

目录

目录	1
Ambari Web基本操作	3
Ambari简介	3
Ambari概述	3
Ambari体系结构	3
KMR中Ambari的入口	3
Ambari Web基本操作	3
仪表盘 (Dashboard)	3
Metrics	3
度量小程序及描述	4
HeatMaps	4
各组件具体信息解释如下:	4
Config History	5
Service管理服务	5
服务管理	5
对各组件操作及解释如下:	5
Summary	6
主机 (Hosts) 管理	6
主机状态	6
主机搜索	7
主机管理	7
主机操作	7
监控管理/Alerts	7
Admin/管理	8
文件查看	9
文件权限	9
Ambari去除Token认证	9
使用Ambari监控集群	10
准备	10
服务级别的监控管理	11
主机 (集群节点) 级别的监控管理	12
警告管理	13
使用Ambari Admin账户安装组件	14
Ambari告警信息	17
告警级别	17
告警 (alert) 类型	18
告警 (alert) 类型	18
Ambari 告警description翻译	18
HDFS说明	18
YARN	19
MapReduce2	20
Hive	20
HBase	20
Oozie	20
ZooKeeper	20
Storm	21
Kafka	21
spark2	21
ElasticSearch	21
Hue	21

Ambari	21
Ambari Metrics	21
去除Ambari页面的token认证	22
Ambari集群配置的实践指南	22
管理集群配置	22
修改 NameNode Java 堆栈大小	22
Hive配置优化实践指南	23
设定Hive集群引擎	23
优化分组参数	23
调优reducer	23
启用并行执行	24
启用矢量化	24
启用基于成本的优化 (CBO)	24
启用中间压缩	25
压缩最终输出	25
启用推理执行	25
优化动态分区	25
启用本地模式	26
设置单个 MapReduce MultiGROUP BY	26
其他 Hive 优化	26
联接优化	26
执行引擎优化	26
HBASE配置优化实践指南	26
设置 HBASE HEAPSIZE	26
优化读取密集型工作负荷	26
块缓存大小	26
Memstore 大小	27
从磁盘扫描时提取的行数	27
优化写入密集型工作负荷	27
最大区域文件大小	27
避免阻止更新	27
定义 Memstore 大小	27
设置 Memstore 本地分配缓冲区	27
开启服务的自动启动	27
设置HDFS再平衡 (rebalance)	28
设置ES balance的阈值	30

Ambari Web基本操作

Ambari简介

Ambari概述

Apache Ambari是一种基于Web的工具，支持Apache Hadoop集群的创建、管理和监控。Ambari已支持大多数Hadoop组件，包括HDFS、MapReduce、Hive、Pig、Hbase、Zookeeper、Sqoop和Hcatalog等；除此之外，Ambari还支持Spark、Storm等计算框架及资源调度平台YARN。

Apache Ambari 从集群节点和服务收集大量信息，并把它们表现为容易使用的，集中化的接口：Ambari Web。

Ambari Web显示诸如服务特定的摘要、图表以及警报信息。可通过Ambari Web对Hadoop集群进行创建、管理、监控、添加主机、更新服务配置等；也可以利用Ambari Web执行集群管理任务，例如启用Kerberos 安全以及执行Stack升级。任何用户都可以查看Ambari Web特性。拥有administrator-level 角色的用户可以访问比 operator-level 或 view-only 的用户能访问的更多选项。例如，Ambari administrator 可以管理集群安全，一个 operator 用户可以监控集群，而 view-only 用户只能访问系统管理员授予他的必要的权限。

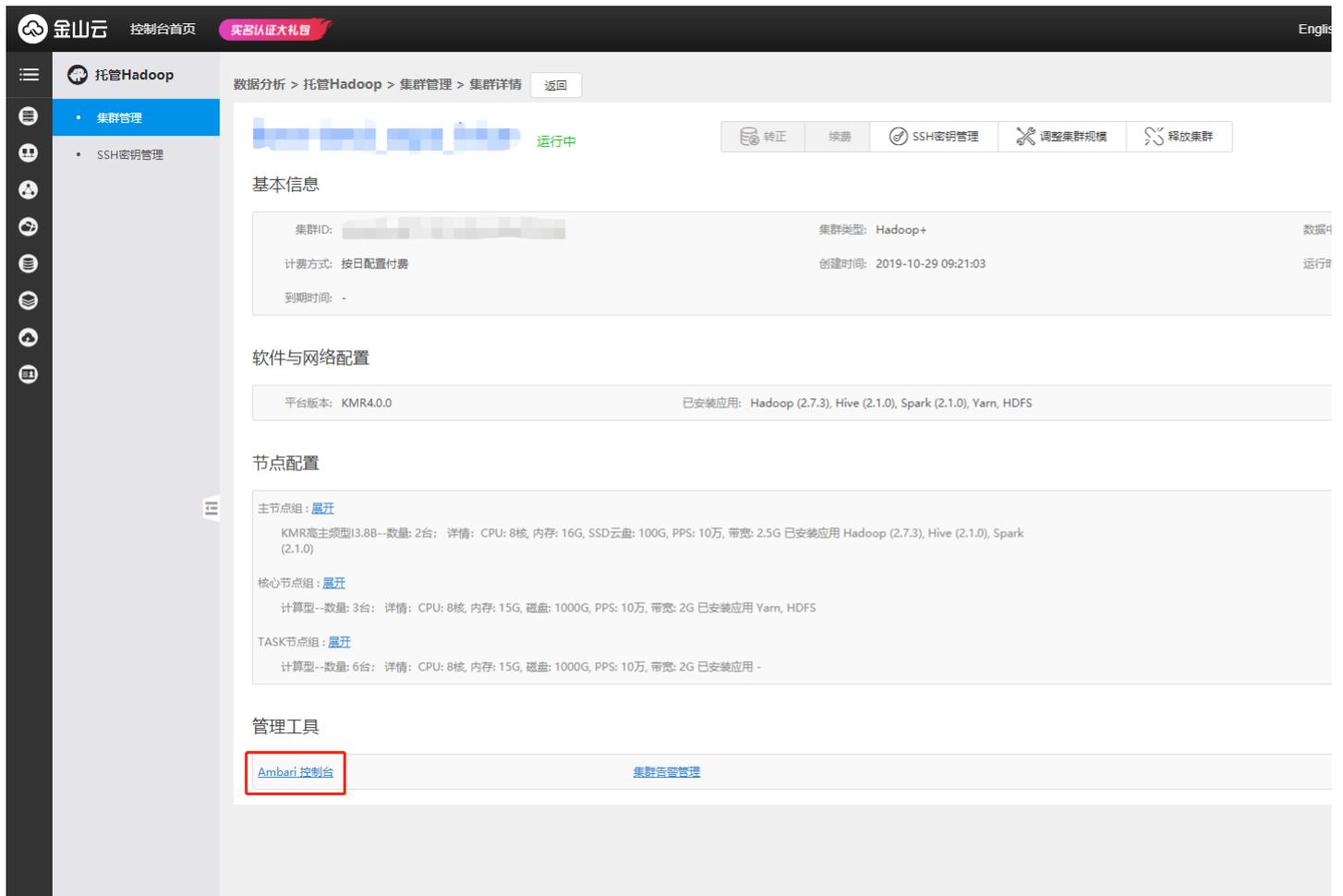
Ambari体系结构

Ambari 自身也是一个分布式架构的软件，主要由两部分组成：Ambari Server 和 Ambari Agent。简单来说，用户通过Ambari Server通知 Ambari Agent 安装对应的软件；Agent 会定时地发送各个机器每个软件模块的状态给 Ambari Server，最终这些状态信息会呈现在 Ambari 的 GUI，方便用户了解到集群的各种状态，并进行相应的维护。

Ambari Server 从整个集群上收集信息。每个主机上都有 Ambari Agent，Ambari Server 通过 Ambari Agent 控制每部主机。

KMR中Ambari的入口

通过KMR管理工具进入Ambari控制台。

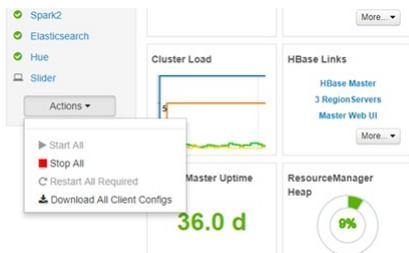


Ambari Web基本操作

仪表盘 (Dashboard)

仪表盘页为Ambari Web主页，用来查看集群的操作状态。

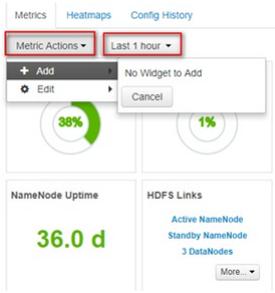
左侧一栏显示当前运行的Hadoop服务列表。在下边Actions下拉菜单中可进行服务列表的全部启动及全部暂停，以及下载全部客户端的配置信息，所下载文件为tar.gz格式，将其解压后可看各组件配置信息。



仪表盘包括Metrics、Heatmaps、Config History选项卡；默认显示 Metrics。

Metrics

在Metrics页面上，有多小程序(widget)，显示 Hadoop集群服务的操作状态信息。小程序以Graph (图)、Gauge (百分比)、Number (数字) 及Template(模板) 四种形式进行展示。如下图，在Metrics中，可进行小程序的添加及查看特定时间段内的信息。



将鼠标移至小程序页面上时，其展现出具体信息内容；点击X，可对其进行移除；点击编辑（铅笔）图标，可对其进行编辑。



度量小程序及描述

模块	小程序	描述
HDFS metrics	HDFS Disk Usage	分布式文件系统(DFS) 已使用的百分比，包括 DFS 和 non-DFS
	Data Nodes Live	运转中的 DataNodes 的数量，由 NameNode 报告
	NameNode Heap	NameNode Java Virtual Machine (JVM) 堆内存使用的百分数
	NameNode RPC	潜在 RPC 队列平均水平 (The average RPC queue latency)
	NameNode CPU WIO	CPU wait I/O 百分比
	NameNode Uptime	NameNode 正常运行时间计算值 (uptime calculation)
YARN metrics (HDP 2.1 or later)	ResourceManager Heap	已使用的 ResourceManager JVM 堆内存百分数
	ResourceManager Uptime	ResourceManager uptime
	NodeManagers Live	运转中的 DataNodes 数量，由 ResourceManager 报告
HBase metrics	YARN Memory	可用的 YARN 内存百分数 (已用/总可用)
	HBase Master Heap	已使用的 NameNode JVM 对内存百分数
	HBase Ave Load	HBase server 上的平均负载
	HBase Master Uptime	HBase master uptime
Storm metrics (Hadoop 2.1 or later)	Region in Transition	转换中的 HBase regions 数量
	Supervisors Live	运转中的 supervisor 的数量，由 Nimbus Server 报告

