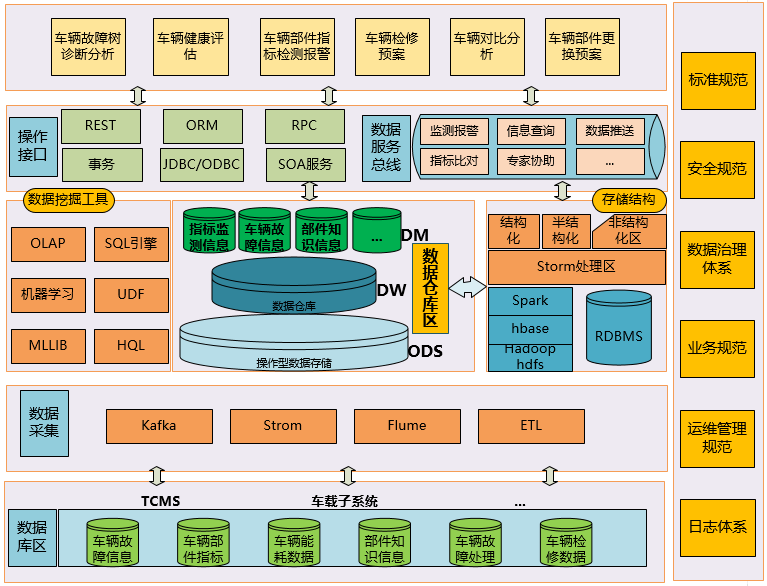
# 数据仓库建设

## 数据仓库总体架构

专家系统接收增购项目车辆TCMS或其他子系统通过车地通信传输的实时或离线数据，经过一系列综合诊断分析，以各种报表图形或信息推送的形式向用户展示分析结果。针对诊断出的车辆故障将给出专家建议处理措施，为车辆的故障根因修复提供必要的支持。

根据专家系统数据仓库建设目标，结合系统数据业务规范，包括数据采集频率、数据采集量等相关因素，设计专家系统数据仓库架构如下：



数据仓库架构从层次结构上分为数据采集、数据存、数据分析、数据服务等几个方面的内容：

数据采集：负责从各业务自系统中汇集信息数据，系统支撑Kafka、Storm、Flume及传统的ETL采集工具。

数据存储：本系统提供Hdfs、Hbase及RDBMS相结合的存储模式，支持海量数据的分布式存储。

数据分析：数据仓库体系支持传统的OLAP分析及基于Spark常规机器学习算法。

数据服务总线：数据系统提供数据服务总线服务，实现对数据资源的统一管理和调度，并对外提供数据服务。

## 数据采集

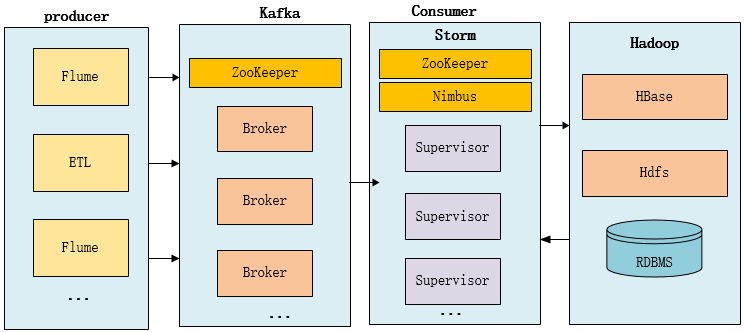
专家系统数据仓库数据采集包括两个部分内容：外部数据汇集、内部各层数据的提取与加载。外部数据汇集是指从TCMS、车载子系统等外部信息系统汇集数据到专家数据仓库的操作型存储层（ODS）；内部各层数据的提取与加载是指数据仓库各存储层间的数据提取、转换与加载。

### 外部数据汇集

专家数据仓库数据源包括列车监控与检测系统（TCMS）、车载子系统等相关子系统，数据采集的内容分为实时数据采集和定时数据采集两大类，实时数据采集主要对于各项检测指标数据；非实时采集包括日检修数据等。

根据项目信息汇集要求，列车指标信息采集具有采集数据量大，采集频率高的特点，考虑到系统后期的扩展，因此在数据数据采集方面，要求采集体系支持高吞吐量、高频率、海量数据采集，同时系统应该灵活可配置，可根据业务的需要进行灵活配置横向扩展。

本方案在数据采集架构采用Flume+Kafka+Storm的组合架构，采用Flume和ETL工具作为Kafka的Producer，采用Storm作为Kafka的Consumer，Storm可实现对海量数据的实时处理，及时对问题指标进行预警。具体采集系统技术结构图如下:



#### 数据汇集架构功能

Flume提供了从console(控制台)、RPC(Thrift-RPC)、text(文件)、tail(UNIX tail)、syslog(syslog日志系统，支持TCP和UDP等2种模式)，exec(命令执行)等数据源上收集数据的能力。Flume的数据接受方，可以是console(控制台)、text(文件)、dfs(HDFS文件)、RPC(Thrift-RPC)和syslogTCP(TCP syslog日志系统)等。在我们系统中由kafka来接收。

Kafka分布式消息队列，支撑系统性能横向扩展，通过增加broker来提高系统的性能。

Storm流处理技术，支撑Supervisor横向扩展以提高系统的扩展性和数据处理的实时性。

#### 采集架构优势

1. **解耦**

在项目中要平衡数据的汇集与数据的处理性能平衡，是极其困难的。消息队列在处理过程中间插入了一个隐含的、基于数据的接口层，两边的处理过程都要实现这一接口。这允许你独立的扩展或修改两边的处理过程，只要确保它们遵守同样的接口约束。

* **冗余**

有些情况下，处理数据的过程会失败。除非数据被持久化，否则将造成丢失。消息队列把数据进行持久化直到它们已经被完全处理，通过这一方式规避了数据丢失风险。在被许多消息队列所采用的“插入-获取-删除”范式中，在把一个消息从队列中删除之前，需要你的处理过程明确的指出该消息已经被处理完毕，确保你的数据被安全的保存直到你使用完毕。

* **扩展性**

因为消息队列解耦了你的处理过程，所以增大消息入队和处理的频率是很容易的；只要另外增加处理过程即可。不需要改变代码、不需要调节参数。扩展就像调大电力按钮一样简单。

* **灵活性 & 峰值处理能力**

在访问量剧增的情况下，应用仍然需要继续发挥作用，但是这样的突发流量并不常见；如果为以能处理这类峰值访问为标准来投入资源随时待命无疑是巨大的浪费。使用消息队列能够使关键组件顶住突发的访问压力，而不会因为突发的超负荷的请求而完全崩溃。

* **可恢复性**

当体系的一部分组件失效，不会影响到整个系统。消息队列降低了进程间的耦合度，所以即使一个处理消息的进程挂掉，加入队列中的消息仍然可以在系统恢复后被处理。而这种允许重试或者延后处理请求的能力通常是造就一个略感不便的用户和一个沮丧透顶的用户之间的区别。

* **送达保证**

消息队列提供的冗余机制保证了消息能被实际的处理，只要一个进程读取了该队列即可。在此基础上，IronMQ提供了一个”只送达一次”保证。无论有多少进程在从队列中领取数据，每一个消息只能被处理一次。这之所以成为可能，是因为获取一个消息只是”预定”了这个消息，暂时把它移出了队列。除非客户端明确的表示已经处理完了这个消息，否则这个消息会被放回队列中去，在一段可配置的时间之后可再次被处理。

* **缓冲**

在任何重要的系统中，都会有需要不同的处理时间的元素。例如,加载一张图片比应用过滤器花费更少的时间。消息队列通过一个缓冲层来帮助任务最高效率的执行—写入队列的处理会尽可能的快速，而不受从队列读的预备处理的约束。该缓冲有助于控制和优化数据流经过系统的速度。

