条件时需要毫秒级响应，生成图标报表应为秒级响应。

4.2 Hadoop 技术架构构建健康大数据平台

Hadoop 是一个由 Apache 基金会所开发的分布式系统基础架构，是一个能够对大量数据进行分布式处理的软件框架。Hadoop 是一整套分布式处理软件，其生态圈包括 HDFS 文件系统，MapReduce 分布式处理框架，HBase 面向列式存储的数据库，Hive 数据仓库等多种组件。Hadoop 的框架最核心的设计就是：HDFS 和 MapReduce。HDFS 为海量的数据提供了存储，则 MapReduce 为海量的数据提供了计算，利用其高可靠性、高扩展性、高效性、高容错性。但由于其门槛及高维护成本也高，在众多条件衡量之下，我们最后选用 EDH (EDH DATA HUB) 大数据平台为底层数据存储架构。其特点是封装了大数据技术非常复杂

的操作,为所有服务提供 SQL 操作。数据存储放在了 EDH 的 Hyperbase 组件中,然后利用 EDH 的 InceptorSQL 提供交互式查询。

建设遵循 Internet/Intranet 国际标准,采用 Webservice、Kafka、Stream 等技术实现数据的采集。通过多种信息采集接口,采用有效的数据抽取技术,实现信息系统间、相关结构间数据资源的有效交流与共享,健康档案数据的实时采集和导入导出。

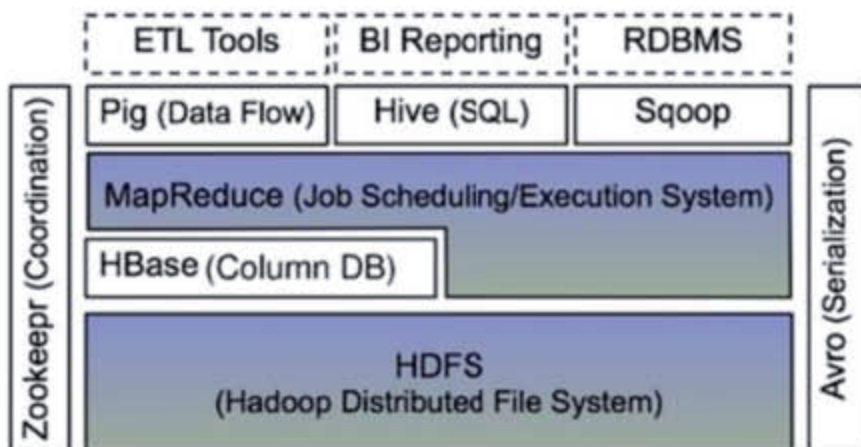


图 4.2 Hadoop 基本框架

Hadoop 技术架构

传统计算结构图

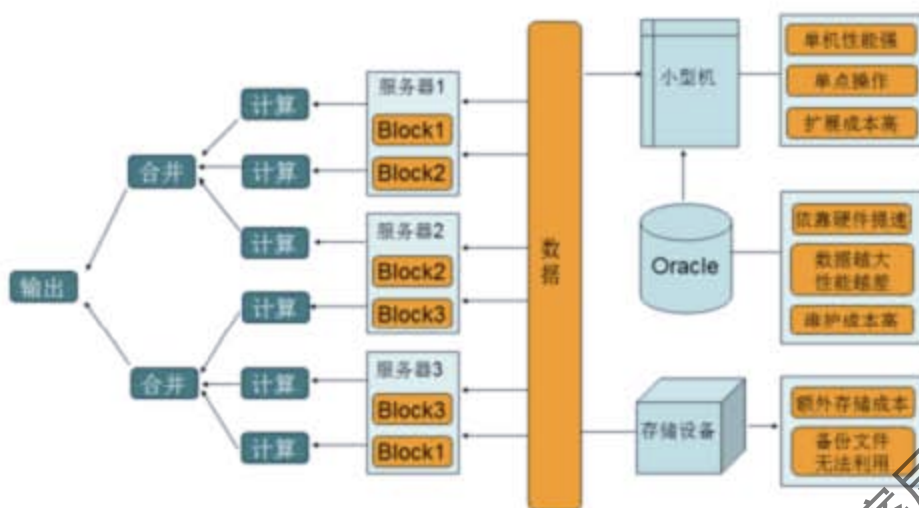


图 4.3 Hadoop 技术架构传统计算结构图比较

4.2.1 大数据平台硬件建设

在大数据数据中心机房建设中，我们设计了三种角色服务器，分别是：控制节点、接入节点、计算节点。三种角色服务器结合成必要的网络设备、存储资源共同构成云平台。

控制节点：对整个大数据平台进行控制与资源调度、运算。包括对 CPU、内存、网络、磁盘等资源的控制和管理。需要强劲的 CPU 处理能力，及较高的磁盘 IO 能力。因此，控制节点会配置 SSD。

接入节点：提供大数据平台所需的网络资源，包括基础网络、私有网络、路由器，以及其他网络功能。

计算节点：用于计算和存储处理所需要的数据，同时采用分布式的架构。

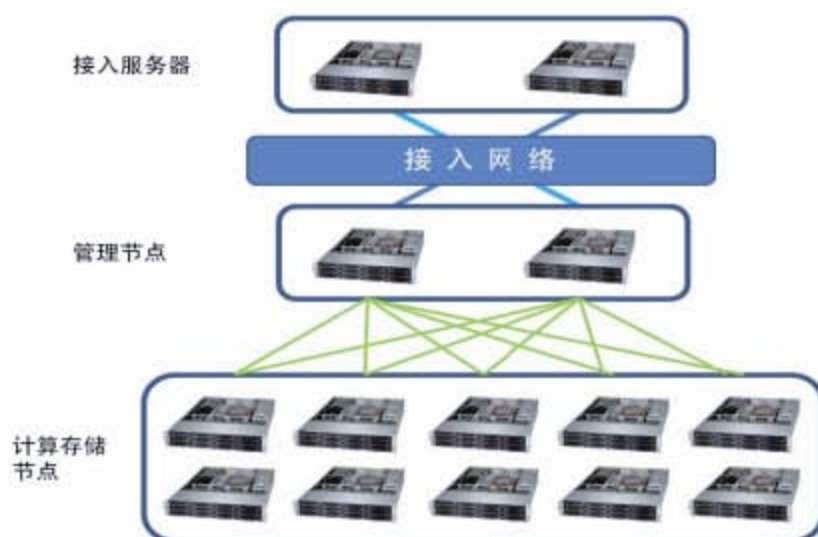


图 4.4 大数据平台物理架构

(1) 硬件规划原则。集中存储和计算平台使用基于 hadoop 分布式块存储和计算系统，因此，不需要采购传统的商业存储。由于计算存储和网络融合，单台机器需要较为强劲的 CPU，每台机器需要 4 口

千兆网卡和 1 个专用管理接口。内部交换机部分，每个机柜需要 2 台带 4 口 10G 上行端口和 48 口千兆下行端口的交换机。

(2) 物理服务器配置

表 4.1 本地大数据平台：

设备	配置说明	数 量 (台)
管理 节点 服务器	架构： 2U；双电源；8 盘位（SSD 盘不占用外部存储槽位）； CPU：E5-2650 v4 * 2； 内存：256G（16 * 16）以上； 硬盘：2TB SAS 硬盘 * 6 硬盘：Intel 480GB SSD 硬盘 * 2； 网卡：板载双万兆，四千兆端口； 远程管理卡：基于以太网的远程管理卡； 电源：1100w 双电（冗余电源）	2
接入 节点 服务器	架构： 2U；双电源；8 盘位（SSD 盘不占用外部存储槽位）； CPU：E5-2630 v4 * 2； 内存：32G（16 * 8）以上； 硬盘：2TB SATA 硬盘 * 4 千兆网卡：板载，双端口； 远程管理卡：基于以太网的远程管理卡； 电源：1100w 双电（冗余电源）	2
计算 节点 服务器	架构： 2U；双电源；8 盘位（SSD 盘不占用外部存储槽位）； CPU：E5-2650 v4 * 2； 内存：256G（16 * 16）以上； 硬盘：4TB SAS 硬盘 * 6 硬盘：Intel 960GB SSD 硬盘 * 2； 网卡：板载双万兆，四千兆端口； 远程管理卡：基于以太网的远程管理卡；	

	电源: 1100w 双电 (冗余电源)	
应用 服务器	架构: 2U; 双电源; 8 盘位 (SSD 盘不占用外部存储槽位); CPU: E5-2620 v4 * 2; 内存: 128G (16 * 8) 以上; 硬盘: 2TB SAS 硬盘 * 6 硬盘: Intel 480GB SSD 硬盘 * 1; 网卡: 板载双万兆, 四千兆端口; 远程管理卡: 基于以太网的远程管理卡; 电源: 1100w 双电 (冗余电源)	2
内部 网络	以太网交换机, 用于内部和外部网络交换, 48 个千兆交换口, 4 个上行万兆交换口	2

*数据库查询指标可以达到千万级秒级查询

4.2.2 技术要求

(1) 软件平台: 项目基于 B/S 架构实现, 采用 MVC 开发模式, 大数据集群要求用 Hadoop 保证数据的高可用、高性能。

(2) 系统平台: 操作系统要求 CentOS 操作系统; 数据库平台要求支持主流大型数据库, 选定 EDH。

(3) 数据要求: 支持数据并行处理技术; 含高可靠性集群, 支持节点间负载均衡; 支持超大规模的数据库设计和存储, 具有处理 TB 级大容量数据的能力; 做到数据“传且只传一次”、“到且仅到一次”即“不重复、不漏包”, 实现高效、稳定、完整传输。

(4) 数据安全性: 对于大数据中的 InceptorSQL、Hyperbase、HDFS、Kafka 进行 Kerberos 加密访问, 其中: HDFS 提供文件和目录的“rwx”权限控制; Hyperbase 提供全局、Namespace, Table, Column Family 和 Column Qualifier 级别的“RWCXA(Read, Write, Create, Exec, Admin)”权限控制; Inceptor 提供基于 SQL 的全局、数据库、表和行级别的 SELECT/CREATE/DELETE/UPDATE/INSERT 权限控制;