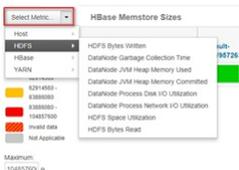
HeatMaps

在HeatMaps中显示小程序的操作状态及条目

颜色	状态
实心绿 (solid green)	所有主节点正常 (All masters are running)
闪烁绿 (blinking green)	启动中 (Starting up)
实心红 (solid red)	至少一个主节点停止 (At least one master is down)
闪烁红 (blinking red)	正在停止 (Stopping)



也可在Select Metrics中查看其他条目状态



各组件具体信息解释如下：

组件	类别	解释
HDFS	HDFS Bytes Written	HDFS写入字节数
	DataNode Garbage Collection Time	DataNode垃圾收集时间
	DataNode JVM Heap Memory Used	DataNode 虚拟机堆内存使用
	DataNode JVM Heap Memory Committed	DataNode 虚拟机已提交堆内存
	DataNode Process Disk I/O Utilization	DataNode进程磁盘I/O使用
	DataNode Process Network I/O Utilization	DataNode进程网络I/O使用
	HDFS Space Utilization	HDFS空间使用
	HDFS Bytes Read	HDFS读取字节数
	YARN	Total Allocatable CPU Utilized per NodeManager
Container Failures		Container故障数
NodeManager GC Time		NodeManger垃圾收集时间
NodeManager JVM Heap Memory Used		NodeManger JVM 堆内存使用
Allocated Containers		所分配Container
NodeManager RAM Utilized		NodeManger RAM使用情况
	NodeManager CPU Utilized	NodeManger CPU使用情况

	Total Allocatable RAM Utilized per NodeManager	每个NodeManager可分配RAM总量
HBase	HBase Memstore Sizes	HBase Memstore量
	HBase Read Request Count	HBase读请求数
	HBase Write Request Count	HBase写请求数
	HBase Regions	HBase分区
	HBase Compaction Queue Size	HBase可压缩队列大小

Config History

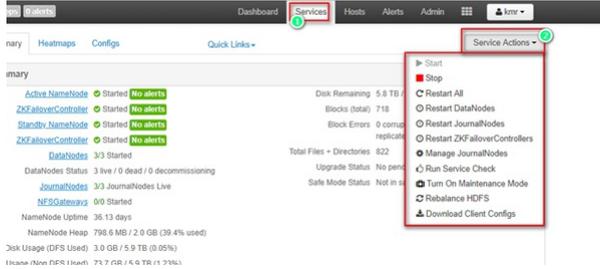
可在此页面中查看各栏目的配置历史信息。

Service管理服务

利用 Services 选项卡可监控和管理运行于集群上选定的服务。集群上安装的所有服务列于左侧的面板上。

服务管理

在Service页面中点击Service Actions下拉菜单，可对左侧所选服务进行启动，停止，重启等操作。



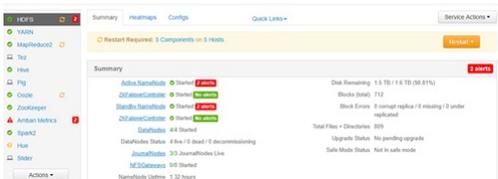
对各组件操作及解释如下：

组件	操作	解释
HDFS	Start	启动
	Stop	停止
	Restart All	全部重启
	Restart DataNodes	重启所有DataNode
	Restart JournalNodes	重启所有JournalNode
	Restart ZKFailoverControllers	重启所有ZKFailoverController
	Manage JournalNodes	管理JournalNode
	Run Service Check	运行服务检查
	Turn On Maintenance Mode	开启维护模式
	Rebalance HDFS	HDFS重平衡
YARN	Download Client Configs	下载客户端配置文件
	Start	启动
	Stop	停止
	Refresh YARN Capacity Scheduler	刷新YARN容量调度
	Restart All	全部重启
	Restart NodeManagers	重启所有NodeManager
MapReduce2	Run Service Check	运行服务检查
	Turn On Maintenance Mode	开启维护模式
	Download Client Configs	下载客户端配置文件
	Start	启动
	Stop	停止
Tez	Restart All	全部重启
	Run Service Check	运行服务检查
	Refresh configs	刷新配置
	Download Client Config	下载客户端配置文件
Hive	Start	启动
	Stop	停止
	Restart All	全部重启
	Run Service Check	运行服务检查
	Turn On Maintenance Mode	开启维护模式
	Restart LLAP	重启LLAP
HBase	Download Client Configs	下载客户端配置文件
	Start	启动
	Stop	停止
	Restart All	全部重启
	Restart RegionServers	重启所有RegionServer
	Run Service Check	运行服务检查
Pig	Turn On Maintenance Mode	开启维护模式
	Download Client Configs	下载客户端配置文件
	Run Service Check	运行服务检查
	Refresh Configs	刷新配置文件
Oozie/Zookeeper	Download Client Configs	下载客户端配置文件
	Start	启动
	Stop	停止
	Restart All	全部重启
	Run Service Check	运行服务检查
Storm	Turn On Maintenance Mode	开启维护模式
	Start	启动
	Stop	停止
	Restart All	全部重启
	Restart Supervisors	重启所有Sepervisor
Ambari Metrics	Run Service Check	运行服务检查
	Turn On Maintenance Mode	开启维护模式
	Start	启动
	Stop	停止

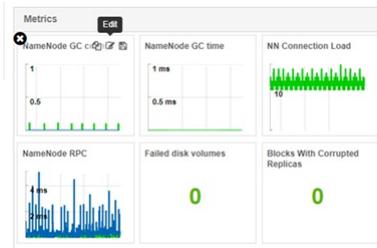
	Restart All	全部重启
	Restart Metrics Monitors	重启所有Metrics Monitor
	Run Service Check	运行服务检查
	Turn On Maintenance Mode	开启维护模式
Kafka	Start	启动
	Stop	停止
	Restart All	全部重启
	Restart Kafka Brokers	重启所有Kafka Broker
	Run Service Check	运行服务检查
	Turn On Maintenance Mode	开启维护模式
Spark2	Start	启动
	Stop	停止
	Restart All	全部重启
	Restart Livy for Spark2 Servers	重启所有Spark2服务器的Livy
	Run Service Check	运行服务检查
	Turn On Maintenance Mode	开启维护模式
ElasticSearch	Download Client Configs	下载客户端配置文件
	Start	启动
	Stop	停止
	Restart All	全部重启
	Run Service Check	运行服务检查
	Turn On Maintenance Mode	开启维护模式
Hue	Start	启动
	Stop	停止
	Restart All	全部重启
	Turn On Maintenance Mode	开启维护模式
	Usersync	用户同步
	Metastoresync	元数据存储同步
Slider	Run Service Check	运行配置检查
	Refresh Configs	刷新配置文件
	Download Client Configs	下载所有客户端配置信息

Summary

在Summary页面下显示有关所选服务操作状态的基本信息，包括警报。要刷新监控面板并显示另一个服务的信息，可以在服务列表上单击一个不同的服务名称。注意服务名称后面带有颜色的图标，指出服务的操作状态和该服务生成的警报。可以单击一个 **View Host** 链接来查看组件和运行选定组件的主机。可单击页面中的 **Alerts** 来查看所有健康检查列表以及所选中服务的状态，重要警报首先显示。要查看警报定义，可以单击列表中每个警报消息的文本标题来查看警报定义。



在仪表盘中对Widget进行添加、移除、创建和删除。将鼠标移至下方Metrics面板中的widget，可对其进行操作。



HeatMaps和Config history信息见Dashboard相关说明，其在两页面中作用相同。

主机（Hosts）管理

作为集群系统管理员或集群操作员，需要知道每部主机的操作状态。也需要知道哪部主机有问题需要处理。可以使用 Ambari Web Hosts 页面来管理多个Hortonworks Data Platform (HDP) 组件，例如运行在整个集群上: DataNodes, NameNodes, NodeManagers, 和 RegionServers。举例来说，可以重启所有的DataNode 组件，可选地控制滚动重启任务。Ambari Hosts 可以过滤进行管理的主机组件选取，基于操作状态，主机健康状况，以及定义的主机分组。

主机状态

可在Ambari Web Hosts页面查看集群上单个主机的状态。主机以 fully qualified domain name (FDQN) 的形式列出，并附有一个带有颜色的图标指示出主机的操作状态。

可点击主机后Components下component数目查看主机上具体组件。

颜色	状态
红色三角形	该主机上至少有一个 master 组件挂掉了，鼠标悬停图标上查看一个工具提示列出受影响的组件
橘色	该主机上至少有一个 slave 组件挂掉了，鼠标悬停图标上查看一个工具提示列出受影响的组件
黄色	Ambari Server 没有从该主机上收到心跳包超过 3 分钟
绿色	正常运行状态
Maintenance Mode	黑色“医药箱”图标指出一部主机处于维护模式
Alert	红色方框带有一个数字指明该主机上的警报数量

在主机状态中，红色图标覆盖橘色图标，橘色图标覆盖黄色图标。换句话说，一部主机有 master component 宕机附有一个红色图标，即便它可能也有slave component和连接问题。主机处于维护模式或遇到警报，图标出现在主机名右侧。

主机搜索

在主机列表页面，可通过上方搜索框对主机由Host、Service及Component中相应栏目进行相应搜索。

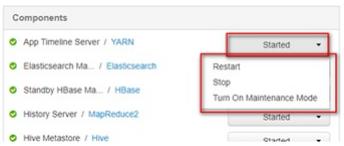


主机管理

点击主机列表页面中主机信息，可查看主机中的具体组件及Host Metrics信息；点击Host Actions下拉菜单，可对所有组件进行启动、暂停、重启等操作。



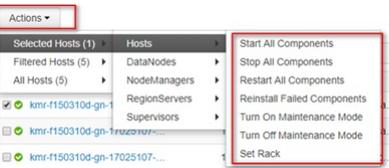
在Components下可对每一个组件进行启动、停止、重启及开启维护模式。



主机操作

选择所需操作的主机，点击上方Actions下拉菜单，选择Selected Hosts，选择Hosts；此后可对此主机组件进行操作：

操作	解释
Start All Components	启动所有组件
Stop All Components	停止所有组件
Restart All Components	重启所有组件
Reinstall Failed Components	重新安装所有失败组件
Turn on Maintenance Mode	开启维护模式
Turn off Maintenance Mode	关闭维护模式
Set Rack	设置Rack