* **异步通信**

很多时候，你不想也不需要立即处理消息。消息队列提供了异步处理机制，允许你把一个消息放入队列，但并不立即处理它。你想向队列中放入多少消息就放多少，然后在你乐意的时候再去处理它们。

### 内部各层数据提取与加载

数据汇集将数据储存于操作型数据存储层（ODS），在数据仓库各层次间数据转换提取加载，采用传统的ETL工具进行采集，数据仓库间的各层次的数据采集的实效性根据具体的数据需求而定，具体ETL建模界面如图：

## 数据加工与处理

对于数据仓库平台，应该建立一套标准化、规范化的数据处理流程，例如：如何采集内部和外部数据、结构化和非结构化数据；如何清洗采集来的脏数据和无效数据；如何对不同来源的数据进行打通；如何对非结构化的数据进行结构化加工；如何在结构化数据的基础上进行商业建模和数据挖掘等等。

大数据管理层在一条数据总线上构建了一条完整的大数据处理流水线。这条流水线从数据的采集、清洗到加工处理，把原始杂乱无章的数据加工成结构化的数据组件，供上层的大数据应用来拼装调用，让企业拥有创造数据资产的能力。



## 存储设计

### 数据量估算

按每列列车平均500毫秒通过车地通信采集监测数据100条，每天运营时间18小时，按每条记录160字节计算(监测数据的数据项相对简单)，初步按照67列列车计算。

单列列车日监测数据=3600\*2\*160\*100\*18/1024/1024/1024≈2G

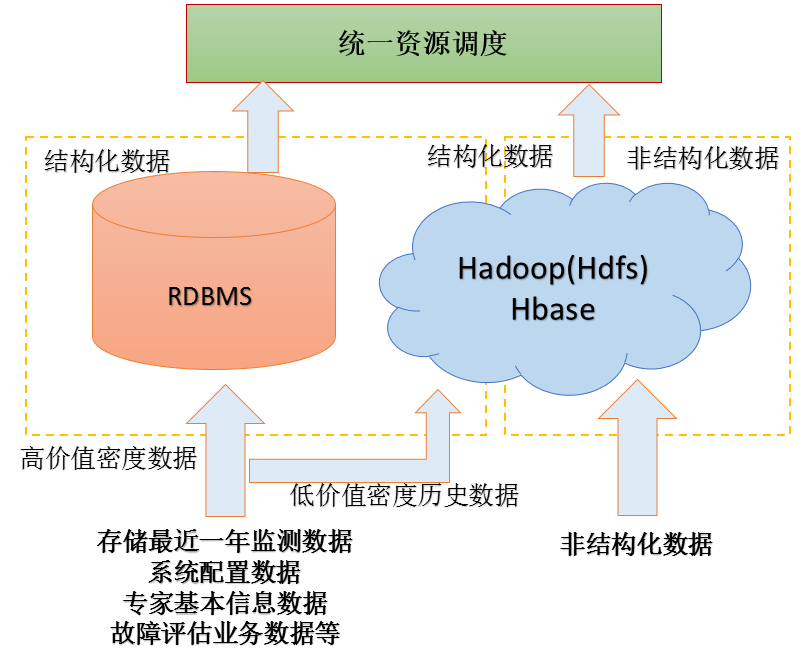
67列列车年数据量=2\*67\*365/1024 ≈ 48T

10年总数据量(乘上增长系数10%)≈530T (含操作系统)

数据规划10年，加上系统用户信息、系统日志信息、专家信息、业务数据及其它不可预测类数据，数据总量预估530T。

### 数据存储

专家系统数据采用混合存储模式进行存储，RDBMS存储专家系统业务基本数据及最近1年的监测数据，10年内历史监测数据采用NoSQL HBase数据库进行存储，以方便查询，HBase基于Hdfs分布式文件系统搭建，具体存储模式如下图。



1. RDBMS数据库，支持专家库的核心业务，存储列车最近1年的监测数据为保证专家系统安全、稳定运行，在数据库系统上支撑各种统计分析及传统的BI业务。考虑到操作系统存储、缓存存储、数据库系统存储、日志存储等因素， RDBMS数据库服务器预计每台60T存储，考虑数据安全及系统稳定因素RDBMS采用双机热备技术互备。
2. 大数据平台规划存储最近10年监测数据，日志文件备份及历史数据采用大数据Hadoop和HBase存储，大数据平台数据采用节点间冗余备份，预设数据2倍冗余存储，

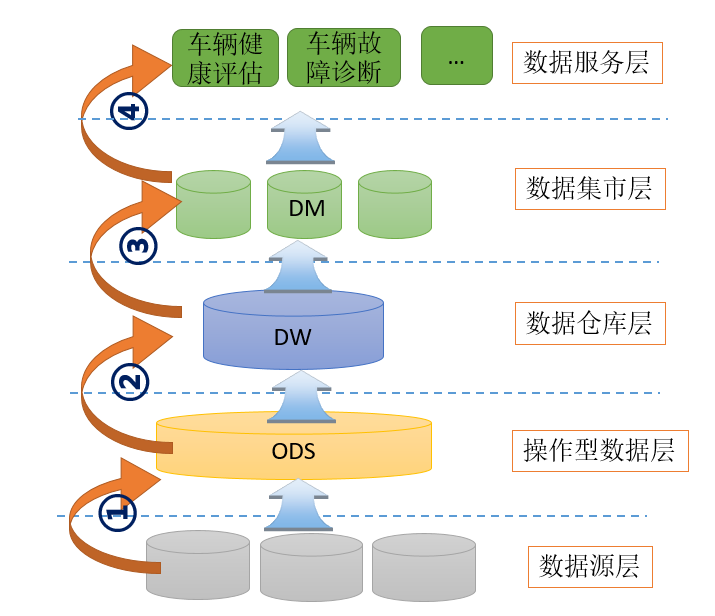
（考虑平台提供的压缩技术，压缩存储可以节省30-55%的空间）。

10年数据量=530T\*1.5≈ 800T (2倍冗余存储)

### 分层存储

专家数据分三个层次进行汇集与存储，分别为ODS层、数据仓库层、主题数据层，各层次数据存储内容如下

* ODS层：数据来源于各生产系统，通过ETL工具对接口文件数据进行编码替换和数据清洗转换，不做关联操作。未来也可用于准实时数据查询。
* 数据仓库层：数据深度汇集层，根据业务有选择的对ODS层的数据进行提取，通过对数据的加工处理，将单一的数据信息转换成体系信息，将点信息数据变成面信息数据。
* 主题数据层：将数据信息体系根据各主题进行提取与转换，主题域内部进行拆分、关联。是对ODS操作型数据按照主题域划分规则进行的拆分及合并。



## 数据分析建模

伴随着大数据时代的悄然来临,数据的价值得到人们的广泛认同,对数据的重视提到了前所未有的高度。数据已经作为企业、事业单位的重要资产被广泛应用于盈利分析与预测、客户关系管理、合规性监管、运营风险管理等业务当中。如何建立大数据分析模型，以提供决策依据是很多用户所迫切解决的问题。

专家数据仓库建立在Hadoop分布式系统之上，提供了多种丰富的算法模型，不同的应用通过借助不同的接口实现数据的多维呈现和结果展示，为用户提供科学的决策支持。

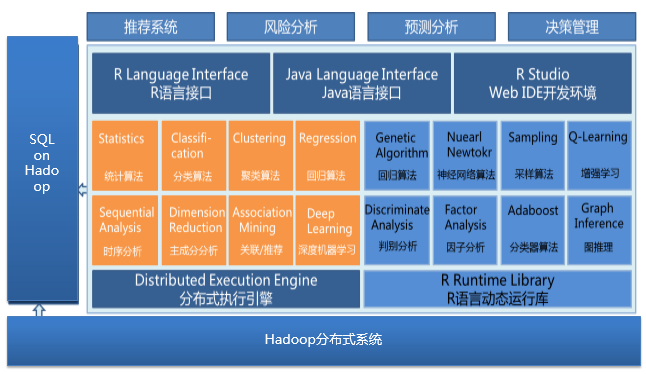


图 10-7 hadoop算法模型图

大数据平台提供数据挖掘模型、分布式计算引擎、高性能机器学习算法库（包含分类 、聚类 、预测、推荐等机器学习算法）、即席查询功能，可以帮助决策者快速建立数据分析模型立方体，便于决策者进行OLAP分析。