Kafka 提供全局或 Topic 级别的 RWCD (READ, WRITE, CREATE, DELETE) 权限控制。

(5) 完整日志：对数据传输、加载、查询及增、删、改的情况进行记录。对于交换中心，能知道本系统内所有经过交换的数据的基本信息；对于某一数据而言，能对数据的传输、访问路径进行跟踪，以检查此数据交换的历史痕迹。

4.2.3 大数据技术架构

EDATA 的大数据综合平台（简称 EDH）包含以下组成部分：Hadoop 企业版，Inceptor 分布式内存分析引擎，Hyperbase 分布式数据库、SlipStream 流处理引擎、Discover 数据挖掘机器学习等主要组件。



图 4.5 EDH (EDATA DATA HUB) 架构图

EDH 具有高模块化和松耦合的五层架构，针对不同的应用领域通过组件之间的灵活组合与高效协作来提供定制化的支撑。

是一个能够对大量数据进行分布式处理的软件框架，能够以一种可靠、高效、可伸缩的方式进行数据处理。Hadoop 通用平台为大数据平台提供分布式系统基础架构，可以在不了解分布式系统细节的情况下，开发分布式程序，充分利用集群的威力进行高速运算和存储。

4.2.3.1 数据分析与处理

数据分析与处理是整个大数据平台主体部分,包含下面几大部分:

数据 ETL: 用来实现对外部数据源的抽取 (Extract)、转换 (Transform)、加载 (Load) 到大数据平台数据仓库中。大数据平台面临的外部数据源数据格式繁杂,包括了结构化、半结构化、非结构化的数据,数据质量的不尽相同,不可避免会存在重复、无效或不完整的情况,经过大数据 ETL,包括数据提取、数据转化、数据加载、数据标准化、数据去重、数据过滤等等操作,最终按照上层应用预定义的数据模型加载到数据仓库中。

NoSQL 数据库: 提供高可用性、良好的扩展性的 NoSQL 数据库服务;支持 SQL,支持 JDBC/ODBC;允许根据具体的业务逻辑灵活的建立二级索引,在实际应用中往往会建立多个二级索引,在使用时会根据查询条件动态计算使用索引的代价,自动选择最合适的索引;支持分布式全文检索,可以为自己的表创建全文索引,并在 SQL 中使用全文检索语法进行搜索。

分析挖掘算法库: 提供上层应用在数据挖掘中所需要通用算法集,这些分析挖掘算法是根据数据创建数据挖掘模型的一组试探法和计算,运行于数据处理引擎之上,针对应用模型提取可行模式和详细统计信息。大数据平台在分析挖掘算法库中,提供了算法管理相关手段,以及用于验证算法准确度的测试框架。

SQL on Hadoop: 为 Hadoop 数据架构提供 SQL 能力。SQL on Hadoop 是目前大数据技术发展的一个重要方向,SQL 与 Hadoop 结合一方面可以发挥 SQL 易于使用的特性,避免编写 MapReduce 任务的复杂度高、部署周期长的问题,上层应用可以直接通过 SQL 进行数据分析;另一方面 Hadoop 平台具备很强的鲁棒性和可扩展性,是之前单独使用数

数据库技术难以获取的能力，兼容 HiveQL 和 SQL2003 语法，适用于数据仓库批量统计应用，支持对数据仓库中二维表的增删查改，可高速处理磁盘表或者缓存在内存或 SSD 上的列式数据。

OLAP：为上层应用提供联机分析处理能力，在这部分共享多维信息，针对特定问题的联机数据进行访问和分析，是上层 BI 类业务应用的重要依赖。OLAP 专门设计用于支持复杂的分析操作，侧重对决策人员和高层管理人员的决策支持，可以根据分析人员的要求快速、灵活地进行大数据量的复杂查询处理，并且以一种直观而易懂的形式将查询结果提供给决策人员，以便他们实现业务状况的准确掌握，了解对象的需求，制定正确的方案。联机分析处理具有灵活的分析功能、直观的数据操作和分析结果可视化表示等突出优点，从而使用户对基于大量复杂数据的分析变得轻松而高效，以利于迅速做出正确判断。它可用于证实人们提出的复杂的假设，其结果是以图形或者表格的形式来表示的对信息的总结。同时可以通过 Inceptor 组件采用 SQL 进行批处理和高并发查询，可从 Hyperbase 的行存储转换成 Holodesk 的列存储，同时支持在线查询和高速 OLAP 分析。

数据管理：围绕大数据平台的数据提供组织策略、操作管理、数据权限、配置管理、发布管理、服务管理等等功能，实现丰富的管理手段，便于对平台中数据进行修改、更新与扩充，同时保证数据的独立性、可靠性与完整性，减少数据冗余，提高数据共享程度及数据管理效率，使管理员能够对整个平台中的数据进行更好的管理和控制。

数据安全：提供应用安全、信息安全、网络安全、身份认证、加密及操作系统安全等等保障，为大数据平台建立安全保护，保护大数据平台整个集群硬件、软件和数据不因偶然和恶意的原因遭到破坏、

更改和泄露。

4.2.3.2 数据服务接口

为适应上层各类智慧健康大数据平台应用的数据服务要求，大数据平台提供的数据服务接口包括：

(1) SQL on Hadoop: SQL 2003 语法支持以及 PL/SQL 过程语言扩展,并且可以自动识别 HiveQL、SQL 2003 和 PL/SQL 语法,在保持跟 Hive 兼容的同时提供更强大的 SQL 支持,实现对 Hadoop 平台包括 HDFS 文件、HBase 数据库中的数据通过 SQL 进行访问,简化上层应用的数据分析操作;

(2) BI: 提供 JDBC/ODBC 支持,完全兼容目前主流 BI 数据展现组件,为上层应用的 BI 开发提供支持。

4.2.3.3 Hadoop 管理

作为 Apache Hadoop 通用 Hadoop 平台功能的补充,提供各类管理功能,为大数据管理员提供丰富的管理手段,满足大数据平台在可用性、易用性等方面的要求。管理的功能主要包括:应用管理,主要是实现对平台上运行的各类数据分析挖掘应用进行提交、审批及发布运行等管理,目的是为了控制整个大数据平台中的使用,避免未经授权的应用随意在大数据平台上执行,浪费不必要的计算资源,同时也能保证整个大数据平台集群的安全;作业调度,对 Hadoop 集群中的作业进行审查,根据资源需求,按照一定的规则进行执行作业;运行监控,主要是方便平台管理员和应用所有者对当前应用的执行情况进行跟踪;状态告警,可以根据告警规则,通过多种报警手段(如邮件、短信等)方式进行实时通知,其中的告警规则由管理员进行自定义,并且可以对告警进行合并、过滤等操作;认证与授权,用于管理整个

大数据平台的用户、角色，实现用户身份认证，通过授予不同用户权限控制对大数据平台的使用权限；资源管理，用于对群集中各节点的资源（包括 CPU、内存、负载、网络 IO 等）使用情况进行图形化展示，便于用户了解当前集群的资源使用情况；安全管理，用于增强 Hadoop 的安全控制机制，通过提供客户端对服务端、服务端对服务端的安全认证方案，可有效避免非法入侵和非授权访问。

4.3 居民电子健康档案数据采集和导入

居民电子健康档案的大数据建设，首先是应满足健康数据能放到大数据平台中。对于历史数据可以直接利用 Sqoop 从原来 Oracle 数据库中直接同步过来，而新数据应该直接从医院端采集，需要采用一套高可用、高性能、高扩展性的分布式采集方案。

4.3.1 数据采集导入技术架构

利用 EDH 的 Stream 流计算引擎来实现数据采集，Stream 是基于 Spark Streaming 的产品级批处理流计算引擎，将实时数据的处理分成一系列的小批量任务，并由 Spark 框架执行。Stream 是一个批处理流计算引擎，其数据处理的延迟在 100 毫秒到 2 秒之间，因此可以满足我们当前对采集实时处理数据的场景。同时，采用 Kafka 作为 Stream 的数据源来实现整体的采集方案，Kafka 是一种分布式的，基于发布/订阅的消息系统。

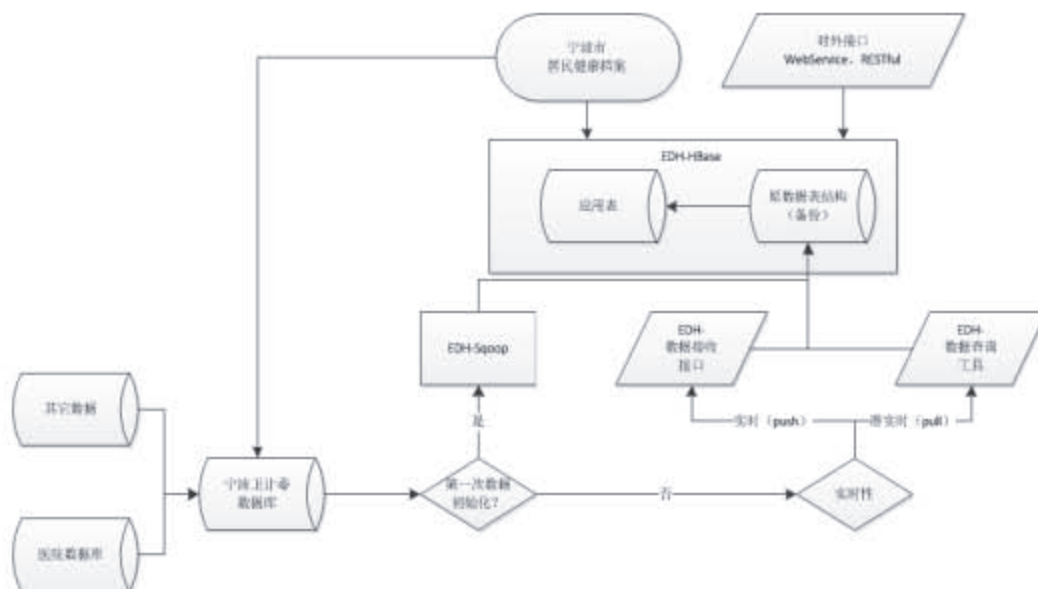


图 4.6 数据采集导入图

4.3.2 采集场景说明

由于医院平台、数据、采集方式的多样化，无法统一采集方案，针对于现有情况，提供五种数据源的采集方式。而对于数据消费，可以有多个消费者同时消费实时的 Kafka 数据，例如流数据实时处理 Stream 做实时分析、原始数据备份至 HBase、HDFS。以下是医院数据采集架构图：