注：

设置维护模式的四个一般场景为：执行维护，测试配置修改，彻底删除一个服务，处理警报。

要在一部主机上执行硬件或操作系统维护

执行维护时，要能够做如下操作：

- 阻止这部主机上所有组件生产警报
- 能够停止、启动、以及重启主机上的每一个组件
- 阻止该主机 host-level 或 service-level 的 starting, stopping, 或 restarting 组件批操作

为了达成这些目标，显示设置主机的维护模式，将这部主机上所有的组件隐式地设置为维护模式。

要测试一个服务配置的修改。应该停止、启动、以及重启服务来测试重启是否激活了配置的变化。

要测试配置信息的变化，要确保如下条件：

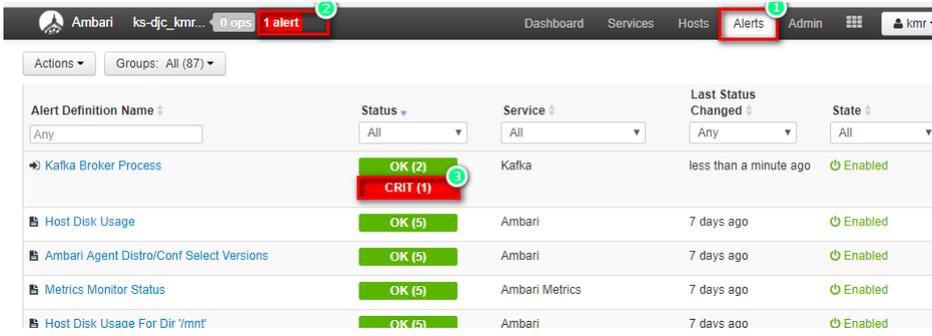
- 这个服务上没有任何组件生成警报
- 这个服务上没有 host-level 或 service-level 的批操作启动、停止、或 重启组件

为了达成这些目标，显示设置服务维护模式。将一个服务设置为维护模式隐式地为该服务的所有组件打开维护模式。

监控管理/Alerts

警告管理展示了集群的全部监控项，请注意查看 “WARNING” 和 “CRITICAL” 状态的警告项。

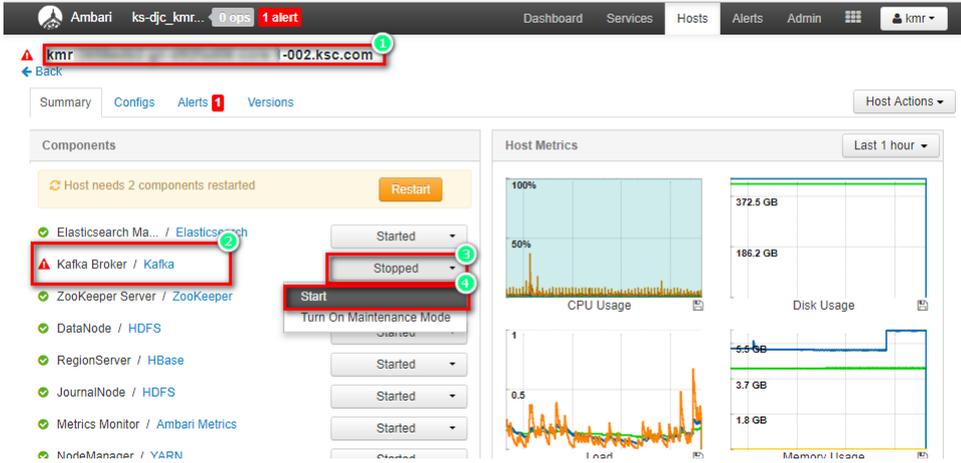
可单击Alerts，看各项状态，可以直接点击集群名称后的红色alert按钮，进行相应报警查看。



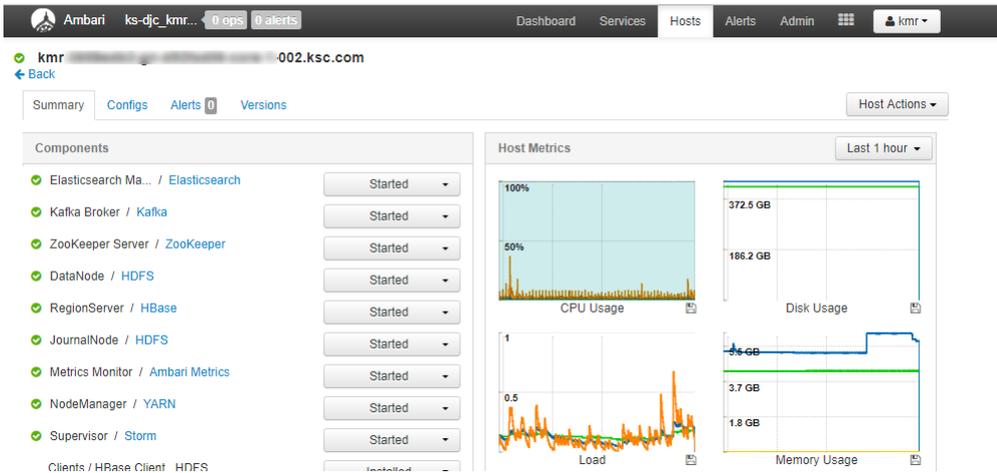
The screenshot shows the Ambari Alerts page. At the top, there are navigation tabs: Dashboard, Services, Hosts, Alerts (highlighted with a red box and a '1' icon), and Admin. Below the navigation, there are filters for Alert Definition Name, Status (set to All), Service (set to All), Last Status Changed, and State. A table lists several alerts:

Alert Definition Name	Status	Service	Last Status Changed	State
Kafka Broker Process	OK (2) CRIT (1)	Kafka	less than a minute ago	Enabled
Host Disk Usage	OK (5)	Ambari	7 days ago	Enabled
Ambari Agent Distro/Conf Select Versions	OK (5)	Ambari	7 days ago	Enabled
Metrics Monitor Status	OK (5)	Ambari Metrics	7 days ago	Enabled
Host Disk Usage For Dir /mnt	OK (5)	Ambari	7 days ago	Enabled

找到报警主机，进入相应报警处，进行处理后报警消除。下例为Kafka Broker停止，启动后报警解除。



The screenshot shows the Ambari Host Alert detail page for host 'kmr-002.ksc.com'. The 'Alerts' tab is active, showing 1 alert. The 'Components' section lists several services with their status: Elasticsearch Master (Started), Kafka Broker (Stopped), ZooKeeper Server (Started), DataNode (Started), RegionServer (Started), JournalNode (Started), Metrics Monitor (Started), and NodeManager (Started). A 'Restart' button is visible. The 'Host Metrics' section shows graphs for CPU Usage, Disk Usage, Load, and Memory Usage. A red box highlights the 'Kafka Broker / Kafka' component, and a dropdown menu is open showing 'Start', 'Stopped', and 'Turn On Maintenance Mode' options.



The screenshot shows the Ambari Host Alert detail page for host 'kmr-002.ksc.com' after the Kafka Broker has been restarted. The 'Alerts' tab now shows 0 alerts. The 'Components' section shows that the Kafka Broker status is now 'Started'. The 'Host Metrics' section shows graphs for CPU Usage, Disk Usage, Load, and Memory Usage.

Admin/管理

单击Admin按钮，可查看相应组件及对应版本信息。

Service	Version	Status	Description
HDFS	2.7.3	Installed	Apache Hadoop Distributed File System
YARN	2.7.3	Installed	Apache Hadoop NextGen MapReduce (YARN)
MapReduce2	2.7.3	Installed	Apache Hadoop NextGen MapReduce (YARN)
Tez	0.7.0	Installed	Tez is the next generation Hadoop Query Processing framework written on top of YARN.
Hive	1.2.1000	Installed	Data warehouse system for ad-hoc queries & analysis of large datasets and table & storage management service
HBase	1.1.2	Installed	A Non-relational distributed database, plus Phoenix, a high performance SQL layer for low latency applications.
Pig	0.16.0	Installed	Scripting platform for analyzing large datasets
Sqoop	1.4.6	Add Service	Tool for transferring bulk data between Apache Hadoop and structured data stores such as relational databases
Oozie	4.2.0	Installed	System for workflow coordination and execution of Apache Hadoop jobs. This also includes the installation of the optional Oozie Web Console which relies on and will install the ExJS Library.
ZooKeeper	3.4.6	Installed	Centralized service which provides highly reliable distributed coordination
Falcon	0.10.0	Add Service	Data management and processing platform
Storm	1.1.0	Installed	Apache Hadoop Stream processing framework
Flume	1.5.2	Add Service	A distributed service for collecting, aggregating, and moving large amounts of streaming data into HDFS
Accumulo	1.7.0	Add Service	Reliable, scalable, high performance distributed key/value store

文件查看

可在Ambari中进行文件查看，新建文件夹，文件搜索及上传文件操作。

点击Admin右侧九宫格形状按钮，选择Files View，可在页面中查看HDFS中文件。

Name	Size	Last Modified	Owner	Group
apps	--	2018-07-10 13:33	hdfs	hdfs
ats	--	2018-07-04 20:13	yarn	hadoop
hdp	--	2018-07-04 20:13	hdfs	hdfs
livy2-recovery	--	2018-07-04 20:14	livy	hdfs
mapred	--	2018-07-04 20:13	mapred	hdfs

一次点击Select All, New Folder和Upload按钮，可选择全部文件（夹），创建新文件夹及上传文件操作；在下方还可以进行文件（夹）搜索操作。

Last Modified	Owner	Group	Permission
2018-07-05 15:56	hue	hdfs	drwxr-xr-x
2018-07-04 20:11	ambari-qa	hdfs	drwxrwx---
2018-07-04 20:11	hbase	hdfs	drwxr-xr-x