**常用算法模型：**

* 分类算法：

分类是找出数据库中的一组数据对象的共同特点并按照分类模式将其划分为不同的类，其目的是通过分类模型，将数据库中的数据项映射到某个给定的类别中。如政务网中将用户在一段时间内的网上办理所遇到的问题划分成不同的类，根据情况向用户推荐关联类的问题解决方案，从而方便用户快速解决网上办事审批中遇到的各类问题。

* 回归算法

回归分析反映了数据库中数据的属性值的特性，通过函数表达数据映射的关系来发现属性值之间的依赖关系。在回归算法中通常将数值结果转化为了0到1之间的概率，数值越大，函数越逼近1，数值越小，函数越逼近0，它可以应用到对数据序列的预测及相关关系的研究中去。如我们根据这个概率可以做垃圾邮件预测，例如概率大于0.5，则这封邮件就是垃圾邮件。

* 聚类算法

聚类类似于分类，但与分类的目的不同，是针对数据的相似性和差异性将一组数据分为几个类别。属于同一类别的数据间的相似性很大，但不同类别之间数据的相似性很小，跨类的数据关联性很低。分类算法中的一个显著特征就是训练数据中包含了标签，训练出的模型可以对其他未知数据预测标签。在聚类的算法中，训练数据都是不含标签的，而算法的目的则是通过训练，推测出这些数据的标签。以二维的数据来说，一个数据就包含两个特征，可通过聚类算法，给他们中不同的种类打上标签，通过聚类算法计算出种群中的距离，根据距离的远近将数据划分为多个族群。

* 关联算法

关联规则是隐藏在数据项之间的关联或相互关系，即可以根据一个数据项的出现推导出其他数据项的出现。关联规则的挖掘过程主要包括两个阶段：第一阶段为从海量原始数据中找出所有的高频项目组;第二极端为从这些高频项目组产生关联规则。

* 推荐算法

推荐算法是目前业界非常火的一种算法，在电商界，如亚马逊，天猫，京东等得到了广泛的运用。推荐算法的主要特征就是可以自动向用户推荐他们最感兴趣的东西，从而增加购买率，提升效益。

* 神经网络模型

神经网络模型，因其自身自行处理、分布存储和高度容错等特性非常适合处理非线性的以及那些以模糊、不完整、不严密的知识或数据为特征的处理问题，它的这一特点十分适合解决数据挖掘的问题。典型的神经网络模型主要分为三大类：第一类是以用于分类预测和模式识别的前馈式神经网络模型;第二类是用于联想记忆和优化算法的反馈式神经网络模型。第三类是用于聚类的自组织映射方法。

* Adaboost算法

其核心思想是针对同一个训练集，训练不同的分类器(弱分类器)，然后把这些弱分类器集合起来，构成一个更强的最终分类器 (强分类器)。其算法本身是通过改变数据分布来实现的，它根据每次训练集之中每个样本的分类是否正确，以及上次的总体分类的准确率，来确定每个样本的权值。将修改过权值的新数据集送给下层分类器进行训练，最后将每次训练得到的分类器最后融合起来，作为最后的决策分类器。

* 深度学习

深度学习算法是对人工神经网络的发展。在计算能力变得日益廉价的今天，深度学习试图建立大得多也复杂得多的神经网络，用来处理存在少量未标识数据的大数据集。

## 数据资源管理

专家系统数据具有数据量大、数据类别多、数据关联关系紧密等特点，随着数据的积累，数据资源的利用价值逐步体现，提高数据的管理，是对数据资源充分利用的前提条件。数据资源管了包括如下几部分内容：数据标准化管理、数据监测管理及元数据管理等。

### 数据标准管理

汇集整理数据资源管理所需的标准规范信息，建立数据标准数据库。利用专家系统数据标准管理系统的接口同步更新标准信息。包括数据元标准以及信息代码标准。

1. 建设数据资源库，实现专家系统发布标准数据元与本地扩展数据元标准的汇集。实现与车辆检修等数据源管理系统接口对接。
2. 建设信息代码资源库，梳理国标、部标和本省定义的标准代码以及各业务信息系统需要使用的其它代码，建立字典代码实体数据库。应具备字典代码定期同步功能。并建设信息代码在线映射维护功能，以便对数据标准化转换提供支持。

### 数据监控管理

大数据运行监控通过对大数据资源库相关服务器、Oracle数据库、分布式存储系统、Hadoop平台等的运行状态、性能指标以及数据更新情况进行持续监控，及时发现存在的问题及隐患，辅助系统管理员及时采取措施，提高大数据资源库的运行可靠性，保障大数据资源库稳定高效运行。发现异常问题时通过短信、邮件等方式通知系统管理员及时处理，实现通过自动、智能、持续的自动监控预警代替人工巡检，降低运维工作量，提高运维效率。通过可视化图表对监控结果进行统计分析直观展现平台运行各类运行指标，辅助管理员从宏观角度掌握平台运行情况。

* 性能指标监控

可以对服务器CPU负载、Oracle数据库连接数、分布式存储IO负载、Hadoop负载等各类性能相关指标进行监控，以便掌握平台负载情况，及时发现性能问题，辅助平台优化。

* 大数据库日志监控

自动采集大数据相关组件运行日志，并根据既定规则进行分析，发现异常及时告警。提供日志查询检索功能，可以按组件类型、时间、关键字等进行过滤。

* 数据量监控

数据量监控通过对数据总量以及增量进行定期监控，可以掌握数据量变化情况，也可以从数据增量角度发现数据入库异常。数据量监测结果可同步到数据台帐，以便数据台帐统计数据总量情况。

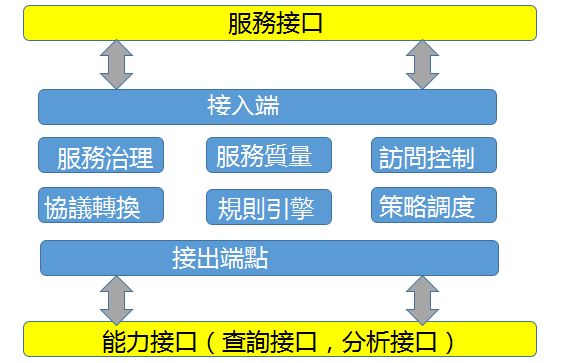
### 元数据管理

元数据是数据仓库中存储的基本单元，实现对元数据的管理，数据仓库的最基本功能之一。元数据管理包括元数据注册登记、元数据存储、元数据建模等多方面功能。

## 数据服务

大数据平台开放存储访问接口，提供基于 Hadoop 技术体系的 HDFS、HBase访问接口，以 OpenAPI 的方式，为应用提供大数据存储服务。

数据服务层主要由数据服务总线来建设，主要负责将大数据平台的能力接口注册进去，再以标准化接口开放给应用系统使用，支持多种协议转换、服务质量 控制、访问控制、规则引擎等。数据服务层将大数据平台的数据服务能力开放出去，供第三方平台使用。



如上图：应用服务系统使用服务接口，来接入数据服务总线，经过数据服务 总线的接入端点，进行过滤。同时根据访问控制、服务质量、协议转换、策略调 度、规则引擎的处理，接出到大数据平台的能力接口。