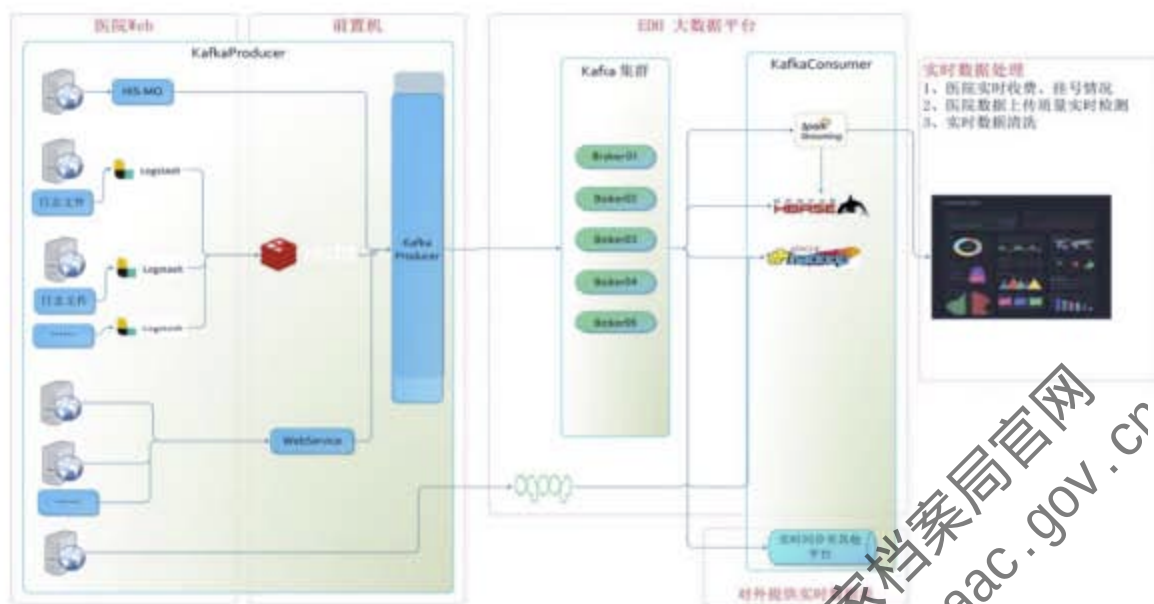


图 4.7 医院数据采集架构图

(1) 实时采集—HIS-MQ

如果医院 HIS 有类似消息队列的，可以把他们消息数据发送到他们的 MQ 中，然后需要推送方的处理程序具备容错和续传机制，保证服务重启后能继续发送数据。这样可以用程序消费 MQ 的数据，并且实时把数据采集到 kafka 中，这样可以有效的保证数据到采集端的完整性。其优点是不需要对原医院系统进行改造，不需要部署额外数据采集器插件，但需要推送方的处理程序。

此方案实施准备工作：需要前置机，不需要部署插件，只需要把数据放到 MQ 中，并提供 MQ 地址，配置 MQ 通道到 Kafka。

(2) 日志采集

如医院采用日志传输数据，首先得统一日志格式规范；并且根据医院应用、服务器性能来选择具体的数据采集工具。此方法对原医院平台改动最小，但需要记录日志，对于服务器数据存储占用了一部分资源，还需要额外部署插件，占用一定服务器资源。

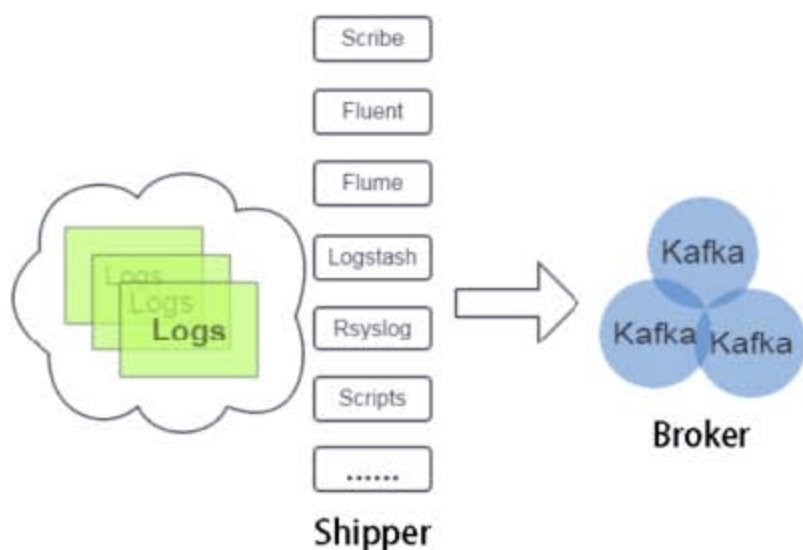


图 4.8 日志采集结构图

对于日志采集,一般采用 logstash/flume/ fluentd,具体根据实际情况而定。

	Flume	LogStash
结构上	Source、Channel、Sink	Shipper、Broker、Indexer
简易程度	很繁琐,要分别作 source、channel、sink 的手工配置,而且涉及到复杂的数据采集环境	简洁清晰,三个部分的属性都定义好了,只需选择就好,而且可以自行开发插件
历史背景	最初设计的目的是为了把数据传入 HDFS 中,侧重传输(多路由),重稳定性	侧重对数据的预处理,因为日志的字段需要大量的预处理,为解析做铺垫
对比	像是散装的台式机,使用较麻烦,工具繁多,需要根据业务选择	更像是组装好的台式机,使用简单方便,搭配 ELK 更高效
优点	对 Java 支持好	对文件读取方便,轻量级,支持 ELK
缺点	对于文件读取不方便,重量级	性能消耗大,不支持缓存,只能用 Redis 和 Kafka 代替缓存

表 4.2 LogStash 和 Flume 的优缺点

例如,如果医院原应用是 Java 开发的,可以选择 Flume,简单配置即可。如果为了统一收集工具,并且医院服务器资源占用不多的情况下,可以考虑 LogStash。

此方案需对医院服务器部署 LogStash/Flume/TdAgent 日志采集工具,对医院的日志格式进行规范化,并且解析不同医院的日志。

(3) Webservice 接口

如觉得写日志压力过大、编写 KafkaProducer 成本过高,由服务商提供 Webservice 接口。这个可以根据定义的 webservice 接口规范,生成客户端完成数据接收转发给 kafka。其优点为:一是不需要写日志;二是不需要额外在客户端部署其他插件或者采集器;三是医院对数据自己可控;四是相比于 Kafka, Webservice 的学习成本更低。缺点为:一是需要医院技术部门或者开发部门编写 KafkaProducer 代码,与业务高度耦合;二是需要对原应用进行小量修改。

此方案准备工作需由开发部门编写 Webservice 接口,底层解析

XML 并传输数据给 Kafka；提供前置机给医院，然后前置机上部署接口客户端。

（4）自定义数据采集

如觉得写日志压力过大，或者为了数据更加精准发送到 Kafka，医院技术部门根据自己业务场景来自定义 KafkaProducer，来发送数据到 Kafka 集群。其优点为：一是不需要写日志；二是不需要额外在客户端部署其他插件或者采集器；三是医院对数据自己可控。缺点为：一是需要医院技术部门或者服务商编写 KafkaProducer 代码，与业务高度耦合；二是需要对原应用进行小量修改。

此方案准备工作：不需要部署前置机；需要医院技术部门或者开发部门编写 KafkaProducer 代码，与业务高度耦合。

（5）数据库离线拉取

如医院几乎不想做任何改造，直接提供数据库地址，可以利用 Hadoop 生态圈的 Sqoop 离线主动拉取数据，前提得知道医院所有表结构，并且每张表能有时间字段和逻辑删除（CreateTime、LastModifyTime、DeleteTime、Flag）等。如果医院原有表就提供时间字段并且表设计和卫计部门需要表结构一致，不要做任何改造。

此方案实施准备工作需要医院提供数据库外网访问地址；在 EDH 集群中安装 Sqoop 组件。

如果数据是第一次导入，可以用 EDH-Sqoop 组件进行关系型数据库到 EDH-HBase 数据库的迁移。

如果不是第一次导入，即增量数据，那么这里又分为两种情况：实时和准实时。实时的数据导入，EDH 提供接口，直接采用推送数据的方式推送过来，即可保证数据的实时性；非实时，如一个小时、

一天数据同步。提供查询接口，EData 主动去获取。该方式需要保证这段时间内，所有数据的操作。例如以一个小时为数据同步时间段，这一个小时内，所有对于原数据库的 CUD 操作都需要记录下来，并提供对应规范。

4.3.3 采集导入内容

(1) 社区卫生服务信息系统数据。已经建档的社区卫生服务个人基本信息：居民个人基本信息包括个人身份信息、地址和联系方式、生物学标识、以及区县（市）建立个人健康档案的建档信息和公共卫生服务信息。

(2) 医疗机构业务运营数据。医疗机构业务数据主要包括门急诊业务信息、住院类业务信息、手术业务信息以及业务数据统计报告信息等。

(3) 临床诊疗数据。包括居民基本信息、门诊就诊信息、住院就诊信息、实验室检验报告、医学影像检查报告、基础字典数据等。

(4) 公卫信息数据。包括高血压、糖尿病报卡、随访等明细信息。

(5) 妇幼保健信息数据。包括妇幼专项个人基本信息、孕妇初次产检、孕妇产时情况、孕妇产后访视、新生儿访视表、孕妇一般情况表信息等。

(6) 体检信息数据。包括体检信息和体检明细信息。

针对采集到的健康档案数据，包括各辖区内社区、二级医院的数据和三级医院的数据，在居民身份唯一识别的基础上，实现区域健康档案数据整合，建立起统一的核心数据库框架，并完成在现有条件下的数据建模、入库，同时通过数据交换平台来实现数据共享、管理、维护、使用的机制。

截止 2018 年 6 月 12 日,健康大数据平台共采集和导入健康档案数据量如下:(表 4.3)