文件权限

HDFS文件权限格式为用户权限组权限其他用户权限，d表示目录，r表示读，w表示写，x表示执行。

例如下图的权限为：

Name	Size	Last Modified
apps	--	2018-08-08 12:

对于apps目录，hdfs用户有用rwx，即可以读写执行这个目录，hdfs用户组的权限为r-x，即只能读和执行而没有写的权限，而其他组的权限也为r-x，即有读和执行的权限而没有写的权限。

您也可以在此栏目中查看YARN Queue Manager, Hive及Tez相关信息。

Ambari去除Token认证

KMR集群为了保证Ambari 访问的安全性，会对访问进行token认证，token过期后的访问会被拒绝。

注意：如果删除token认证，请注意管理帐号及访问权限（在安全组中仅开放限定IP的访问权限）

如果需要删除token认证机制，需要进行如下操作：

- 1、登入到KMR集群master1节点中。
- 2、编辑ngnix.conf。

vim /opt/nginx/conf/nginx.conf 210行附近

```
location / {
    # secure_link $arg_kmr_token,$arg_kmr_exp;
    # secure_link_md5 4d145a527973363cf5f96da451727bf2$arg_kmr_exp;
    # if ( $secure_link = "" ) {
    #     return 402;
    # }
    # if ( $secure_link = "0" ) {
    #     return 405;
    # }
```

注释掉secure_link相关的几行，210-217行。

3、重新加载nginx配置，重启nginx服务。

```
service nginx reload
service nginx restart
```

使用Ambari监控集群

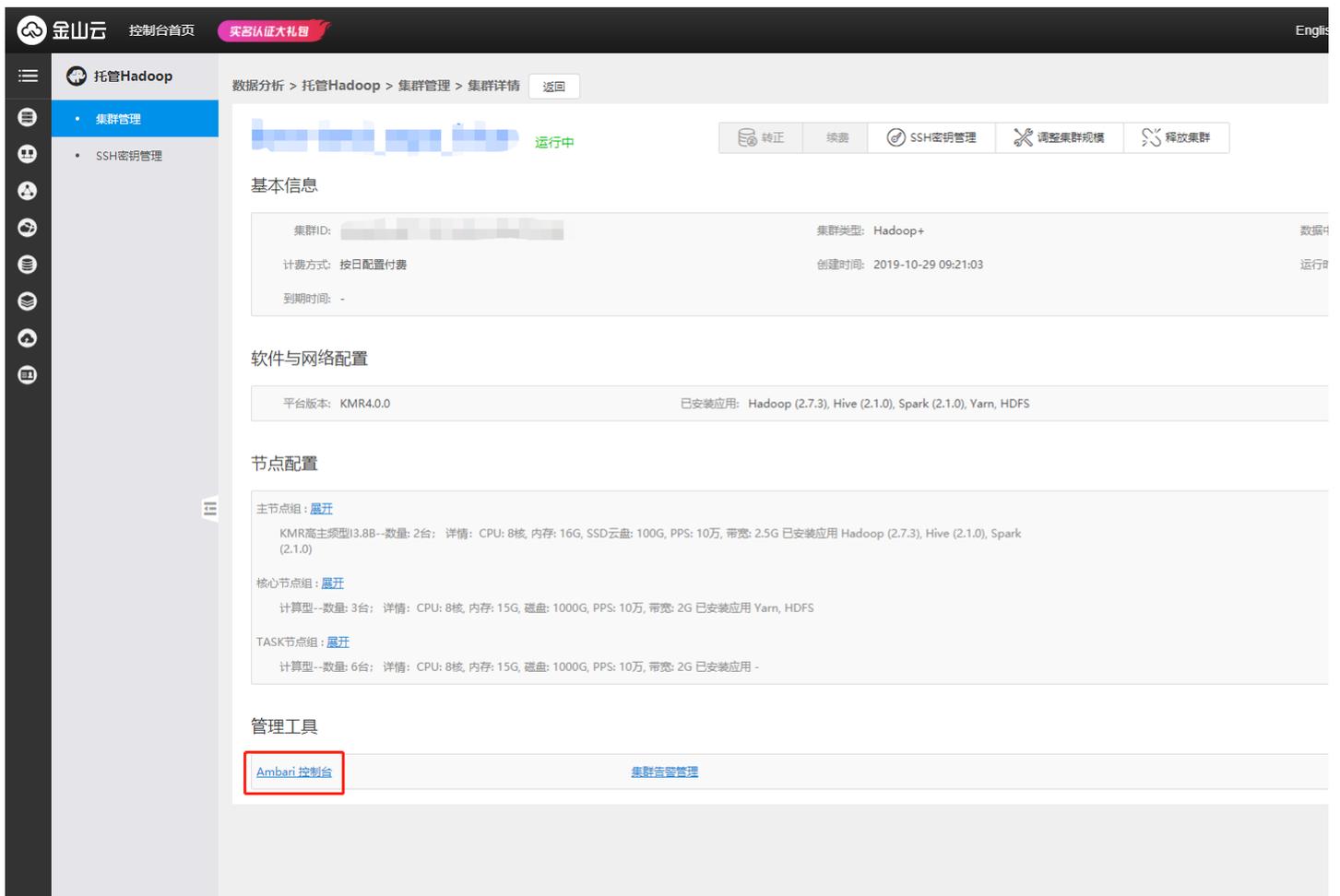
Apache Ambari是一种基于Web的工具，支持Hadoop集群的供应、监控和管理。Ambari支持HDFS、MapReduce、Hive、Spark、Storm、Hbase、Zookeeper等的集中管理，也是5个顶级Hadoop管理工具之一。