# 大数据平台

## 大数据平台基础架构

大数据基础平台基于烽火自主知识产权FitData产品，FitData主要集成了基础计算资源、网络资源、存储资源，在统一的安全体管理体系下，将这些资源再进行深度加工、处理、关联，形成多种类型的基础服务能力，构建基础资源层，向应用提供基础资源的服务能力。数据服务总线通过服务治理来维护基础资源服务能力，并通过访 问控制、服务质量、协议转换等，对应用提供多协议支持。平台支撑体系的运维体系提供整体运维能力，保障平台的正常运行；安全体系提供整体安全能力，保障平台的数据安全和使用安全；平台采用分布式架构，支持巨量数据存储与分析， 保障专家管理系统的高性能、高可用性和易扩展性。FitData大数据基础平台结构如下图红线标出部分。



* + 数据计算与存储：是FitData 大数据平台的核心内容，提供分布式存储能力和分布式计算能力。提供的存储框架能力，包括基于结构化数据存储、非结构化数据存储和半结构化数据存储，其计算框架与存储框架均是分布式集群方式部署，可以平滑的进行弹性扩容。
  + 数据服务层：数据服务层主要由数据服务接口来实现，对应用提供数据支撑。通过数据服务接口将平台的数据资源以标准 API 接口的方式开放出来，供不同的应用系统使用。数据应用层主要提供基于该平台来构建的专家系统应用。采用平台的标准API，数据资源层获取数据服务，目前API 接口包括资源目录浏览、数据查询搜索等。
  + 数据汇聚层：提供各层之间数据交换能力，由ETL数据集成工具来实现。平台支持多中异构数据源，针对不同数据源的不同数据，也提供多种数据抽取方式，例如数据库直 连抽取、Sqoop 抽取等。提供计算框架能力，主要集成了批处理计算框 架、流式计算框架、内存计算框架等能力，还提供了像 Hive、Mahout、 Spark 等二次计算能力框架。平台可将这些计算能力开放，供数据模型、数据挖掘、应用系统来使用。
  + 运维体系：运维体系提供面向专家系统完整运维方案， 涵盖了运行监控到使用操作。安全体系提供面向专家系统大数据平台的用户权限管理、终 端访问控制、日志安全审计等能力。

数据存与计算是 FitData 大数据平台核心能力，将目前专家系统内部业务数据源进行有效整合，集成以数据为核心的查询、 分析和管理能力。采用分层整合，灵活配置，横向扩展，纵向贯穿的大数据平台服务能力，其计算框架、存储框架都以容器的方式，可轻松灵活的在线进行装卸，以平滑扩充大数据平台的集成能力。除此还集成了二级计算框架、通用的数据处理算法库和数据仓库，将大数据平台的数据进行清洗、加工和分析挖掘，处理后的数据可订阅，充分体现数据即服务的大数据思想。

• 分布式存储框架：主要负责针对巨量数据的存储，以分布式存储技术， 支持快速、巨量、多种类型的数据存取。支持从数据源抽取数据到大数 据平台存储，集成多种存储方式，有针对结构化数据、非结构化数据和 半结构化数据的存储。

• 计算框架：主要提供批处理计算、内存计算、流式计算框架，由数据处 理管理驱动来分配和调度计算框架，加载数据处理算法，完成数据处理。

• 数据仓库：主要对计算框架完成后的结果进行存储，支持 Hbase、MS SQL Server 等存储，同时将数据以接口的形式开放出去。

• 数据处理算法库：集成通用的数据分析算法、能够插入用户自定义的数 据模型算法，配合以资源管理系统为主的计算存储框架，进行数据处理。

• 资源管理系统，以容器的方式，来为计算框架和存储框架分配资源，并 支持资源调度，弹性伸缩。

• 数据服务总线：主要将基础平台的能力和数据服务接口，以 API 的方式开放出去，形成一个共享的、供应用使用的服务总线。

## FitData特点

* 广泛适应性：支持结构化、半结构化、非结构化数据；支持实时数据。
* 巨量数据：数据处理能力在PB级以上。
* 线性扩展：存储、计算均可增加节点进行线性扩展。
* 统一运维管理：降低安装部署、运营、维护成本。
* 经济性：可运行在普通X86服务器上，硬件成本低。
* 高可靠性：支持容灾容错、备份恢复机制，支持自动告警。支持节点可靠性、数据可靠性。
* 高性能：高效数据处理性能，支持Spark、Storm、R。
* 认证安全：支持Kerberos安全认证、LDAP账户管理控制。
* 数据安全：支持数据加密。
* 负载均衡：支持节点间存储、技术负载均衡。
* 开放性：支持符合Hadoop规范的第三方组件或工具。

## FitData主要功能

FitData是基于开源Hadoop开发的企业级大数据产品，提供PB级数据的采集、存储和处理能力，支持数据加载、查询、分析、挖掘等功能。

### 节点批量自动部署

通过以Web管理，以图形界面的方式实现大数据平台节点批量自动部署，只需添加主机名(或者IP地址)即可实现将节点服务器添加到集群中，截图如下：



图 向集群中添加节点

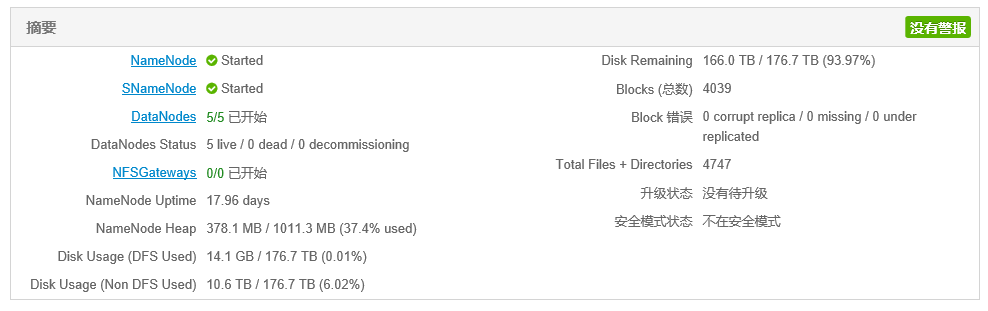
### 节点动态管理

通过web管理实现节点的动态添加、删除，当存储空间或者计算资源不足时，支持向集群中添加同等配置的服务器，实现大数据平台在线动态扩容，而不需要停机处理，不影响平台正常运行。

大数据平台以Web图形界面实现Hadoop集群监控，包括大数据平台的硬件资源、软件资源、数据资源的监控，以及整个Hadoop集群的工作负载。主要包括以下几个方面：