表名	数据量(条)
基本信息	129,408,113
挂号信息	331,250,909
门诊工作日志	300,098,542
门诊诊断信息	387,914,092
门诊卡	48,090,387
门诊处方信息	304,995,898
门诊处方明细	901,096,821
门诊收费信息	622,792,067
门诊费用明细	1,836,960,784
住院登记	4,916,012
住院医嘱	478,186,471
住院费用	1,130,105,757
出院小结	4,074,943
住院诊断	9,180,783
实验室检验报告信息	137,002,535
检验报告单明细	1,718,856,215
医学影像检查报告	51,677,379
体检信息	7,130,886
体检明细信息	552,973,245
健康档案个人信息	10,676,824
高血压随访记录	28,404,358
高血压报告卡	1,448,344
高血压随访记录明细表	29,918,670
糖尿病病例随访记录表	6,822,165
糖尿病病例卡	436,022
糖尿病病例随访记录明细表	7,039,709
妇幼专项个人基本信息	2,419,703
孕妇初次产检	738,501
孕妇产时情况	575,114
孕妇产后访视	1,352,437
新生儿访视表	828,372
孕妇一般情况表	694,367
肿瘤报告卡	46,100
住院病案首页	2,742,678

4.4 电子健康档案数据的交换

在健康大数据平台上搭建大数据交换平台，遵循标准基于先进的企业服务总线 ESB 技术，遵循先进技术标准和规范，为跨地域、跨部门、跨平台不同应用系统、不同数据库之间的互连互通提供包含提取、转换、传输等操作的数据交换服务，实现扩展性良好的“松耦合”结构的应用和数据集成；同时数据交换平台能够通过分布式部署和集中式管理架构，可以有效解决各节点之间数据的及时、高效地上传下达，在安全、方便、快捷、顺畅的进行信息交换的同时精准的保证数据的一致性和准确性，实现数据的一次采集、多系统共享；要求数据交换平台服务配置可视化，快速实现不同业务系统、不同数据库之间基于不同传输协议的数据交换与信息共享，为各种应用和决策支持提供良好的数据环境。

建设交换平台的总体目标是利用大数据集群特性，大量数据导入和导出不影响运行在大数据平台上的业务系统，动态配置数据接口实现真正的活数据，实现交换平台集群负载，为各业务系统数据交换提供实时有效的保障。总体需求为：一是提供交换规范，包括各业务的技术规范以及交换权限认证；二是实现高可用性，包括交换服务、接口实现负载均衡、失败转移等集群技术保障系统服务的高可用性。