为防止因为误操作引起的服务不稳定，KMR对Ambari对了一定的权限控制。如需Ambari的管理权限，请联系售后支持人员。

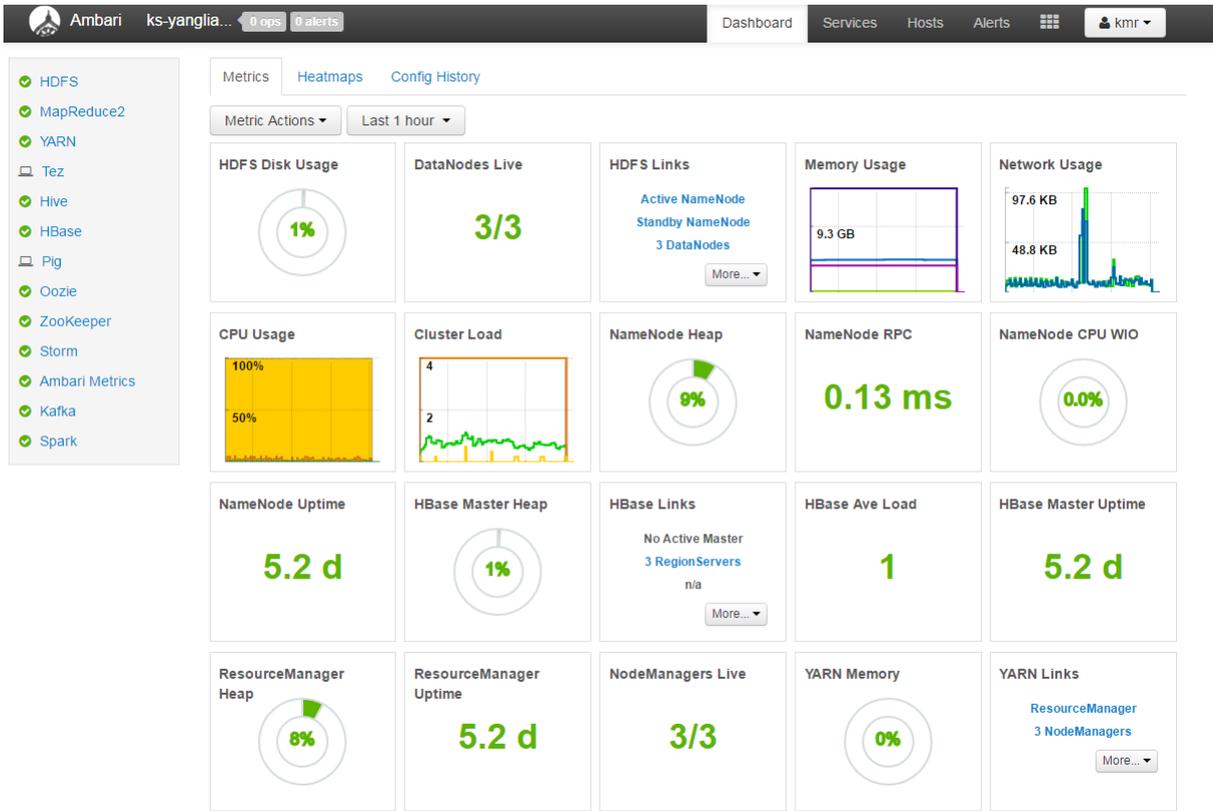
更多资料参见 [Apache Ambari](#) 官网。

准备

您已成功创建一个常驻集群，登录[KMR控制台](#)，进入集群详情页面，通过管理工具进入Ambari控制台。



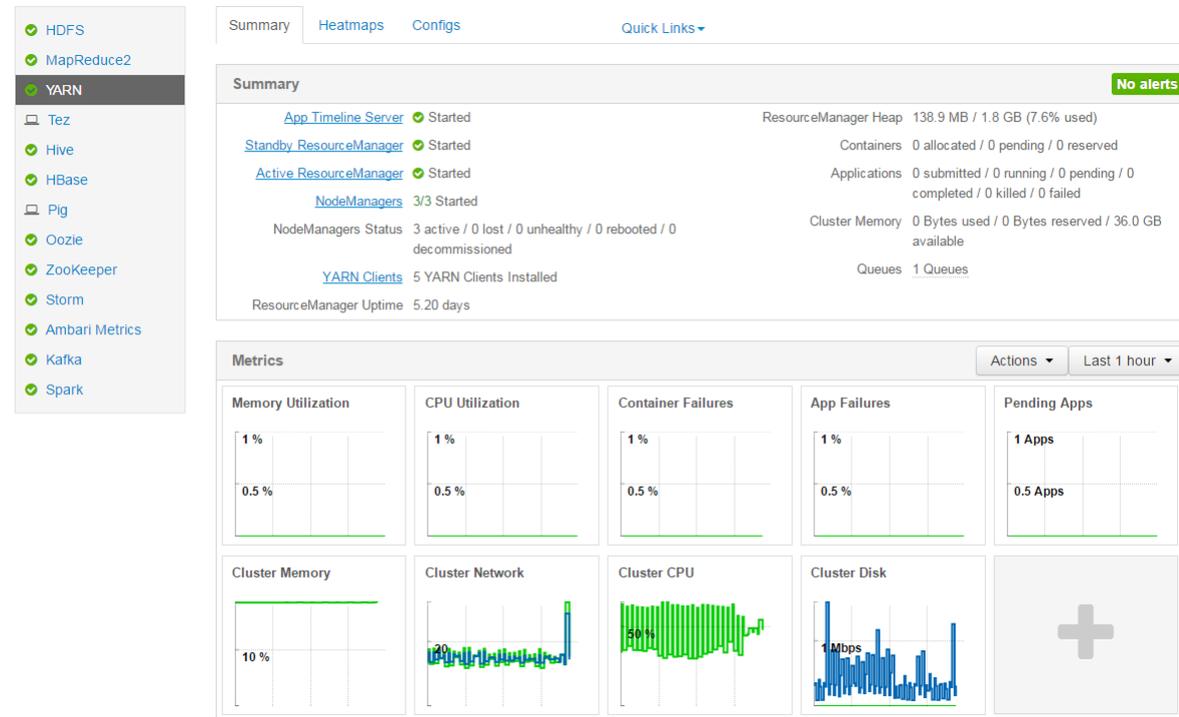
进入Ambari Dashboard页面展示了集群的整体情况，可以点击各个图表查看具体信息。



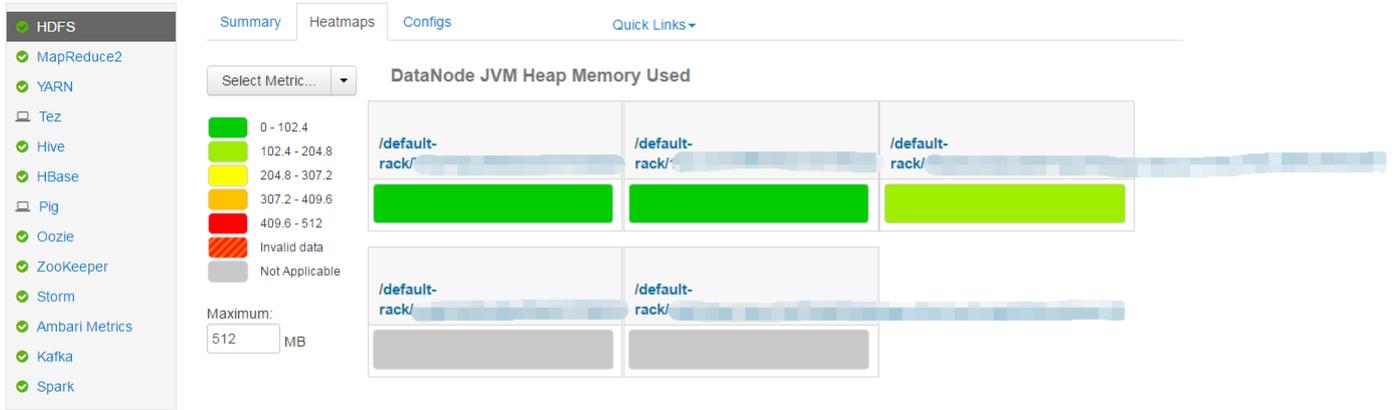
服务级别的监控管理

服务级别监控管理包括对HDFS, MapReduce, Storm, Spark等的管理, 通过左边导航点击对应的服务, 可以查看该服务的使用情况。

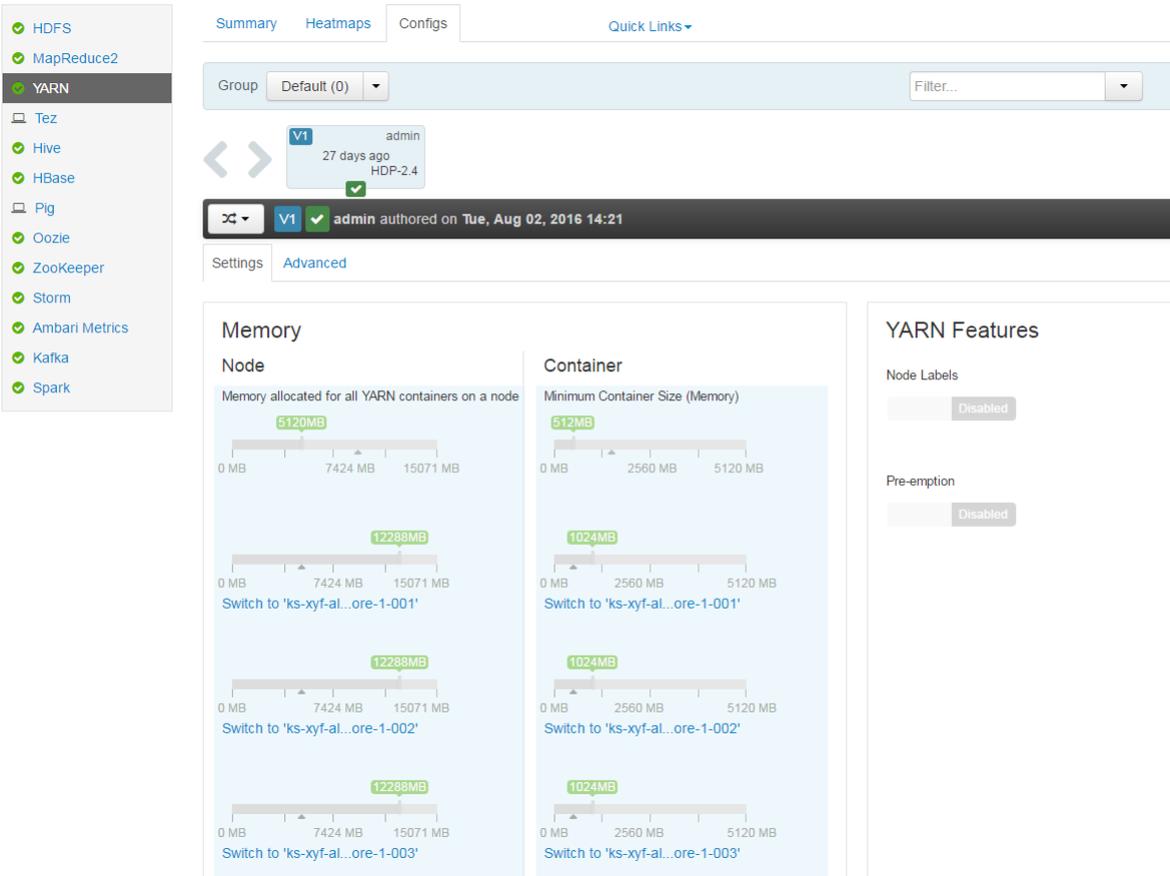
Summary展示了服务的基本情况和图标, 也可以通过+选择需要显示的指标。



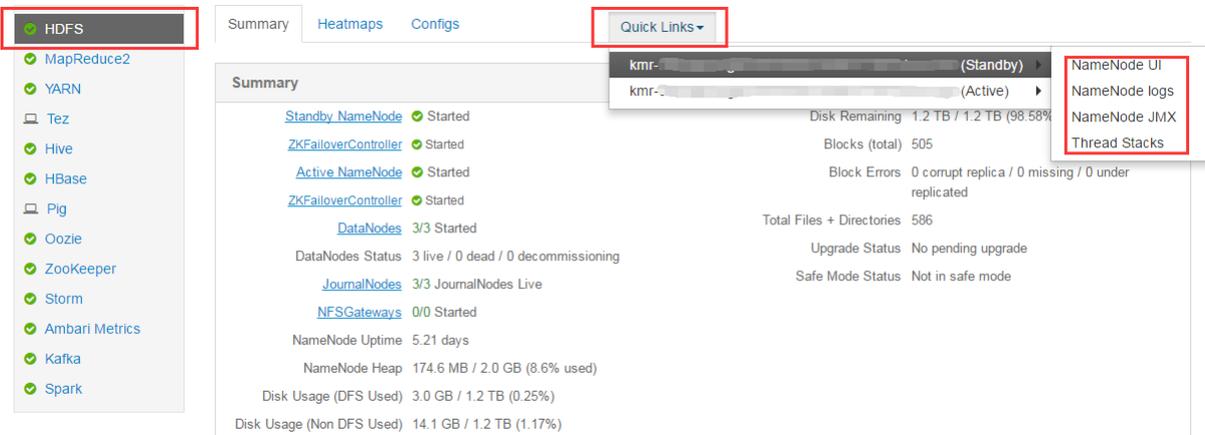
Heatmaps展示了服务各个角色使用磁盘IO、网卡IO、CPU以及JVM等性能状况。



Configs展示了服务的配置信息



点击服务详情页面的Quick Links可查看集群应用的原生监控工具。



主机（集群节点）级别的监控管理

点击上方导航栏的Hosts按钮，可以查看Ambari所管理的主机列表，通过输入过滤规则，可以筛选主机。

Filter: All (5)

Name	IP Address	Rack	Cores	RAM	Disk Usage	Load Avg	Versions	Components
kmr-...	...	/default-rack/7de77...	4 (4)	14.72GB		0.42		17 Components
kmr-...	...	/default-rack/1bbc1...	4 (4)	14.72GB		0.66		17 Components
kmr-...	...	/default-rack/80de8...	4 (4)	14.72GB		0.43		17 Components
kmr-...	...	/default-rack/cce82...	4 (4)	14.72GB		0.50		23 Components
kmr-...	...	/default-rack/c6d02...	4 (4)	14.72GB		1.03		22 Components

Show: 10 1 - 5 of 5

点击进入具体的机器，查看机器详情，对该机器的服务角色进行监控管理，可查看配置，警告和版本信息。

kmr-... Back

Summary **Configs** Alerts 0 Versions

Components

- Kafka Broker / Kafka
- ZooKeeper Server / ZooKeeper
- DataNode / HDFS
- RegionServer / HBase
- JournalNode / HDFS
- Metrics Monitor / Ambari Metrics
- NodeManager / YARN
- Supervisor / Storm

Clients / HBase Client, HDFS Client, Hive Client, MapReduce2 Client, Oozie Client, Spark Client, Tez Client, YARN Client, ZooKeeper Client

Host Metrics Last 1 hour

Summary

Hostname: kmr-...