### 服务组件状态监控

通过管理平台可以看到所有目前已安装的服务组件的健康状况。



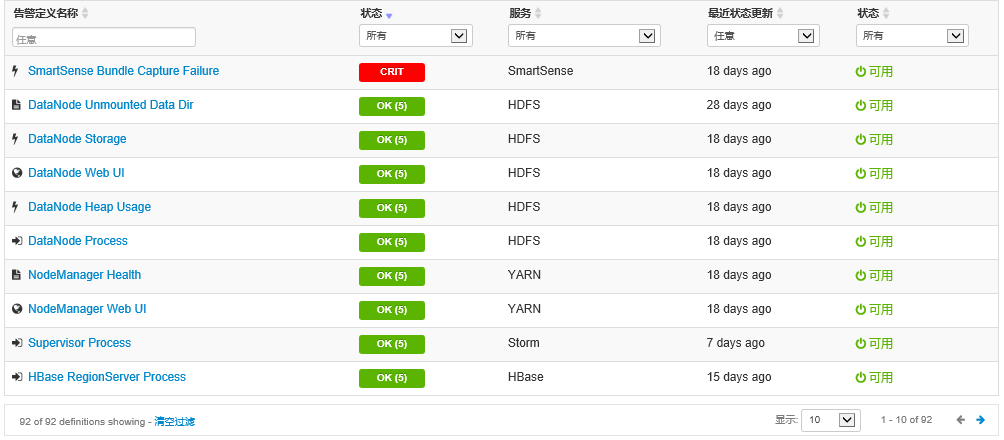


图 服务组件运行状况

### 计算资源负载监控

通过管理平台可以实时看到整个平台的资源负载情况，包括集群的CPU、集群磁盘IO、集群网络IO、HDFS IO，如下图所示：

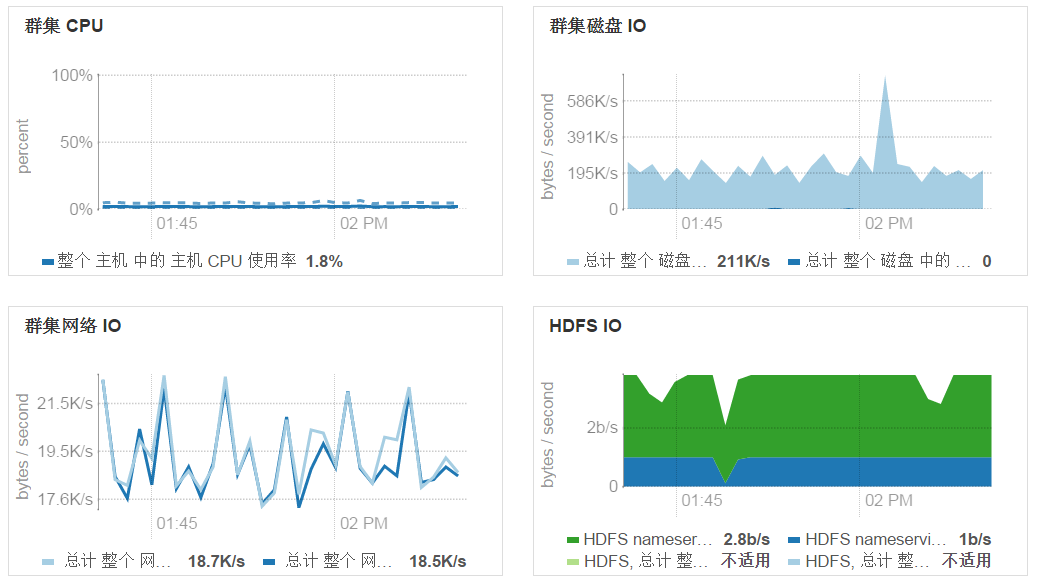


图 计算资源监控

### 多任务实时监控

通过对集群运行任务的实时监测，并根据任务优先级和耗时不同对任务进行动态调度，减少出现大量任务等待和重要任务无法及时完成的可能，可以使Hadoop集群的运行变得更加高效合理。

（1）、系统根据各队列资源的最小值分配集群资源，这样可以按照需求对各任务队列获取的集群资源进行分配，而且不会出现集群资源的闲置浪费。

（2）、可以实现对各任务队列获取的集群资源大小实时动态调整，及时保证高优先级任务所在队列获得更多的集群资源。

（3）、可以实现在某个任务队列出现空闲时，将该任务队列获取的集群资源自动分配给其他繁忙的任务队列，以使得集群资源利用最大化。

### 磁盘性能监控

对集群机器的硬盘进行监控，如下图所示，详细的展示出磁盘IO的利用率，读写速度，磁盘的等待时间。



图：磁盘性能监控

### 故障快速定位

大数据平台具备完整的告警监控和故障快速定位能力。能够将计算框架的每个作业进度、状态、资源利用情况进行监控，并通过可视化图形界面进行展示。

当大数据平台出现异常情况时，平台能够通过监控系统，对服务器节点宕机、集群异常、安全异常等异常事件进行预警、报警，并通过邮件、短信报警手段进行告警通知。提供预制的恢复规则和安全规则，对集群异常进行自动修复、自动限制非安全行为的操作。

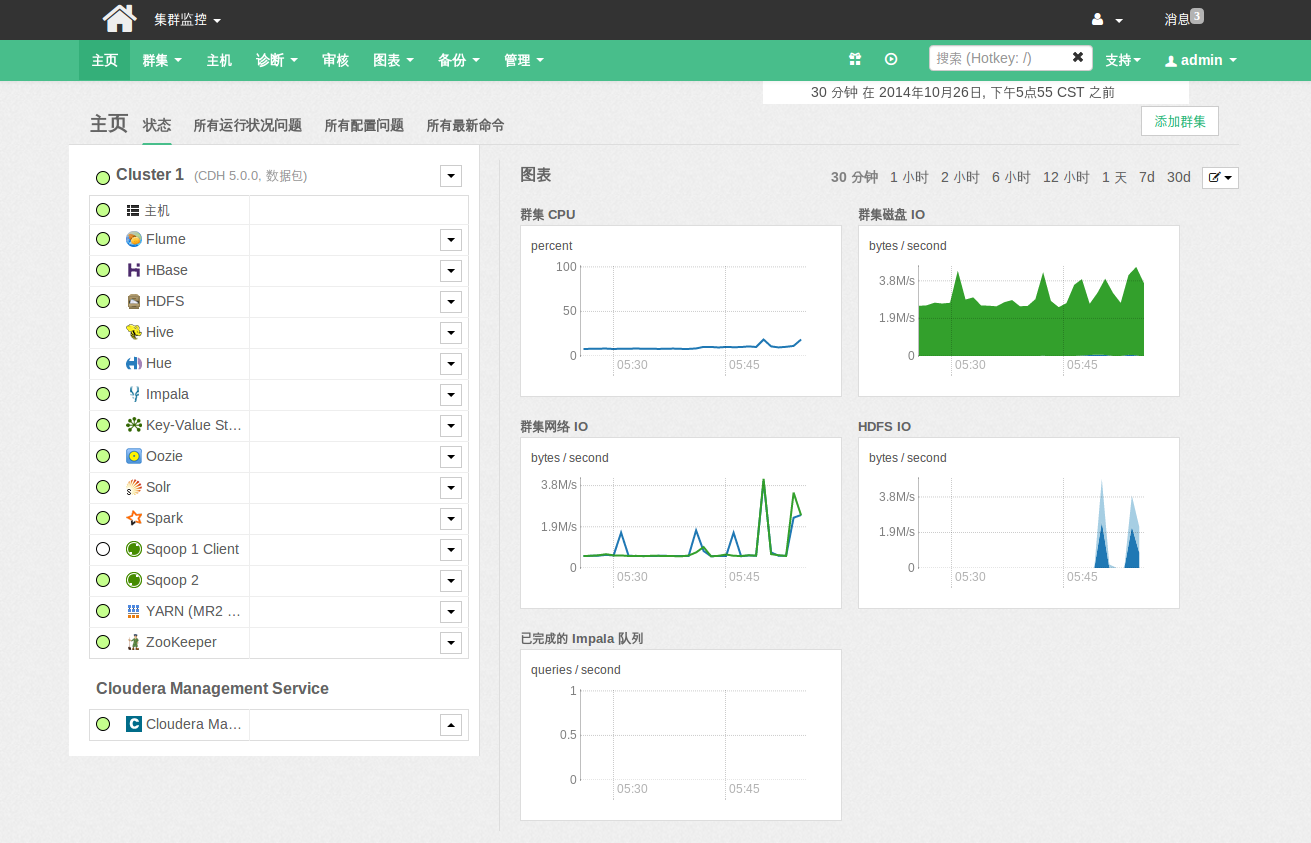
大数据平台能够通过对告警信息的分析，快速定位平台内部出现故障的节点，对于因故障无法继续提供服务器的节点进行标记，将平台的作业任务自动分配到其他的节点上运行，同时，大数据平台采用分布式体系结构及无单点故障设计，平台内任何节点的宕机都不会影响平台的稳定运行和业务的正常使用。待故障节点恢复正常后，再将该节点纳入平台的资源中，将作业任务分配到恢复后的节点上运行。