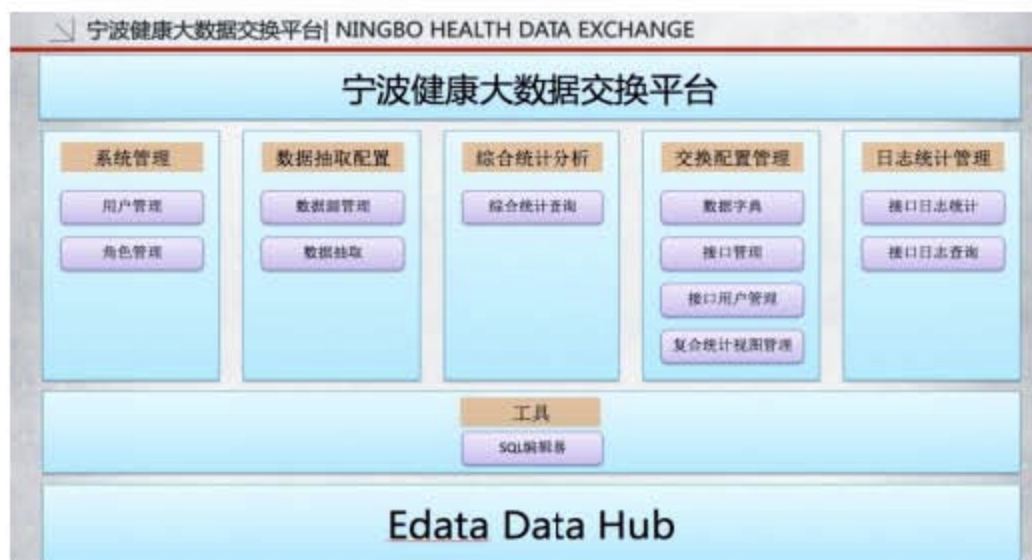


图 4.9 健康大数据交换平台框架

4.4.1 功能描述

(1) 系统管理

用户管理：实现系统用户基本信息进行增加、删除、修改、查询等功能。

角色管理：维护系统角色与用户之间的关系，还需管理其对整个系统的权限，并实现角色的增加、删除、修改、查询等基础管理功能。

(2) 数据抽取配置

满足主动推送数据到其他平台，需要开发数据抽取功能。

数据源管理：配置目的库的 JDBC 信息，含数据库类型、驱动包名、JDBC-URL、数据库名称、用户名、密码，对这些信息进行增加、删除、修改、查询等功能。

数据抽取：能配置到目的库需要抽取的表和字段，能进行初始化数据导出和定时增量数据导出。

(3) 综合统计查询

该功能不需要书写 SQL，直接能提供对单表某些字段的相等、包含、区间查询。

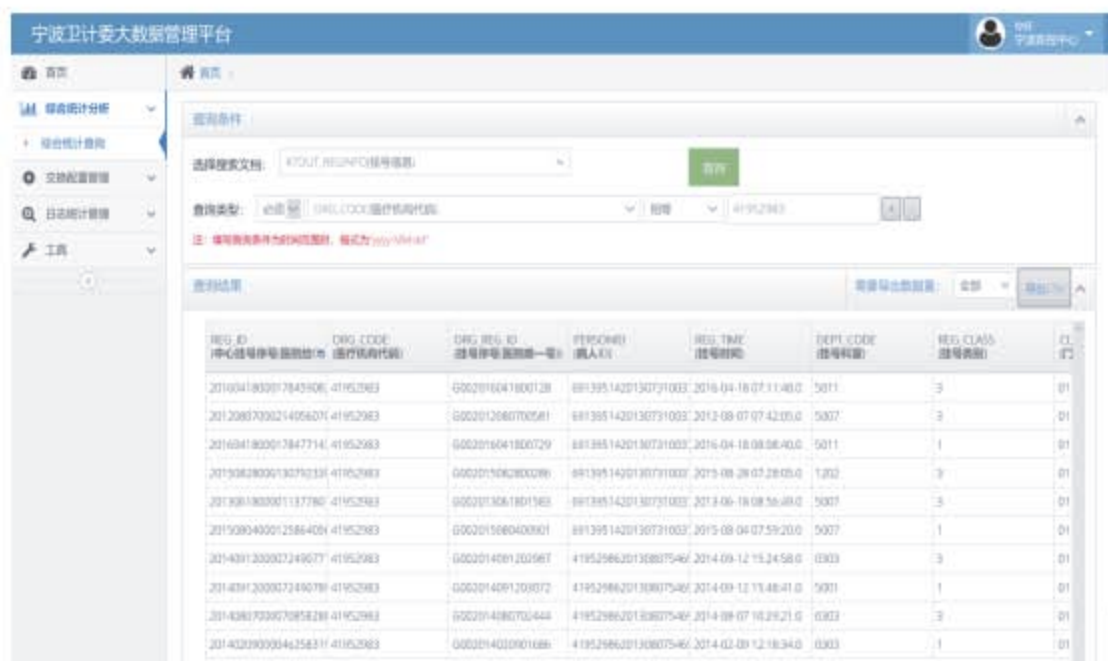


图 4.10 综合统计查询模块

(4) 交换配置管理

数据字典：实现对所有表、字段的增加、删除、修改、查询等操作。

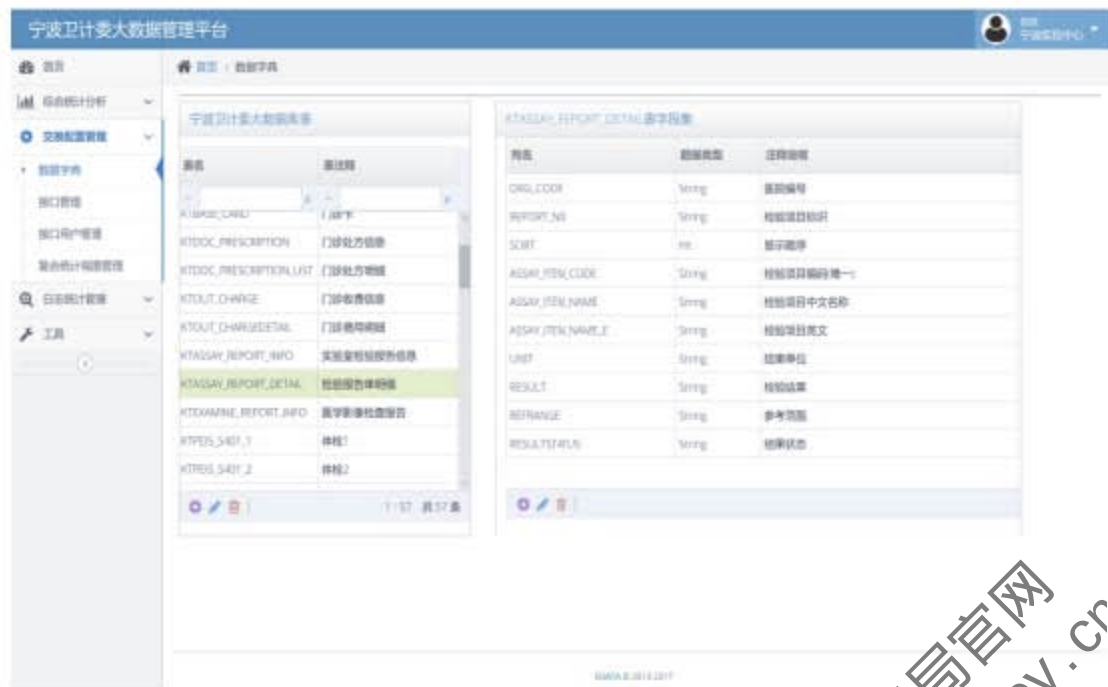
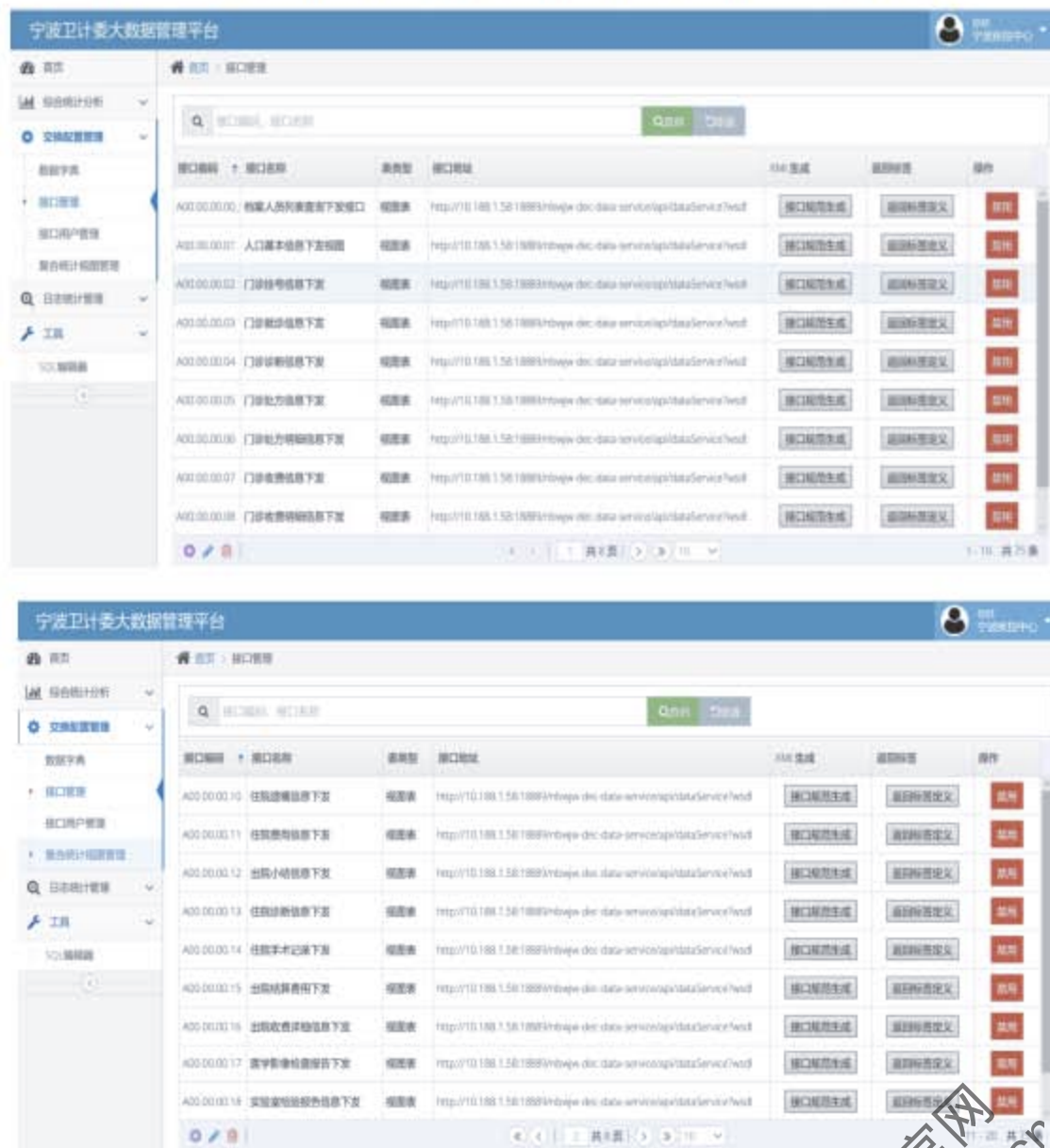


图 4.11 数据字典管理模块

接口管理：提供统一的 API 接口，实现对各个接口的用户，动态

的配置对应的表接口，接口是以本平台的视图进行直接操作。接口以WebService来实现，数据格式以XML作为标准。对以上功能实现增加、修改、查询管理。对接口管理可以实现以上功能的增加、删除、修改、查询等操作。

以下是各类健康档案数据接口管理配置界面：（图 4.12）



国家档案局官网
www.saac.gov.cn

The screenshot displays the "宁波卫计委大数据管理平台" (Ningbo Health Commission Big Data Management Platform). The left sidebar contains navigation menus such as "首页" (Home), "数据统计分析" (Data Statistical Analysis), "交换配置管理" (Exchange Configuration Management), "接口字典" (Interface Dictionary), "接口管理" (Interface Management), "接口用户管理" (Interface User Management), "接口统计和管理" (Interface Statistics and Management), "日志统计管理" (Log Statistics Management), and "工具" (Tools). The main area shows the "接口管理" (Interface Management) page, which includes a search bar and a table listing various interfaces.

接口名称	接口名称	数据类型	接口地址	API生成	返回预览	操作
B03-1.2.00.00	健康档案信息 基础信息下发	数据表	http://10.188.1.58/188HISmwy-dec-data-service/api/dataService.html	接口规范生成	返回预览定义	删除
B03-1.3.00.00	健康档案信息 联系地址下发	数据表	http://10.188.1.58/188HISmwy-dec-data-service/api/dataService.html	接口规范生成	返回预览定义	删除
B03-1.4.00.00	健康档案信息 就医方式下发	数据表	http://10.188.1.58/188HISmwy-dec-data-service/api/dataService.html	接口规范生成	返回预览定义	删除
B03-1.5.00.00	行为习惯下发	数据表	http://10.188.1.58/188HISmwy-dec-data-service/api/dataService.html	接口规范生成	返回预览定义	删除
B03-2.1.00.00	高血压随访信息推送下发	数据表	http://10.188.1.58/188HISmwy-dec-data-service/api/dataService.html	接口规范生成	返回预览定义	删除
B03-2.1.00.00	文惠科信息下发	数据表	http://10.188.1.58/188HISmwy-dec-data-service/api/dataService.html	接口规范生成	返回预览定义	删除
B03-2.1.00.00	高血压随访记录下发	数据表	http://10.188.1.58/188HISmwy-dec-data-service/api/dataService.html	接口规范生成	返回预览定义	删除
B03-2.2.00.00	糖尿病随访记录下发	数据表	http://10.188.1.58/188HISmwy-dec-data-service/api/dataService.html	接口规范生成	返回预览定义	删除
B03-2.3.00.00	慢阻肺随访信息下发	数据表	http://10.188.1.58/188HISmwy-dec-data-service/api/dataService.html	接口规范生成	返回预览定义	删除

国家档案局官网
www.saac.gov.cn

宁波卫计委大数据管理平台

首页 > 接口管理

接口名称: 接口名称

Q 查询 清除

接口名称	接口名称	数据类型	接口地址	API 生成	返回数据	操作
C01.02.02.00	挂号记录查询接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C01.02.01.00	查询患者历史就诊信息 接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C01.02.02.00	就诊信息查询接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C01.04.01.00	查询患者处方详细数据接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C01.04.02.00	查询患者处方处方信息接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C01.04.03.00	查询地方门诊信息接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C01.04.04.00	查询患者的处方信息接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C01.04.05.00	查询患者处方信息接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C02.01.01.00	查询患者就诊记录接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用

1/1 共 75 条

宁波卫计委大数据管理平台

首页 > 接口管理

接口名称: 接口名称

Q 查询 清除

接口名称	接口名称	数据类型	接口地址	API 生成	返回数据	操作
C02.04.01.00	查询患者出外小数据接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C02.05.01.00	查询患者健康档案信息接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C02.07.01.00	查询患者出外小数据接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C03.02.01.00	查询实验室检验信息接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
C03.02.02.00	查询实验室检验信息接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
902.00.00.00	血站住院病人登记信息接口	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
Z001_001	浙江共享 出生医学证明	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
Z001_002	浙江共享 门诊信息	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
Z001_003	浙江共享 体检信息	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用

1/1 共 75 条

宁波卫计委大数据管理平台

首页 > 接口管理

接口名称: 接口名称

Q 查询 清除

接口名称	接口名称	数据类型	接口地址	API 生成	返回数据	操作
Z001_003	浙江共享 门诊信息	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
Z001_004	浙江共享 出外小结	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
Z001_005	浙江共享 检验信息	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
Z001_006	浙江共享 检验信息	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用
Z001_007	浙江共享 检验信息	数据表	http://10.188.1.58:18818/ncwepi-dec-data-service/api/dataService.html	接口规范生成	返回数据定义	禁用

1/1 共 75 条

复合统计视图管理：用数据字典中的表能动态创建 SQL 视图，并能实现接口之后的模板参数，能校验生成之后的视图，并对该功能实现增加、删除、修改、查询等操作。

(5) 日志统计管理

接口日志统计：能对人、或者某个接口进行操作量的一个统计，以图表形式展现。

接口日志查询：可以根据用户、接口编码、业务类型、交换状态、等多条件查询所有历史的交换记录详情。

接口编码	用户名	接口类型	接口名称	请求时间	请求内容	响应内容
HIS00100101	root	HIS	HIS_MESSAGE	2018-01-10 10:00:46.215	read	请求详情 响应详情
HIS00100101	root	HIS	HIS_MESSAGE	2018-01-10 10:00:46.400	read	请求详情 响应详情
HIS00100101	root	HIS	HIS_MESSAGE	2018-01-10 10:00:46.175	read	请求详情 响应详情
HIS00100101	pharmex	HIS	HIS_MESSAGE	2018-01-10 10:00:46.400	read	请求详情 响应详情
HIS00100101	pharmex	HIS	HIS_MESSAGE	2018-01-10 10:00:46.175	read	请求详情 响应详情
HIS00100101	pharmex	HIS	HIS_MESSAGE	2018-01-10 10:00:46.215	read	请求详情 响应详情
HIS00100101	pharmex	HIS	HIS_MESSAGE	2018-01-10 10:00:46.400	read	请求详情 响应详情
HIS00100101	pharmex	HIS	HIS_MESSAGE	2018-01-10 10:00:46.175	read	请求详情 响应详情
HIS00100101	pharmex	HIS	HIS_MESSAGE	2018-01-10 10:00:46.215	read	请求详情 响应详情
HIS00100101	pharmex	HIS	HIS_MESSAGE	2018-01-10 10:00:46.400	read	请求详情 响应详情