IP Address: ...

Rack: /default-rack/...

OS: centos6 (x86_64)

Cores (CPU): 4 (4)

Disk: 11.9GB/418.56GB (2.84% used)

Memory: 14.72GB

Load Avg: 0.62

Heartbeat: a moment ago

Current Version: ...

警告管理

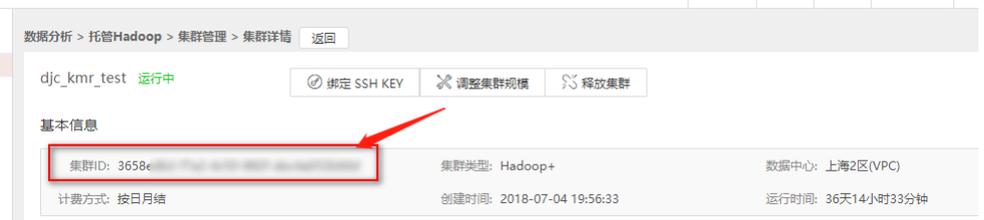
警告管理展示了集群的全部监控项，请注意查看“WARNING”和“CRITICAL”状态的警告项。

Groups: All (77) ▾

Alert Definition Name	Status	Service	Last Status Changed	State
Host Disk Usage	OK (5)	Ambari	5 days ago	Enabled
Metrics Monitor Status	OK (5)	Ambari Metrics	5 days ago	Enabled
Ambari Agent Heartbeat	OK (5)	Ambari	5 days ago	Enabled
DataNode Web UI	OK (3)	HDFS	5 days ago	Enabled
Supervisor Process	OK (3)	Storm	5 days ago	Enabled
JournalNode Web UI	OK (3)	HDFS	5 days ago	Enabled
DataNode Unmounted Data Dir	OK (3)	HDFS	5 days ago	Enabled
Kafka Broker Process	OK (3)	Kafka	5 days ago	Enabled
HBase RegionServer Process	OK (3)	HBase	5 days ago	Enabled
NodeManager Health	OK (3)	YARN	5 days ago	Enabled
DataNode Heap Usage	OK (3)	HDFS	5 days ago	Enabled
ZooKeeper Server Process	OK (3)	ZooKeeper	5 days ago	Enabled
DataNode Process	OK (3)	HDFS	5 days ago	Enabled
DataNode Storage	OK (3)	HDFS	5 days ago	Enabled
NodeManager Web UI	OK (3)	YARN	5 days ago	Enabled
HiveServer2 Process	OK (2)	Hive	5 days ago	Enabled
NameNode RPC Latency	OK (2)	HDFS	5 days ago	Enabled
ZooKeeper Failover Controller Process	OK (2)	HDFS	5 days ago	Enabled
Percent Supervisors Available	OK (2)	Storm	5 days ago	Enabled
DataNode Health Summary	OK (2)	HDFS	5 days ago	Enabled
WebHCat Server Status	OK (2)	Hive	5 days ago	Enabled

使用Ambari Admin账户安装组件

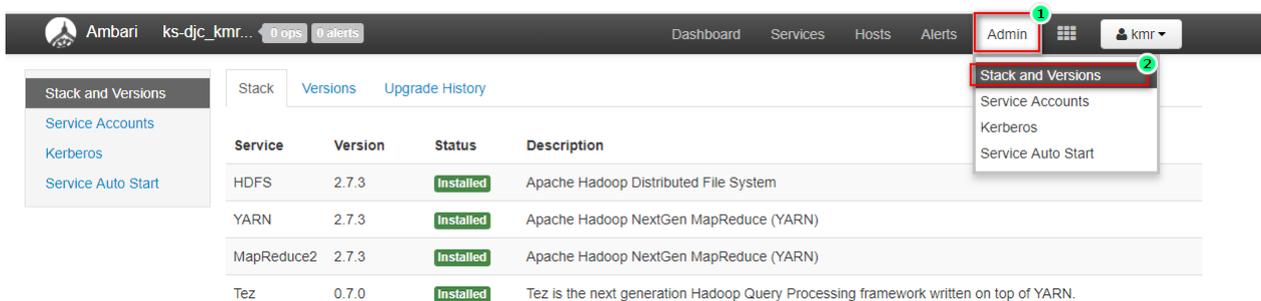
KMR集群中的默认Ambari用户为kmr用户，kmr用户只有组件配置管理，组件启停的权限，要进行组件的安装必须以admin账户登录；使用admin账户登录并安装组件操作如下：



1. 联系金山云售后索取admin账户密码（需提供集群id）
2. 进入KMR控制台>集群详情页面下部管理工具中点击Ambari控制台



3. 手动将 url 中的 "&kmr_auto_login=1" 进行删除后，重新加载页面；否则kmr账号会自动登录以致不能退出
4. 在 Ambari 右上方的账户管理中选择 sign out，使用 admin 账户进行登录
5. 在顶部Admin中选择Stack and Versions



6. 在Stack选项下找到所需安装组件，点击Add service

Service	Version	Status	Description
HDFS	2.7.3	Installed	Apache Hadoop Dis
YARN	2.7.3	Installed	Apache Hadoop Ne
MapReduce2	2.7.3	Installed	Apache Hadoop Ne
Tez	0.7.0	Installed	Tez is the next gene
Hive	1.2.1000	Installed	Data warehouse sy: service
HBase	1.1.2	Installed	A Non-relational dis applications.
Pig	0.16.0	Installed	Scripting platform fc
Sqoop	1.4.6	Add Service	Tool for transferring databases
Oozie	4.2.0	Installed	System for workflow of the optional Oozie
ZooKeeper	3.4.6	Installed	Centralized service

7. 在choose Services中可以看到刚才所选择安装组件已自动勾选，此时也可以勾选其他组件一同进行安装；选择完毕后，选择最下方的next

Add Service Wizard

ADD SERVICE WIZARD

- Choose Services
- Assign Masters
- Assign Slaves and Clients
- Customize Services
- Configure Identities
- Review
- Install, Start and Test
- Summary

Choose Services

Choose which services you want to install on your cluster.

Service	Version	Description
<input checked="" type="checkbox"/> HDFS	2.7.3	Apache Hadoop Distributed File System
<input checked="" type="checkbox"/> YARN + MapReduce2	2.7.3	Apache Hadoop NextGen MapReduce (YARN)
<input checked="" type="checkbox"/> Tez	0.7.0	Tez is the next generation Hadoop Query Processing framework written on top of YARN.
<input checked="" type="checkbox"/> Hive	1.2.1000	Data warehouse system for ad-hoc queries & analysis of large datasets and table & storage management service
<input checked="" type="checkbox"/> HBase	1.1.2	A Non-relational distributed database, plus Phoenix, a high performance SQL layer for low latency applications.
<input checked="" type="checkbox"/> Pig	0.16.0	Scripting platform for analyzing large datasets
<input checked="" type="checkbox"/> Sqoop	1.4.6	Tool for transferring bulk data between Apache Hadoop and structured data stores such as relational databases

<input checked="" type="checkbox"/> Slider	0.92.0	A framework for deploying, managing and monitoring existing distributed applications on YARN.
--	--------	---

[Next →](#)

8. 在Assign Slaves and Clients中Host下放上鼠标可以看到主机信息；在列表最右方选择需要进行安装的主机，点击next进行

Add Service Wizard

ADD SERVICE WIZARD

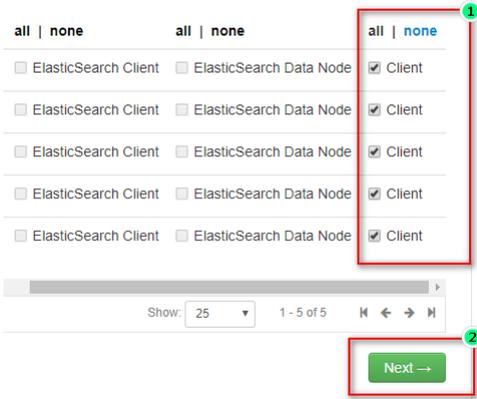
- Choose Services
- Assign Masters
- Assign Slaves and Clients
- Customize Services
- Configure Identities
- Review
- Install, Start and Test
- Summary

Assign Slaves and Clients

Assign slave and client components to hosts you want to run them on. Hosts that are assigned master components are shown with *.

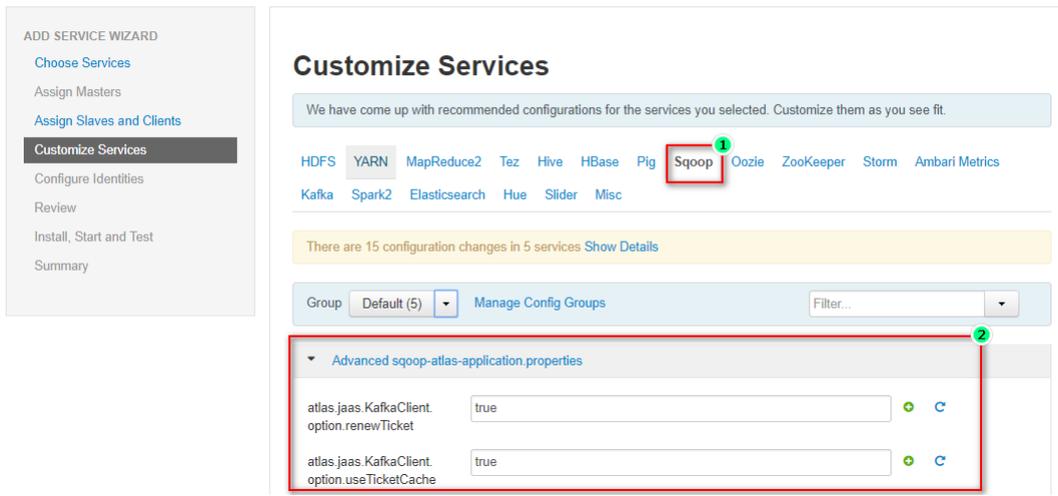
"Client" will install Sqoop Client

| Host | all none |
|-----------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| kmr-3658edb2-gn-d92fad0...* | <input type="checkbox"/> |
| kmr-3658edb2-gn-d92fad0...* | <input type="checkbox"/> |
| kmr-3658edb2-gn-d92fad0...* | <input type="checkbox"/> |
| kmr-3658edb2-gn-d92fad0...* | <input type="checkbox"/> |
| kmr-3658edb2-gn-d92fad0...* | <input type="checkbox"/> |



9. 在Customize Services中选择所需安装组件，在下方进行文件配置；配置完成后点击Next

Add Service Wizard



10. 点击next后，系统默认推荐配置改变值，在左侧选择所需改变的配置后，点击OK

Dependent Configurations

Recommended Changes

Based on your configuration changes, Ambari is recommending the following dependent configuration changes. Ambari will update all checked configuration changes to the Recommended Value. Uncheck any configuration to retain the Current Value.