### 日常运维监控

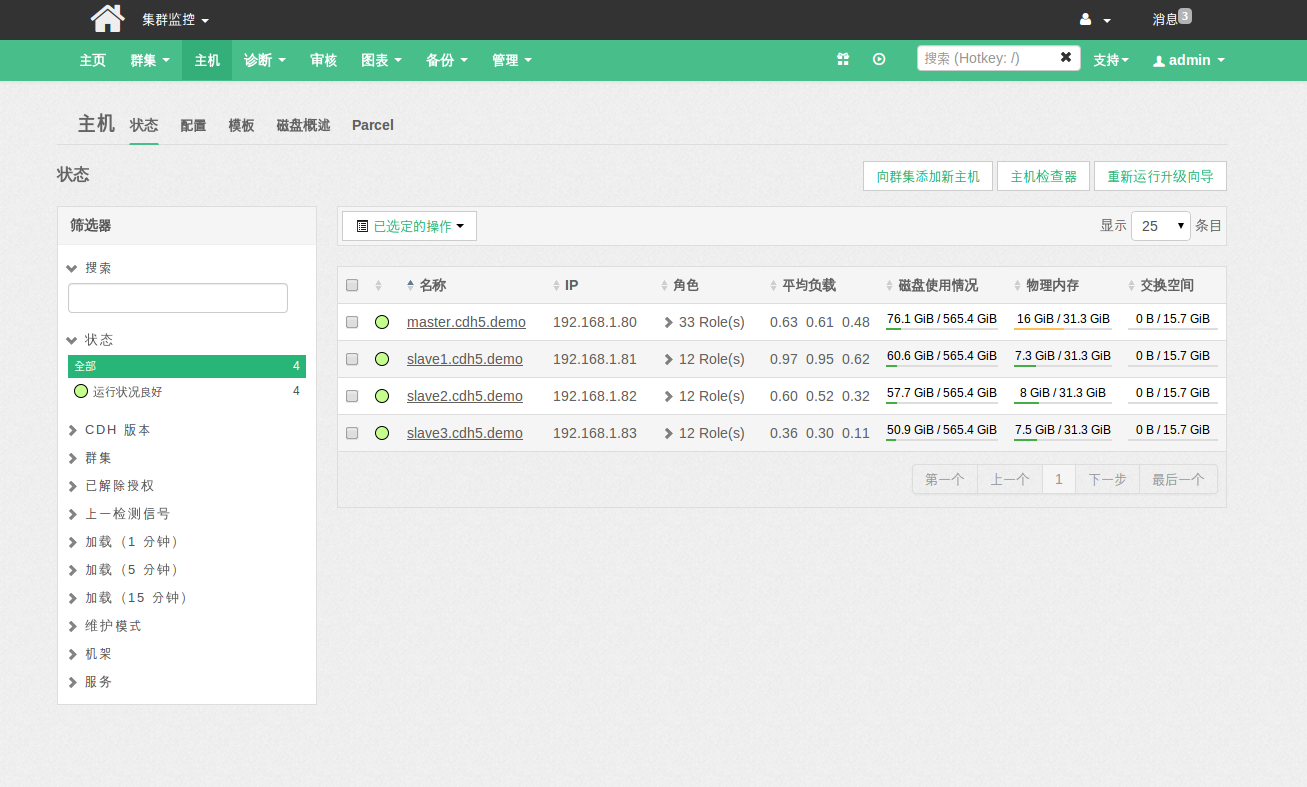
大数据综合平台提供完整的日常运维监控的服务能力，针对从上层应用平台到底层基础平台的各个功能模块和组件均提供有监控能力，能够分析系统的运行日志和用户日志，并且能够将监控数据通过文件接口或webservice接口的方式汇总到平台管理运维模块的监控管理界面中进行统一呈现和管理使用。系统能够根据监控到的数据进行分析判断，对异常的数据触发告警，在前台界面提醒，直至出发通知和处理等进一步动作。

平台的监控范围涵盖有：

* 平台管理资源的使用与分配
* 服务器视图：提供针对各服务器和存储等设备的资源使用情况的实时查看，包括当前设备的CPU负荷，内存占用情况，存储空间使用情况，网络带宽占用情况、设备运行状态等。管理员能够根据监控信息在管理平台上有效调度分配系统资源。其中集群的监控如下图所示：



针对服务器的监控如下图所示：



* 服务视图：提供系统中各服务资源使用情况的实时查看，包括连接数、当前作业数，I/O情况，运行状态等。
* 监控系统的运行情况
* 接口服务运行监控：提供针对数据源和应用层的监控服务，包括运行状态和流量等信息；
* 数据存取过程监控：提供针对数据存储过程的监控服务，包括系统平台的I/O情况（整体I/O和具体各节点I/O以及具体的各作业的I/O情况）和数据存取过程的任务列表；
* 数据汇聚过程监控：监控系统的数据汇聚过程，包括使用资源信息，使用的数据源信息，作业进程运行状况信息，使用时间/计划完成时间等信息；
* 数据处理过程监控（作业监控）：监控系统的数据处理（作业）过程，包括使用资源信息，使用的数据源信息，作业进程运行状况信息，使用时间/计划完成时间等信息；
* 应用监控：针对运行在平台上的应用进行监控，包括各应用当前的运行状态、应用对数据的使用状况，应用为用户提供的查询数量等；
* 系统异常告警与处理
* 用户告警：对用户操作使用过程中的异常行为进行告警，例如某用户访问了超过其正常权限的数据等。
* 系统告警：对系统中存在的服务节点宕机，系统接口异常，数据存储报错，系统资源紧张等系统运行异常情况进行告警触发，并提醒用户进行操作处理。

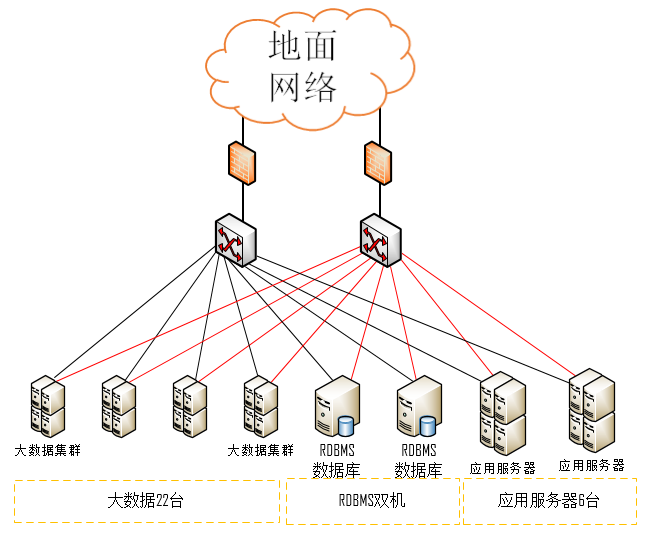
## FitData优势

烽火大数据平台FitData借助先进开源的大数据存储及处理技术，成功实施了公安大数据平台、楚天云政务大数据平台，通过大数据项目的实施，逐步沉淀了大量的算法模型及分析与展示工具，在平台性能及稳定性上经历了实战的考验，逐步总结出一套FitData自己的系统优化策略及系统运维策略，平台经受住了单节点超过1000台集群的实战考验，并支持HA高可用性运行策略，经过四年时间及高强度项目的锤炼，FitData大数据平台已经走出了自己的路。在数据处理上支持PB及超大量数据的秒级查询及汇集。

SmartAS是企业级基础开发平台，它基于FitData平台之上，采用微服务架构，支持分布式部署，是成熟可靠的多终端应用开发框架。它集成业界流行和成熟的技术框架，通过应用系统使用，反馈的情况不断完善应用框架的通用功能，满足业务系统快熟构建的目标，具备良好用户体验

# 硬件部署

按照专家系统安装接口规范要求，结合专家管理系统数据量估算值和数据存储特点，本着数据安全、系统稳定可靠的核心设计思路，设计专家系统大数据平台数据节点服务器22台，其中管理节点服务器2台，数据节点服务器19台，监控节点一台，系统RDBMS数据库服务器台，应用服务器6台，绘制专家系统部署逻辑结构图如下:



# 硬件清单

根据系统规划及安装接口规范要求，初步规划服务器如下：系统应用服务器需求6台；大数据平台设计节点22个，其中管理节点2个，数据节点19个，监控节点服务器1台，RDBMS数据库服务器两台双机热备。具体各服务器硬件需求如下表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 编号 | 服务器名 | 配置 | 数量 | 说明 |
| 1 | RDBMS数据库服务器 | 4\*Intel Xeon E7-4800/8800 v3  最大可扩展至4 CPU，72 核  支持8GB/16GB/32GB/64GB DDR4 高速内存  配置128GB DDR4 内存  配置9 块900GB 15K SAS，14\*4T NL SAS 硬盘。 | 2 | 双机备份 |
| 2 | 大数据平台管理节点 | 2\*Intel Xeon E7-4800/8800 v3  最大可扩展至4 CPU，72 核  支持8GB/16GB/32GB/64GB DDR4 高速内存  配置128GB DDR4 内存  配置6 块600GB 15K SAS，3\*4T NL SAS 硬盘。 | 1 | Active |
| 3 | 大数据平台管理节点 | 2\*Intel Xeon E7-4800/8800 v3  最大可扩展至4 CPU，72 核  支持8GB/16GB/32GB/64GB DDR4 高速内存  配置128GB DDR4 内存  配置6 块600GB 15K SAS，3\*4T NL SAS 硬盘。 | 1 | Standby |
| 4 | 大数据平台数据节点 | 2\*Intel Xeon E7-4800/8800 v3  最大可扩展至4 CPU，72 核  支持8GB/16GB/32GB/64GB DDR4 高速内存  配置128GB DDR4 内存  配置6 块600GB 15K SAS，12\*4T NL SAS 硬盘。 | 19 | 数据节点 |
| 5 | 大数据集群性能检测服务器 | 2\*Intel Xeon E7-4800/8800 v3  最大可扩展至4 CPU，72 核  支持8GB/16GB/32GB/64GB DDR4 高速内存  配置128GB DDR4 内存  配置6 块600GB 15K SAS，3\*4T NL SAS 硬盘。 | 1 | 监控节点 |
| 6 | 应用服务器 | CPU：2 颗E5-2630 v3  ≥24 个内存插槽，最大支持1.5TB 内存，支持2133 MHz 内存。当前配置64GB 内存。  支持SAS、SSD 和PCIe SSD 硬盘，支持2.5  寸和3.5 寸硬盘混插。  支持24+2 个2.5 寸 SAS/SATA 或者 14 个3.5  寸 SAS/SATA + 2 个2.5 寸SAS/SATA +16 个  1.8" SSD。  硬盘：配置6 块600GB 15K SAS 硬盘 | 2 | 应用服务器 |
| 7 | 交换机 | 48 10/100/1000Base-TX，  4 100/1000Base-X SFP | 2 | 网络设备 |
| 8 | 防火墙 | 多功能防火墙，4口以上 | 2 | 安防设备 |
| 9 | 工作站 | Intel(R)Xeon CPU E5，配置1T SATA 硬盘。内存：8GB | 2 |  |

**说明：硬件部分交换机、防火强及工作站，请根据标书确认！大数据服务器、RDBMS数据库服务器及应用服务器的具体配置参数请硬件朋友和标书上进行重新确认，这边只对内存量、CPU颗数及存储空间大小做了要求。**

# 个人介绍

吴宏勋：“烽火集成”高级大数据架构师，曾担任医疗大数据、公安大数据、财税大数据项目大数据架构师，具有丰富的大数据项目实施经验，对高吞吐、高并发、海量数据实时汇集，TB、PB级海量数据即席查询与实时处理具有针对性方案和经验，研读过部分Hadoop、HBase、Spark源码，对Hadoop、HBase、Spark的原理有很深的理解，曾从事多个项目大数据平台的调优工作！

# 专家系统架构设计



本系统总共分为四个层次，从下到上依次为数据采集层、基础平台层、应用支撑层、应用及展示层，各层在专家系统统一业务规范、技术规范、安全规范下进行数据通信及集成。

1. 数据采集层：负责专家系统信息数据的汇集、转换与加载，数据采集层提供多种数据采集方法:ETL、Flume、Kafka等，系统支持Flume+Kafka+Storm混合架构的数据采集模式，以提高数据采集系统的吞吐量和并发量。
2. 基础平台层：基础平台层为专家数据仓库提供大数据基础平台支撑，包括分布式存储系统、Hbase数据库系统、Yarn并行计算资源管理与监控等，同时支持Spark 机器学习算法库，支持R等行业分析库。
3. 应用支撑层：应用支撑层为系统各类应用提供支撑，是系统数据层和应用层的连接纽带。应用支撑层包括基础平台和常规算法两个部分，基础平台负责数据的存储与并行计算，数据存储支持分布式存储、RDBMS存储等存储方式，常规算法负责数据分析与业务建模。
4. 应用及展示层：应用层是系统各项业务功能的集合，主要包括资车辆故障诊断、车辆健康评估、车辆部件检修、车辆故障处理及车辆对比分析等。展示层是用户同系统交互的窗口，是应用层对外提供服务的主要手段。支持多种图表展示如饼图、柱状图、曲线图、热力图、气泡图和散点图等可视化展示。

# 平台运维管理

## Hadoop集群监控

大数据平台以Web图形界面实现Hadoop集群监控，包括大数据平台的硬件资源、软件资源、数据资源的监控，以及整个Hadoop集群的工作负载。主要包括以下几个方面：

### 服务组件状态监控

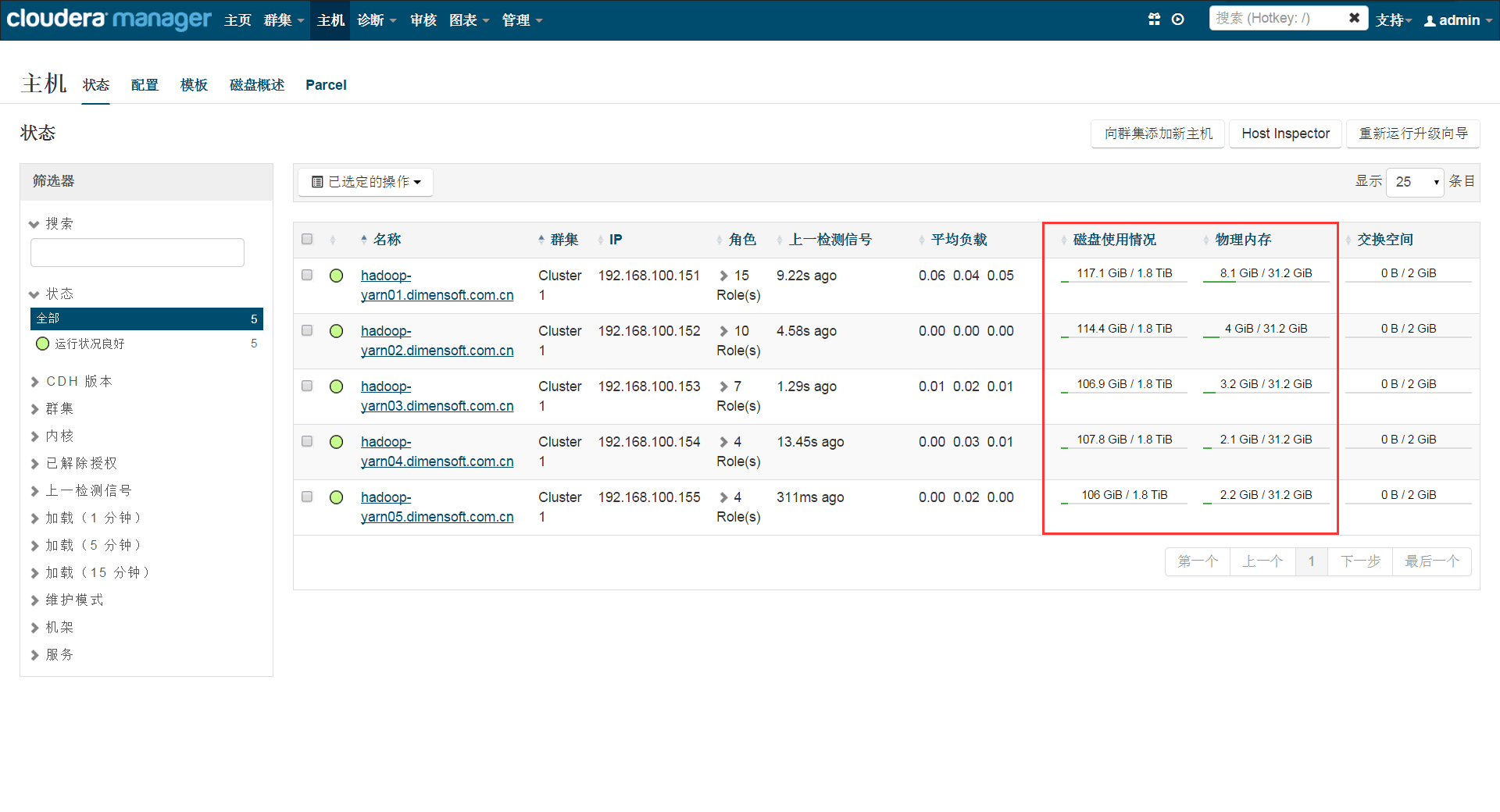
通过管理平台可以看到所有目前已安装的服务组件的健康状况，绿色圈表示运行状态健康。