图 4.13 日志统计管理模块

4.4.2 系统体系结构

(1) 技术架构

采用 B/S 架构进行建设，采用目前经典主流 MVC 开发模式，如图 4.12 所示。

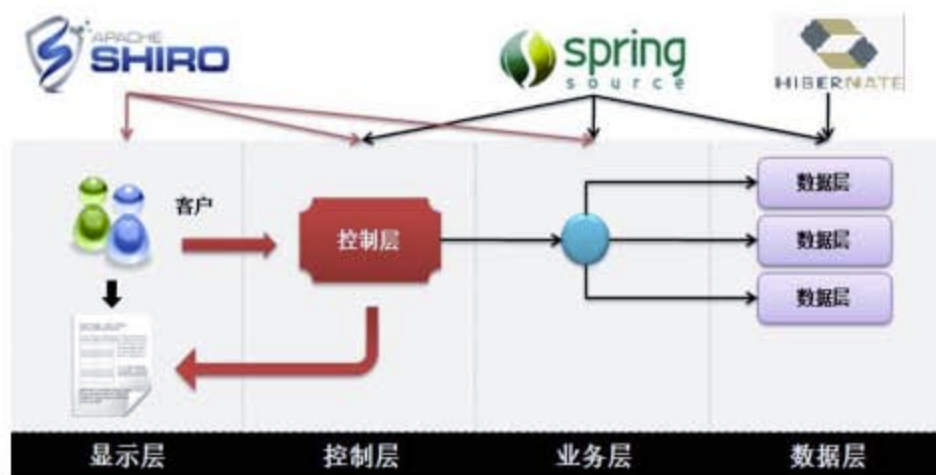
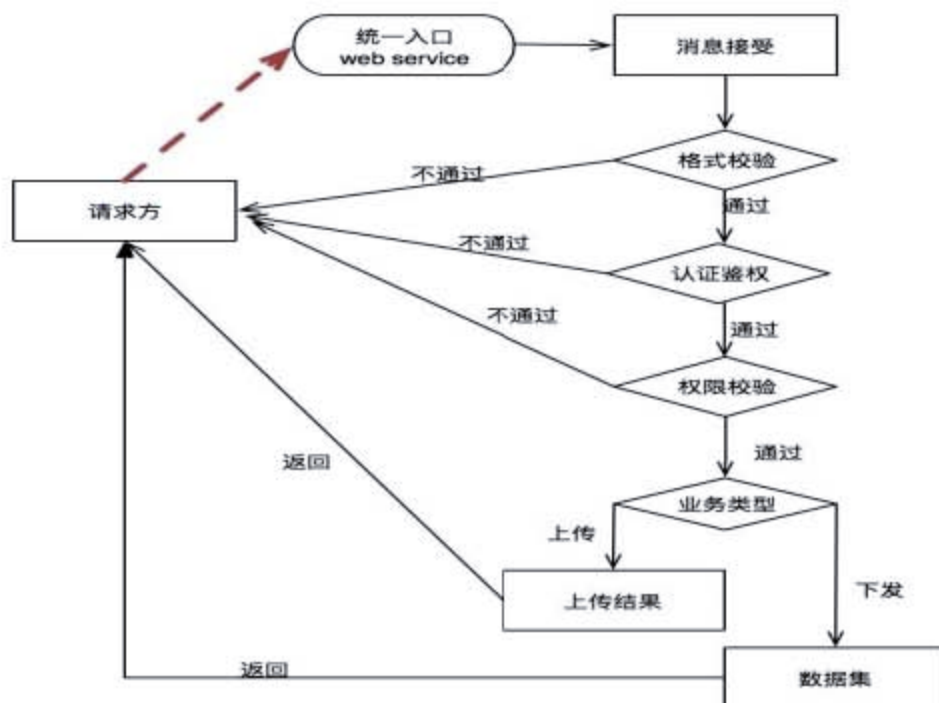


图 4.14 框架设计

基于 J2EE 技术规范，按照三层架构界面展现、业务逻辑、数据访问，利用 JAVA 技术实现基于 B/S 的应用开发。使得系统能够在 Windows、Linux、Solaris 系统上都能运行。设计良好的 web 应用，可以被按职责分为四层。这些层次是表现层、持久层、业务层、和领域模型层。通过分层可以实现层与层之间松散耦合，增加代码重用率。各层分工明确，利于团队的明确分工，提高开发效率。

（2）交换技术架构

用户通过统一的接口入口 Webservice，接口接受消息，经过格式校验，认证鉴权（用户名、密码），权限校验（校验对应的访问数据的权限），业务类型（上传，下发），上传把上传结果返回（是否成功），下发把对应的条件查询数据结果集返回。权限校验流程中，任何一个不通过即返回。交换实现场景是直连目标服务实现数据交换。如图 4.13 显示：



4.15 数据交换业务流程图

4.5 居民电子健康档案的展现

为实现电子健康档案各类数据有机的统一，并能有效得到调阅展现，必须建立健康档案浏览器。基于健康大数据平台构建，实现对居民的健康档案基本信息、医疗服务、公卫专项、体检专项、妇幼专区等档案信息进行分类展现。居民电子健康档案展现功能基于 B/S 架构实现，采用 MVC 开发模式，使用 J2EE 进行开发。健康档案数据存放在大数据平台中，利用其高可靠、高性能、可伸缩的方式进行数据处理。平台能实现扩展性良好的“松耦合”结构的应用和数据集成。居民标准化健康档案将通过一定的逻辑性、层次性及时序性来全面反映所需关注的健康相关信息。



图 4.16 电子健康档案基本框架

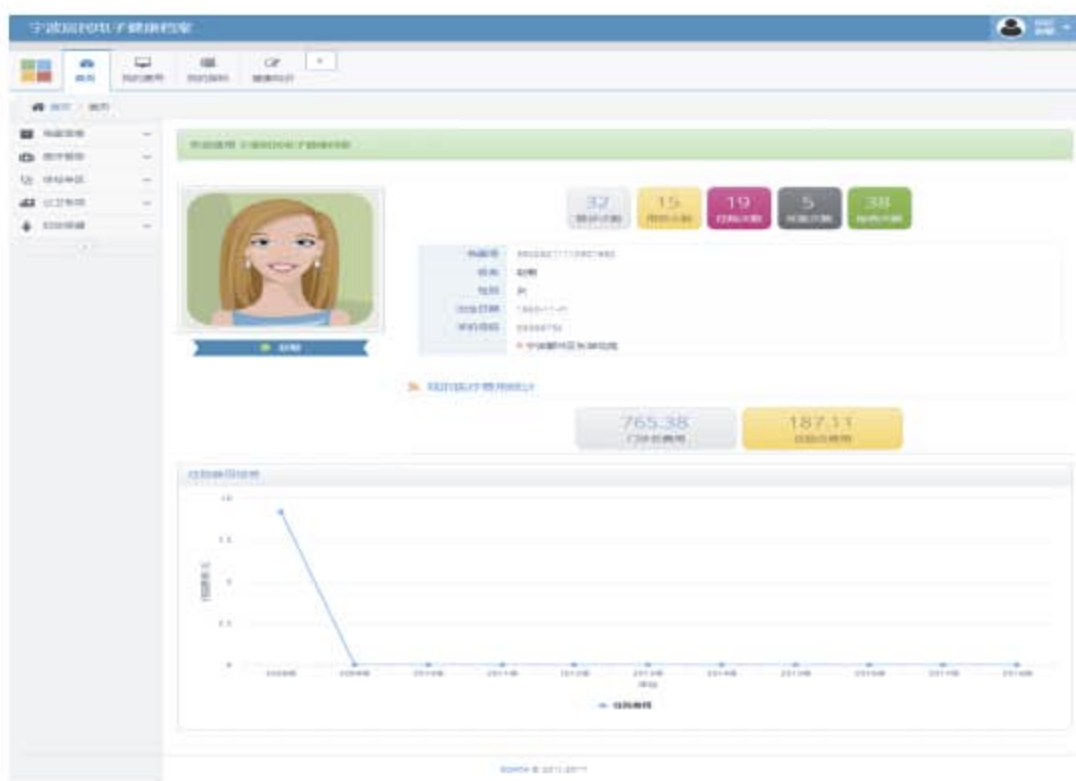


图 4.17 居民电子健康档案首页展示

国家档案局官网
www.saac.gov.cn

国家档案局官网
www.saac.gov.cn

居民健康档案封面

健康档案编号: 【330200 02-22 162】

居民健康档案

姓 名:	朱山
现 住 址:	浙江省宁波市鄞州区塘 村委会
户籍地址:	浙江省宁波市鄞州区塘 村委会
联系电话:	88000168
乡镇(街道)名称:	塘 镇
村(居)委会名称:	塘 村委会
建档单位:	鄞州区 社区卫生服务中心
建 档 人:	system
责任医生:	国
建档日期:	2011-15