<input checked="" type="checkbox"/>	Property	Service	Config Group	File Name	Current Value	Recommended Value
<input checked="" type="checkbox"/>	hadoop.proxyuser.hcat.hosts	HDFS	Default	core-site	*	kmr-3658edb2-gn-d92fad06-master-1-001.ksc.com, kmr-3658edb2-gn-d92fad06-master-2-001.ksc.com
<input checked="" type="checkbox"/>	hadoop.proxyuser.hive.hosts	HDFS	Default	core-site	*	kmr-3658edb2-gn-d92fad06-master-1-001.ksc.com, kmr-3658edb2-gn-d92fad06-master-2-001.ksc.com
<input checked="" type="checkbox"/>	hadoop.proxyuser.oozie.hosts	HDFS	Default	core-site	*	kmr-3658edb2-gn-d92fad06-master-1-001.ksc.com
<input checked="" type="checkbox"/>	hadoop.proxyuser.root.hosts	HDFS	Default	core-site	*	kmr-3658edb2-gn-d92fad06-master-1-001.ksc.com
<input checked="" type="checkbox"/>	dfs.encryption.key.provider.uri	HDFS	Default	hdfs-site		Property removed
<input checked="" type="checkbox"/>	hadoop.security.key.provider.path	HDFS	Default	core-site		

11. 在Configure Identities中进行确认

12. 在Review中再次确认信息，点击Deploy

Add Service Wizard

Configure identities

- Review
- Install, Start and Test
- Summary

Cluster Name : ks-djc_kmr_test

Total Hosts : 5 (0 new)

Repositories:

- debian7 (HDP-2.6): http://public-repo-1.hortonworks.com/HDP/debian7/2.x/updates/2.6.5.0
- debian7 (HDP-UTILS-1.1.0.21): http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.21/repos/debian6
- redhat-ppc7 (HDP-2.6): http://public-repo-1.hortonworks.com/HDP/centos7-ppc/2.x/updates/2.6.5.0
- redhat-ppc7 (HDP-UTILS-1.1.0.21): http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.21/repos/ppc64le
- redhat6 (HDP-2.6): http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.6.5.0
- redhat6 (HDP-UTILS-1.1.0.21): http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.21/repos/centos6
- redhat7 (HDP-2.6): http://kmr-sh-hdp.ks3-cn-shanghai-internal.ksyun.com/HDP/centos7/2.x/updates/2.6.1.0

← Back Print Deploy →

13. 之后系统进行组件安装，启动与测试；需等待一段时间（约5分钟，不同集群情况可能不同）；在Statue为100%后，点击Next

Add Service Wizard

ADD SERVICE WIZARD

- Choose Services
- Assign Masters
- Assign Slaves and Clients
- Customize Services
- Configure Identities
- Review
- Install, Start and Test
- Summary

Install, Start and Test

Please wait while the selected services are installed and started.

12 % overall

Host	Status	Message
kmr-3658ec...	12%	Installing Sqoop Client
kmr-3658ec...	12%	Installing Sqoop Client
kmr-3658ec...	12%	Installing Sqoop Client
kmr-3658ec...	12%	Installing Sqoop Client
kmr-3658ec...	12%	Installing Sqoop Client

5 of 5 hosts showing - Show All Show: 25 1 - 5 of 5

Next →

14. 安装完成后出现以下界面；注意此时应重启所有服务新安装组件才能生效；此时点击Complete

Summary

Important: You may also need to restart other services for the newly added services to function properly (for example, HDFS and YARN/MapReduce need to be restarted after adding Oozie). After closing this wizard, please restart all services that have the restart indicator next to the service name.

Here is the summary of the install process.

The cluster consists of 5 hosts
Installed and started services successfully on 5 new hosts
Install and start completed in 1 minutes and 31 seconds

Complete →

15. 点击complete后，选择Services→Service Actions→Restart All进行所有组件的重启；至此，新安装组件完成

The screenshot shows the Ambari dashboard for cluster 'ks-djc_kmr...'. The 'Services' tab is active. A yellow banner at the top indicates 'Restart Required: 15 Components on 5 Hosts'. The 'Service Actions' dropdown menu is open, and the 'Restart All' option is selected. The dashboard also shows a list of services on the left, including HDFS, YARN, MapReduce2, Tez, Hive, HBase, Pig, Sqoop, Oozie, ZooKeeper, and Storm.

Ambari告警信息

告警级别

告警级别	解释	说明
OK	ok	集群运行良好
WARNING	警告	集群指标到了一定阈值，需关注
CRITICAL	危险	集群的状态可能存在问题，需进行一定的处理
UNKNOWN	未知	状态未知
NONE	无	无

告警 (alert) 类型

类型	用途	告警级别	阈值是否可配置	单位
PORT	用来监测机器上的一个端口是否可用	OK, WARN, CRIT 是		秒
METRIC	用来监测 Metric 相关的配置属性	OK, WARN, CRIT 是		变量
AGGREGATE	用于收集其他某些 Alert 的状态	OK, WARN, CRIT 是		百分比
WEB	用于监测一个 WEB UI (URL) 地址是否可用	OK, WARN, CRIT 否		无
SCRIPT	Alert 的监测逻辑由一个自定义的 python 脚本执行	OK, WARN, CRIT 否		无

告警 (alert) 类型

报警状态如下：现对其进行解释说明

Instances

Service	Host	Status	24-Hour	Response	
Ranger	kmr-3658edb2-gn-d92fad06-core-1-001.ksc.com	CRIT	for about a day	0	Connection failed to http://kmr-3658edb2-gn-d92fad06-core-1-001.ksc...

Show: 10 1 - 1 of 1

字段	解释	备注
Service	服务，在其中可选具体组件以查看其告警状态	
Host	主机id，显示所告警的虚拟机id	
Status	状态，如告警级别中所示，总分为5种	
24-Hour	告警时长	
Response	响应，报警具体内容	点击可展示具体所告警全部内容

Ambari 告警description翻译

注：WARNING 用警告表示；CRITICAL 用危险表示

HDFS说明

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL) 默认值
HDFS Storage Capacity Usage(Weekly)	This service-level alert is triggered if the increase in storage capacity usage deviation has grown beyond the specified threshold within a week period. 一周内存增量偏移值超过设定阈值则触发此服务级告警。	20%
DataNode Unmounted DataDir	This host-level alert is triggered if one of the data directories on a host was previously on a mount point and became unmounted. If the mount history file does not exist, then report an error if a host has one or more mounted data directories as well as one or more unmounted data directories on the root partition. This may indicate that a data directory is writing to the root partition, which is undesirable. 主机上有数据目录在原先挂载点上且未卸载，则触发此主机级告警。若安装历史文件不存在，主机有一个或这个挂载点或未挂载点数据目录在根分区上则报错。这意味着数据目录正在写入根分区，这是不可取的。	2分钟
JournalNode Web UI	This host-level alert is triggered if the JournalNode Web UI is unreachable. 不能访问 JournalNode Web UI 时触发此主机级告警。	Connection failed to (1) (13)
DataNode Process	This host-level alert is triggered if the individual DataNode processes cannot be established to be up and listening on the network. 不能启动单个DataNode进程及在网络上监听单个DataNode进程时触发此主机级告警。	5
DataNode Web UI	This host-level alert is triggered if the DataNode Web UI is unreachable. 不能访问DataNode Web UI 时触发此主机级告警。	Connection failed to (1) (13)
DataNode Storage	This host-level alert is triggered if storage capacity is full on the DataNode. It checks the DataNode JMX Servlet for the Capacity and Remaining properties. The threshold values are in percent. DataNode上存储空间满时触发此主机级告警。会检查DataNode JMX服务上已存储和可存储容量。阈值以百分比形式展示。	80%
DataNode Heap Usage	This host-level alert is triggered if heap usage goes past thresholds on the DataNode. It checks the DataNode JMXServlet for the MemHeapUsedM and MemHeapMaxM properties. The threshold values are in percent. DataNode上堆使用情况超过设定阈值时触发此主机级告警。会检查DataNode JMX服务中已使用堆及堆最大量情况。阈值以百分比形式展示。	90%
HDFS Pending Deletion Blocks	This service-level alert is triggered if the number of blocks pending deletion in HDFS exceeds the configured warning and critical thresholds. It checks the NameNode JMX Servlet for the PendingDeletionBlock property. HDFS中待删除块的数量超过所配置的警告和告警阈值，则触发此服务级告警。会检查NameNode JMX 中的挂起的需删除块的数量。	100000
NameNode Client RPC Queue Latency (Daily)	This service-level alert is triggered if the deviation of RPC latency on client port has grown beyond the specified threshold within a day period. 在一天中客户端接口的RPC延迟偏移量增长率超过设定阈值时触发此服务级告警。	200%
DataNode Health Summary	This service-level alert is triggered if there are unhealthy DataNodes. 有不健康的DataNode时触发此服务级告警。	1
HDFS Upgrade Finalized State	This service-level alert is triggered if HDFS is not in the finalized state. HDFS不在完成状态时触发此服务级告警。	1
NameNode Client RPC Processing Latency (Daily)	This service-level alert is triggered if the deviation of RPC latency on client port has grown beyond the specified threshold within a day period. 一天中客户端接口的RPC延迟增量增长率超过设定阈值时触发此服务级告警。	200%
NameNode Blocks Health	This service-level alert is triggered if there are unhealthy DataNodes 有不健康的DataNode时触发此服务级告警。	1