图：服务组件运行状况

### 存储与内存资源监控

包括获取存储量、剩余存储量以及存储系统整体情况信息。如果集群中的某台机器的磁盘或者内存的使用率达到指定的阀值，系统可以通过邮件或者短信的方式进行预警。

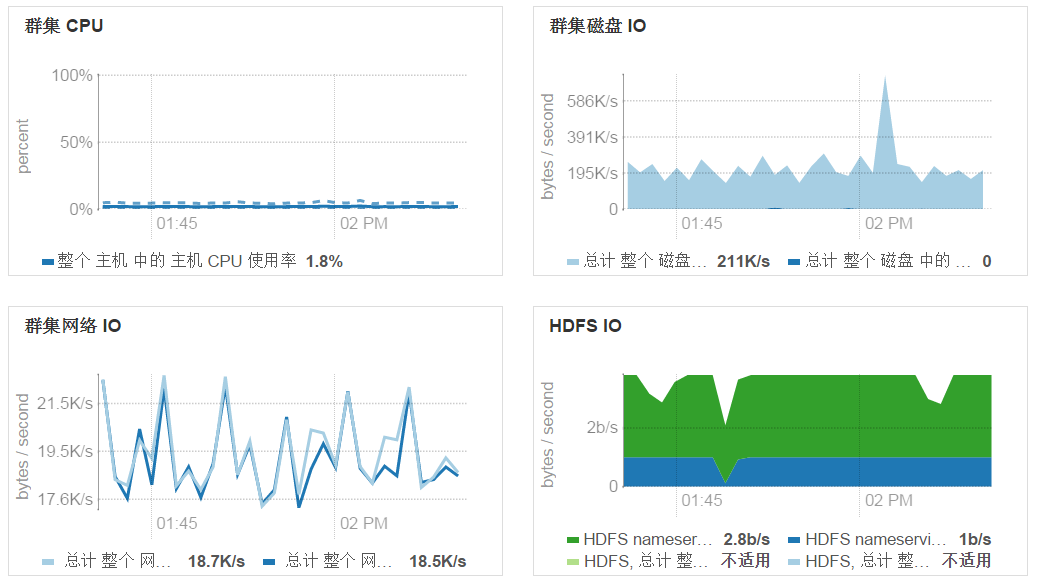


图：存储和内存资源监控

## 系统负载管理

I

通过管理平台可以实时看到整个平台的资源负载情况，包括集群的CPU、集群磁盘IO、集群网络IO、HDFS IO，如下图所示：



通过对集群运行任务的实时监测，并根据任务优先级和耗时不同对任务进行动态调度，减少出现大量任务等待和重要任务无法及时完成的可能，可以使Hadoop集群的运行变得更加高效合理。

（1）、系统根据各队列资源的最小值分配集群资源，这样可以按照需求对各任务队列获取的集群资源进行分配，而且不会出现集群资源的闲置浪费。

（2）、可以实现对各任务队列获取的集群资源大小实时动态调整，及时保证高优先级任务所在队列获得更多的集群资源。

（3）、可以实现在某个任务队列出现空闲时，将该任务队列获取的集群资源自动分配给其他繁忙的任务队列，以使得集群资源利用最大化。

## 操作系统管理

### 磁盘性能监控

对集群机器的硬盘进行监控，如下图所示，详细的展示出磁盘IO的利用率，读写速度，磁盘的等待时间。



图：磁盘性能监控

### 故障快速定位

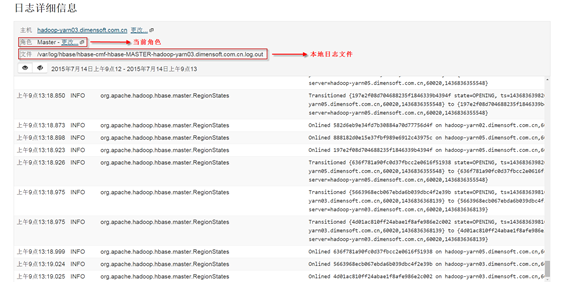
大数据平台具备完整的告警监控和故障快速定位能力。能够将计算框架的每个作业进度、状态、资源利用情况进行监控，并通过可视化图形界面进行展示。

当大数据平台出现异常情况时，平台能够通过监控系统，对服务器节点宕机等集群异常、安全异常等异常事件进行预警、报警，并通过邮件、短信等报警手段进行告警通知。提供预制的恢复规则和安全规则，对集群异常进行自动修复、自动限制非安全行为的操作。

大数据平台能够通过对告警信息的分析，快速定位平台内部出现故障的节点，对于因故障无法继续提供服务器的节点进行标记，将平台的作业任务自动分配到其他的节点上运行，同时，大数据平台采用分布式体系结构及无单点故障设计，平台内任何节点的宕机都不会影响平台的稳定运行和业务的正常使用。待故障节点恢复正常后，再将该节点纳入平台的资源中，将作业任务分配到恢复后的节点上运行。

### 运行日志监控

针对每个服务组件运行的实时日志信息可以从平台中查看，便于在服务组件运行中断时查找和追踪原因。例如，我们想要查看HBase服务组件中Mater角色的日志信息，如下图所示：



## 平台安全管理

在Hadoop 2.x中加入了Kerberos认证机制。Kerberos可以将认证的密钥在集群部署时事先放到可靠的节点上。集群运行时，集群内的节点使用密钥得到认证。只有被认证过节点才能正常使用，防止恶意的使用或篡改Hadoop集群的问题，确保Hadoop集群的可靠安全。

## 数据质量管理

### 数据标准化

数据标准化包括数据标准制定及数据标准化处理两个部分，数据标准制定是在专家系统业务统一规范前提下，指导专家系统大数据标准，包括数据格式标准、数据交换标准、数据共享标准等；数据标准规范化是指按照统一专家系统数据标准格式。将专家信息数据进行标准化处理，生成符合专家系统数据标准要求的信息数据。

### 数据质量检测

根据数据质量监测规则，通过数据质量检测引擎，对数据表中的增量数据进行扫描，调用规则算法或扩展程序进行数据质量检测，并提供问题数据库的建立、数据质量报告的生成、问题数据的处理、以及对问题数据的通报和反馈来保证数据的质量和实效性等功能。

### 数据关联

对采集的数据库根据数据间的业务关联关系实现数据的关联，通过数据的关联，增加实体数据的维度，将单个的数据扩展成行业信息资源，提高数据的价值。