图 4.19 居民电子健康档案电子书封面

4.5.1 功能模块

(1) 档案信息

对于个人的基本信息、档案信息、体质信息等进行维护和展现；
对于个人的行为习惯进行展现，例如：吸烟、饮酒、饮食、身体锻炼情况等；

对于个人的健康信息进行展现，例如疾病史、过敏史、预防接种史、用血史等；

对于个人的家庭档案、家庭成员等信息进行展现；

对于不同年龄段建立不同的档案专项信息，例如：儿童、妇女、老年、残疾等档案信息展现。

（2）医疗服务

对于个人所有门诊就诊信息、诊断结果展现；

对于个人所有住院信息、入院诊断结果展现；

对于个人所有化验检查的项目和时间等指标进行展现；

对于个人所有化验检查的项目和时间等指标进行展现。

（3）公卫专项

对于个人不同的公卫专项进行展现，例如不同病种专项的高血压、糖尿病、心脑血管等展现；

对于不同病种的基本信息、诊断信息、最近检查信息、最近用药信息、随访卡信息进行展现；

对于近期常规检查的血压曲线图进行展现。

（4）体检专项

对于所有普通体检进行展现，并能查看详细的体检报告，有其报告封面、医生检查、功能检查、体检总结、报告封底等版面；

对于所有健康体检进行展现，并能查看详细的体检报告，有其基本检查、生活方式、身体检查、辅助检查、体质及问题、医疗信息、健康评估等相关信息。

（5）妇幼专项

对于妇女建立妇女保健专项，中间应含孕卡建册、初孕产检、产前检查、产前筛选、高危情况、产时情况、产妇访视、新生儿访视、42天检查等专项信息展现；

对于儿童建立儿童保健专区，中间应含儿童保健卡、新生儿访视、保健体检记录、周岁健康小结、营养评价等专项信息展现。

4.5.2 设计方案

4.5.2.1 分析原表结构

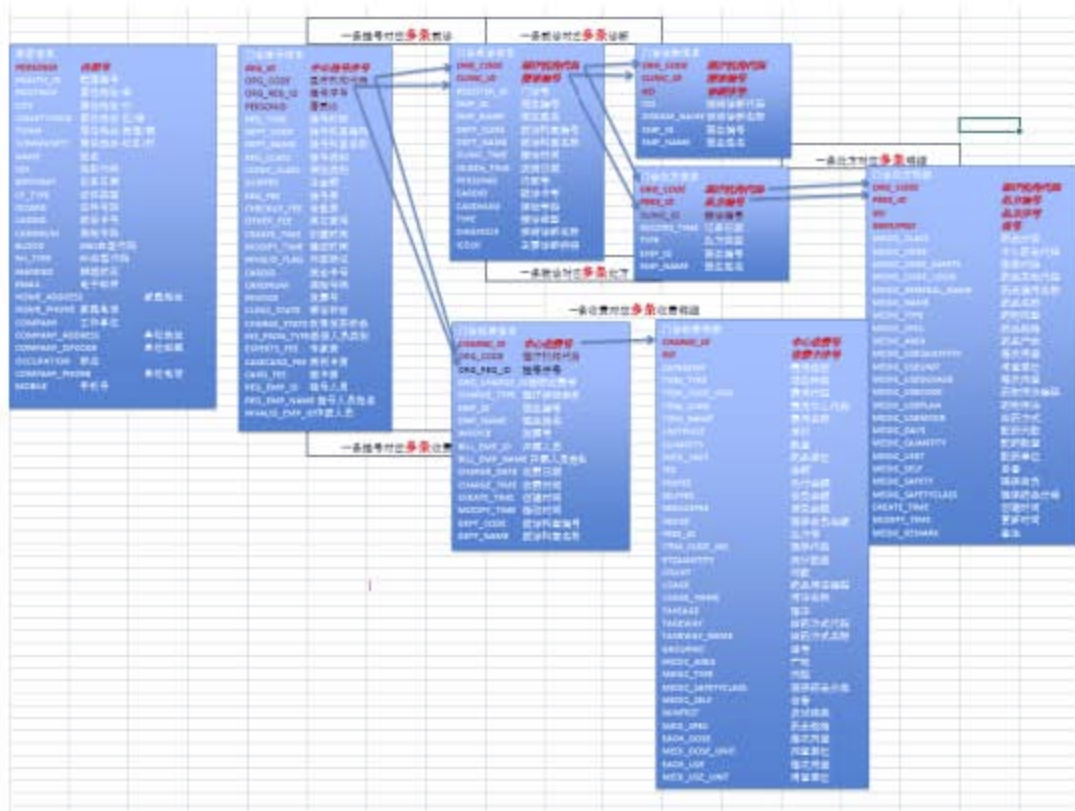


图 4.20 原表结构

从上图看出，原门诊档案信息表结构都是一对多的表设计，从居民健康档案浏览器操作特点上来分析，查询条件都是根据居民身份证和姓名查询个人、公卫、医疗三个维度的信息。而表数据都是以内部号 PersonID 作为底层最核心的 ID。档案信息部分是以档案号 HealthID 来作为最核心的 ID。

4.5.2.2 主键实现方式

电子健康档案数据选用了 EDH 中的 InceptorSQL 与 Hyperbase

组件进行存储和交互式分析。档案数据存储在 Hyperbase 中。根据 Hyperbase 主键即索引,通过主键查询能达到 PB 级别数据毫秒级别响应。

为了数据能负载均衡,避免热点数据,Region 大量拆分等对集群性能产生大量影响的操作,需预先对表进行分区,分区的 splitkey 是通过历史数据总数和预分区数计算出来的。此方法不需要 hash 也能保证数据的负载均衡,没有热点数据。Hyperbase 的 Rowkey 是把各种业务中常用的内部号、档案号作为前缀。因为 EHR 原来的档案数据表都是大部分都是通过内部号、档案号来关联匹配的,作为前缀匹配,不走索引,数据是 100ms 以内响应的,包含原表最大的两个业务,处方明细、费用明细。对于原来应执行的 Hyperbase API 的操作,现在用 InceptorSQL 对 Hyperbase 建立外部表,以 SQL 方式实现分布式数据仓库和面向数据的交互式查询。对不同表有不同 Rowkey 设计,主要是以下三种场景。

(1) 含查询条件的表

对于有查询条件的表,无需额外创建业务库表,“患者基本信息”、“门诊挂号信息”,“门诊就诊信息”直接利用原始库表,无需创建额外的业务库表,直接对其内部号作为主键的前缀即可。

(2) 不含查询条件的表

对于没有查询条件的表,即没有内部号 PersonID 的表,需要对上层关联,清洗无用数据。如“门诊诊断信息”和“门诊就诊信息”关联,在“门诊诊断信息”表中加入“门诊就诊信息”的内部号,并且作为主键前缀。

(3) 宽表紧耦合的表

对于在一对多的表中，一方仅对多方提供关联，一方字段少，可以完全将一方表的字段冗余的放到多表中。例如“门诊处方信息”和“门诊处方明细”，“门诊处方信息”仅多出“接诊号”、“记录日期”、“处方类型”、“医生编号”、“医生姓名”、“处方分类”6个字段，完全可以全部放到“门诊处方明细”表中。并且对最新的“门诊处方明细”和“门诊就诊日志”关联，加入内部号 PersonID，作为“门诊处方明细”的主键前缀。

4.5.2.3 数据存储和安全性

对于数据存储空间，原历史档案数据有 4.5T 的纯文本数据，如果不进行任何优化导入 Hyperbase 平台，数据将以 (7~10) 倍*3 副本的倍数增长，这是因为 Hyperbase 有其 Rowkey 长度，列族、列限定符、时间戳、版本等信息以及乘 3 个副本的原因。对于该情况，我们选用 Snappy 压缩，压缩率达到 90%，再简化列族命名、列限定符命名。即最终效果就是 4.5T 数据只要 13.5T 的存储空间，含副本。

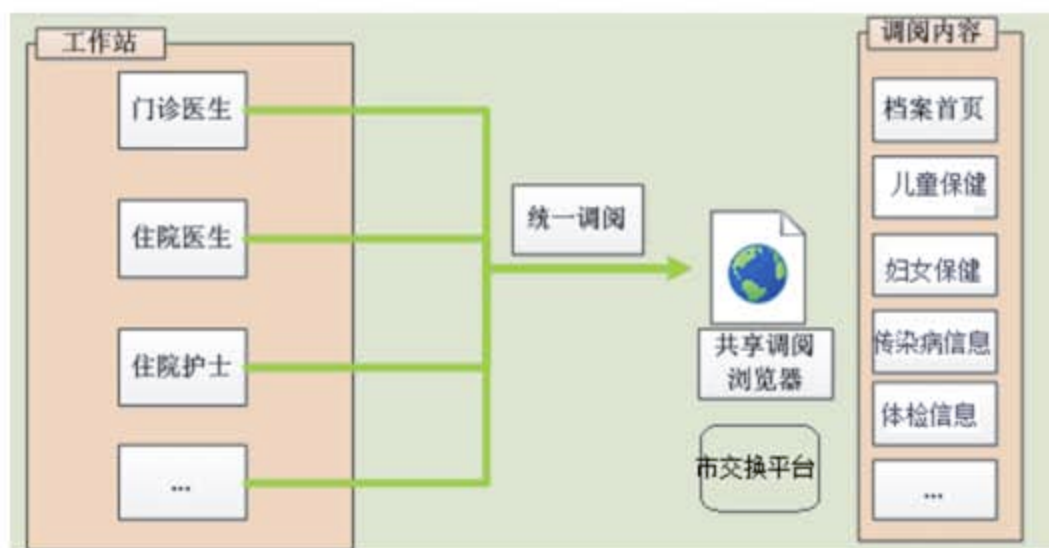
数据安全方面，EDH 提供 Guardian 组件，对 InceptorSQL、Hyperbase 进行 Kerberos 加密访问，其中 Hyperbase 提供全局、Namespace, Table, Column Family 和 Column Qualifier 级别的“RWCXA(Read, Write, Create, Exec, Admin)”权限控制；Inceptor 提供基于 SQL 的全局、数据库、表和行级别的 SELECT/CREATE/DELETE/UPDATE/INSERT 权限控制。

4.6 居民电子健康档案的共享方式

全市电子健康档案是按照全市统一标准制定，是建立在大数据 Hadoop 技术来保证档案的安全性、可靠性、高效性，通过健康大数据底层架构、交换平台和电子健康档案浏览器的建设，居民电子健康

档案的共享开放应用在技术上主要通过 2 种方式实现。一种是运用 WebAPI 技术基于健康档案浏览器的调阅，这种主要应用在医生工作站，通过集成视图的方式和医院医生工作站深度融合，实现就诊环节的健康档案信息的共享；第二种是以 WebService 接口提供统一服务实现健康信息调取。APP、微信、第三方需要信息共享部门通过调用统一接口服务，通过各自的应用展示出健康信息，达到信息共享。

健康档案浏览器是在医疗机构内部提供跨机构的健康信息整合视图。健康档案浏览器是一种只能查阅，数据不能落地修改的调阅模式，是基于 Web 访问的应用程序，使用方便，集成快，从数据安全角度上，是一种相对理想的信息共享方式。在中心搭建电子健康档案共享服务，建立相应的共享调阅服务站点，供市属医院和区县（市）卫生机构共享调阅。统一浏览器调阅提供了统一的访问方式，便于统一管理和维护。



4.21 健康信息调阅功能架构图

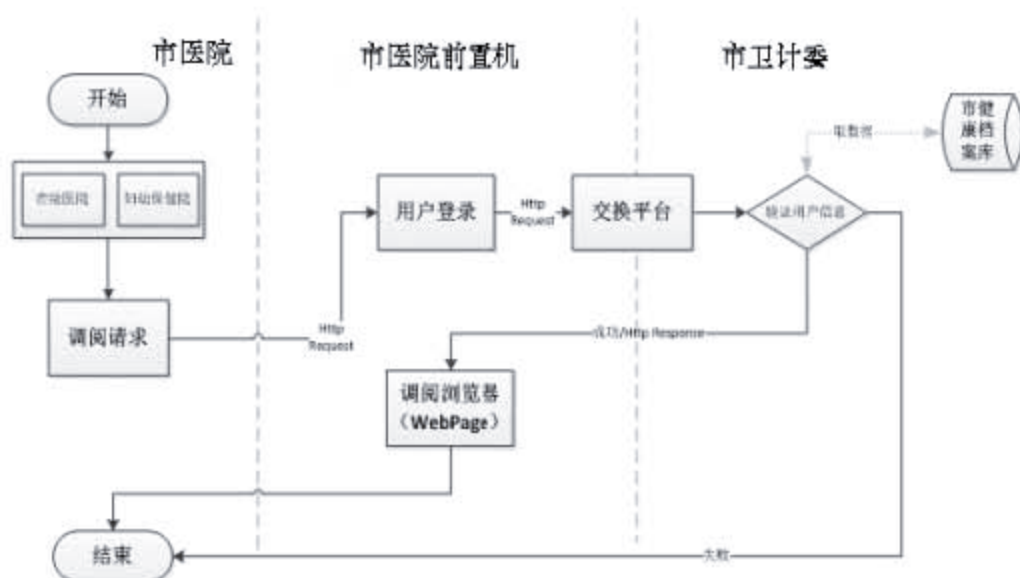


图 4.22 健康信息调阅功能实现图

基于 Web 站点提供的 Web API 接口,利用 HTTP GET/POST 请求发送功能实现基于健康档案浏览器的 Web 服务。访问记录索引请求格式如下:

https://ip:port/PersonInfo.aspx?visittype=< 模块>&orgcode=<机构号>&cid=<主键号>&empcode=< 医生工号 >

按照不同的信息类型分类,采用不同的接口入参: (表 4.4)

项目	访问方式
访问就诊记录	PersonInfo.aspx?visittype=clinic&orgcode=&password=&cid=&headtype=&empcode=
访问住院记录	PersonInfo.aspx?visittype=inhos&orgcode=&password=&cid=&headtype=&empcode=
访问化验记录	PersonInfo.aspx?visittype= assay & orgcode=&password=&cid=&headtype=&empcode=
访问检查记录	PersonInfo.aspx?visittype=examine&orgcode=&password=&cid=&headtype=&empcode=
访问体检记录	PersonInfo.aspx?visittype= clinic& orgcode=&password=&cid=&headtype=&empcode=
访问个人首页	PersonInfo.aspx?visittype= home& idcard=& cardid=&headtype=&empcode=
访问妇女保健	PersonInfo.aspx?visittype= womenhealth& idcard=& headtype=&empcode=
访问幼儿保健	PersonInfo.aspx?visittype= childrenhealth& idcard=& headtype=&empcode=
访问化验信息	PersonInfo.aspx?visittype=climict& idcard=& headtype=&sdate=&edate=&empcode=
访问就诊信息	PersonInfo.aspx?visittype=assayt& idcard=& headtype=&sdate=&edate=&empcode=
访问住院信息	PersonInfo.aspx?visittype=inhos& idcard=& headtype=&sdate=&edate=&empcode=
访问检查信息	PersonInfo.aspx?visittype=examinet& idcard=& headtype=&sdate=&edate=&empcode=
访问体检信息	PersonInfo.aspx?visittype=checkt& idcard=& headtype=&sdate=&edate=&empcode=

在下表中对访问方式的接口参数 visittype、idcard 等进行了详细的解释,如访问类型 clinic 代表就诊记录,check 代表体检记录。

参数 idcard 代表患者的身份证号等,可调阅一条健康信息,也可调阅一批健康信息。

接口参数说明: (表 4.5)

参数	描述	是否允许为空	允许值及综合考虑
PersonInfo.aspx	页面访问地址文件	否	/
visittype	访问类型	否	clinic, 就诊记录 in hos, 住院记录 assay, 化验记录 examine, 检查记录 check, 体检记录 home, 个人首页 womenhealth, 妇女保健 childrenhealth, 幼儿保健 climict, 就诊信息 in host, 住院信息 assayt, 化验信息 examinet, 检查信息 checkt, 体检信息
cid	就诊记录: 接诊号 住院记录: 住院号 化验记录: 化验报告单号 检查记录: 检查报告单号 体检记录: 体检号	是	
idcard	身份证号	是	/
cardid	就诊卡号	是	社区人员就诊卡号允许为空
patientcode	住院号	是	/
patienttime	住院次数	是	第几次
headtype	显示头底部控制	是	1 默认 2 共享调阅头部
empcode	职工代码	是	规则: 机构号+医院人员标识 例如: 机构号: yyyyyyyy 医院人员标识: xxxxxx 职工代码: yyyyyyyyxxxxxx
sdate	开始时间	是	规则: yyyy-MM-dd
sdate	结束时间	是	规则: yyyy-MM-dd

健康档案信息共享的另一种方式是大数据交换平台的 Web Service 接口, 是直接将 ESB 服务以 Web Service 的方式发布, 这种

方式主要用户实现异构环境间数据的交互。

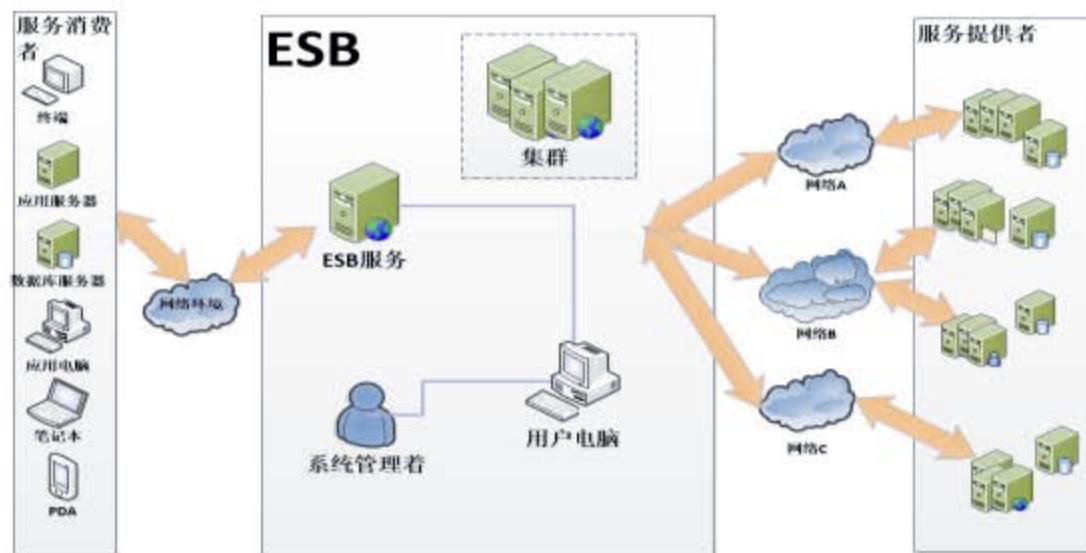


图 4.23 健康档案信息调取功能架构图

健康档案信息调取通过 ESB 总线的功能中转实现。前端服务消费者访问 ESB 总线服务，后端服务提供者通过配置 ESB 服务，完成数据调取功能实现。



图 4.24 健康档案信息调取模式

技术实现方式是通过 ESB 配置管理工具与业务相结合，基于 XML 标准规范进行信息交互。利用 ESB 配管管理工具，实现基于WebService服务数据调取。下表为详细调阅病人挂号信息的 XML 结构：（表 4.6）

<body>
<head>
<userid></userid>

```

    <password></password>
    <trans_no>A00.00.00.02</trans_no>
</head>
<request>
<A00.00.00.02>
    <!--医疗机构代码-->
    <ORG_CODE></ORG_CODE>
    <!--病人 ID-->
    <PERSONID></PERSONID>
    <!--每页大小-->
    <PAGE_SIZE></PAGE_SIZE>
    <!--当前需要返回第几页-->
    <PAGE_INDEX></PAGE_INDEX>
    <!--其他区域的医疗信息标志[Y: 非本区域, N: 本区域]-->
    <FLAG></FLAG>
    <!--获取数据开始时间 格式: yyyy-MM-dd-->
    <BEGINDATE></BEGINDATE>
    <!--获取数据结束时间 格式: yyyy-MM-dd-->
    <ENDDATE></ENDDATE>
</A00.00.00.02>
</request>
</body>

```

前端是调用 Web service 服务。服务地址是: <http://localhost:端口/nbwjw-dec-data-service/api/dataService?wsdl>, 通过统一接口服务调取相应健康信息。

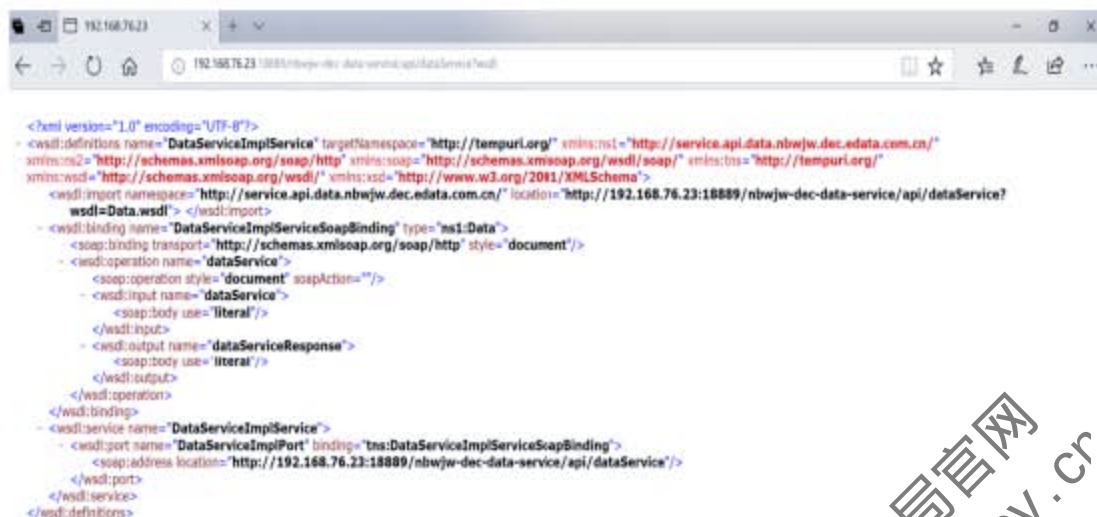


图 4.25 接口服务展示