NameNode Web UI	NameNode Blocks Health NameNode块健康状态 This host-level alert is triggered if the NameNode Web UI is unreachable. 不能访问NameNode Web UI 时触发此主机级告警。	Connection failed to {1} ({3})
NameNode Heap Usage (Daily) NameNode堆使用 (每天)	This service-level alert is triggered if the NameNode heap usage deviation has grown beyond the specified threshold within a day period. 一天中NameNode堆使用增量增长率超过所设定阈值时触发此服务级告警。	50%
NameNode Last Checkpoint NameNode最后检查	This service-level alert will trigger if the last time that the NameNode performed a checkpoint was too long ago. It will also trigger if the number of uncommitted transactions is beyond a certain threshold. 距上次NameNode检查时间太长时触发此服务级告警。未提交事务超过某个阈值时也会触发此告警。	200%
NameNode RPC Latency NameNode RPC延迟	This host-level alert is triggered if the NameNode RPC latency exceeds the configured critical threshold. Typically an increase in the RPC processing time increases the RPC queue length, causing the average queue wait time to increase for NameNode operations. The threshold values are in milliseconds. NameNode RPC延迟超过所设定阈值时触发此主机级告警。比较典型的是RPC进程时间增加会增加RPC队列长度,造成NameNode操作平均队列等待时间增长。阈值以毫秒级设计。	500
HDFS Storage Capacity Usage (Daily) HDFS存储使用量 (每天)	This service-level alert is triggered if the increase in storage capacity usage deviation has grown beyond the specified threshold within a day period. 一天中存储容量使用率增量超过特定阈值时触发此服务级告警。	50%
NameNode client RPC Queue Latency (Hourly) NameNode客户端RPC队列延迟 (每小时)	This service-level alert is triggered if the deviation of RPC queue latency on client port has grown beyond the specified threshold within an hour period. 一小时中客户端接口的RPC队列延迟增长率超过特定阈值时触发此服务级告警。	200%
HDFS Capacity Utilization HDFS容量使用	This service-level alert is triggered if the HDFS capacity utilization exceeds the configured warning and critical thresholds. It checks the NameNode JMX Servlet for the CapacityUsed and CapacityRemaining properties. The threshold values are in percent. HDFS容量使用超过设定警告和告警阈值时触发此服务级告警。会检查NameNode JMX中的容量使用和容量存留。阈值以百分比形式展示。	80%
NameNode Heap Usage (Weekly) NameNode 堆使用 (每周)	This service-level alert is triggered if the NameNode heap usage deviation has grown beyond the specified threshold within a week period. 一周中NameNode 堆使用偏移量增长率超过特定阈值时触发此服务级告警。	50%
NameNode Directory Status NameNode文档状态	This host-level alert is triggered if the NameNode NameDirStatuses metric (name=NameNodeInfo/NameDirStatuses) reports a failed directory. The threshold values are in the number of directories that are not healthy. NameNode 名字文档中有失效文档时触发此主机级告警。阈值是不健康文档数量。	1
NameNode Client RPC Latency (Hourly) NameNode客户端RPC延迟 (每小时)	This service-level alert is triggered if the deviation of RPC latency on client port has grown beyond the specified threshold within an hour period. 一小时中客户端接口RPC延迟偏移量增长率超过特定阈值时触发此服务级告警。	200%
NameNode Host CPU Utilization NameNode主机CPU使用	This host-level alert is triggered if CPU utilization of the NameNode exceeds certain warning and critical thresholds. It checks the NameNode JMX Servlet for the SystemCPULoad property. The threshold values are in percent. NameNode的CPU使用超过警告和告警阈值时触发此主机级告警。会检查NameNode JMX中的系统CPU载量。阈值以百分比形式展示。	250%
Zookeeper Failover Controller Process zookeeper故障转移控制进程	This host-level alert is triggered if the ZooKeeper Failover Controller process cannot be confirmed to be up and listening on the network. zookeeper故障转移控制进程不能被确认已启动或被网络监听时触发此主机级告警。	6
NameNode High Availability Health NameNode高可用健康状态	This service-level alert is triggered if either the Active NameNode or Standby NameNode are not running. 主节点或备用节点的NameNode都不运行时触发此服务级告警。	1
Percent DataNodes Available DataNode可用百分比	This alert is triggered if the number of down DataNodes in the cluster is greater than the configured critical threshold. It aggregates the results of DataNode process checks. 集群中挂掉DataNode数目超过所设定阈值时触发此告警。这会聚合DataNode进程检查结果。	30%
Percent DataNodes With Available Space DataNode中可用空间百分比	This service-level alert is triggered if the storage on a certain percentage of DataNodes exceeds either the warning or critical threshold values. 一定比例的DataNode中存储量超过警告或告警阈值时触发此服务级告警。	30%
Percent JournalNodes Available JournalNode可用百分比	This alert is triggered if the number of down JournalNodes in the cluster is greater than the configured critical threshold. It aggregates the results of JournalNode process checks. 集群中所挂掉JournalNode数目多于所设定阈值时触发此告警。会聚合JournalNode进程检查结果。	50%
NameNode Service RPC Processing Latency (Hourly) NameNode服务RPC延迟 (每小时)	This service-level alert is triggered if the deviation of RPC latency on datanode port has grown beyond the specified threshold within an hour period. 一小时中DataNode接口总的RPC延迟偏移量超过所设定阈值时触发此服务级告警。	200%
NameNode Service RPC Queue Latency (Hourly) NameNode服务RPC队列延迟 (每小时)	This service-level alert is triggered if the deviation of RPC queue latency on datanode port has grown beyond the specified threshold within an hour period. 一小时中datanote接口中RPC队列延迟偏移增长率超过设定阈值时触发此服务级告警。	200%
NameNode Service RPC Queue Latency (Daily) NameNode服务RPC队列延迟 (每天)	This service-level alert is triggered if the deviation of RPC latency on datanode port has grown beyond the specified threshold within a day period. 一天中datanote接口中RPC延迟偏移增长率超过设定阈值时触发此服务级告警。	200%
Secondary NameNode Process NameNode副本进程	This host-level alert is triggered if the Secondary NameNode process cannot be confirmed to be up and listening on the network. NameNode副本进程不能被确认已启动或被网络监听时触发此主机级告警。	Connection failed to {1} ({3})
NFS Gateway Process NFS网关进程	This host-level alert is triggered if the NFS Gateway process cannot be confirmed to be up and listening on the network. NFS网关进程不能被确认已启动或被网络监听时触发此主机级告警。	5
NameNode Service RPC Processing Latency (Daily) NameNode服务RPC进程延迟 (每天)	This service-level alert is triggered if the deviation of RPC latency on datanode port has grown beyond the specified threshold within a day period. 一天中DataNode接口的RPC延迟偏移量增长率超过设定阈值时触发此主机级告警。	200%

YARN

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL) 默认值
NodeManager Web UI	This host-level alert is triggered if the NodeManager Web UI is unreachable. 不能访问 NodeManager Web UI时触发此主机级告警。	Connection failed to {1} ({3})

NodeManager Health	This host-level alert checks the node health property available from the NodeManager component.	1
NodeManager健康状态	此主机级告警检查NodeManager组件中的节点健康状态。	
ResourceManager Web UI	This host-level alert is triggered if the ResourceManager Web UI is unreachable.	Connection failed to {1} ({3})
ResourceManager Web UI	不能访问 ResourceManager Web UI时触发此主机级告警。	
ResourceManager CPU Utilization	This host-level alert is triggered if CPU utilization of the ResourceManager exceeds certain warning and critical thresholds. It checks the ResourceManager JMX Servlet for the SystemCPULoad property. The threshold values are in percent.	250%
ResourceManager CPU 使用情况	ResourceManager CPU 使用增长率超过警告及告警阈值时触发此主机级告警。会检查ResourceManager JMX的系统CPU负载能力。阈值以百分比形式展示。	
ResourceManager RPC Latency	This host-level alert is triggered if the ResourceManager operations RPC latency exceeds the configured critical threshold. Typically an increase in the RPC processing time increases the RPC queue length, causing the average queue wait time to increase for ResourceManager operations. The threshold values are in milliseconds.	5000
ResourceManager RPC 延迟	ResourceManager RPC 延迟超过设定告警阈值时触发此主机级告警。典型情况下增加RPC进程时间会增加RPC队列长度, 使ResourceManager操作的平均队列等待时间增加。此阈值为毫秒级。	
NodeManager Health Summary	This service-level alert is triggered if there are unhealthy NodeManagers	1
NodeManager健康状态	有不健康的NodeManager时触发此服务级高级。	
Percent NodeManagers Available	This alert is triggered if the number of down NodeManagers in the cluster is greater than the configured critical threshold. It aggregates the results of NodeManager process checks.	30%
可用NodeManager百分比	集群中挂掉的NodeManager数量超过所设定告警阈值时触发此告警。会聚合NodeManager进程检查结果。	
App Timeline Web UI	This host-level alert is triggered if the App Timeline Server Web UI is unreachable.	Connection failed to {1} ({3})
App Timeline Web UI	不能访问App Timeline Server Web UI时触发此主机级告警。	
Failed Apps Check	This service-level alert is triggered if failed yarn apps is beyond the specified threshold within a given time span.	2
失败的App检查	在给定时间内失败的yarn app数超过阈值时触发此服务级告警。	

MapReduce2

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL) 默认值
History Server Web UI	This host-level alert is triggered if the History Server Web UI is unreachable.	Connection failed to {1} ({3})
历史服务器Web UI	不能访问历史服务器Web UI时触发此主机级告警。	
History Server RPC Latency	This host-level alert is triggered if the History Server operations RPC latency exceeds the configured critical threshold. Typically an increase in the RPC processing time increases the RPC queue length, causing the average queue wait time to increase for operations. The threshold values are in milliseconds.	5000
历史服务器RPC延迟	历史服务器RPC延迟超过设定阈值时触发此主机级告警。一般增减RPC进程时间会增加RPC队列长度, 使操作的平均队列等待时间增加。阈值为毫秒级。	
History Server CPU Utilization	This host-level alert is triggered if the percent of CPU utilization on the History Server exceeds the configured critical threshold. The threshold values are in percent.	250%
历史服务器CPU使用情况	历史服务器的CPU使用百分比超过阈值时触发此主机级告警。阈值以百分比形式展示。	
History Server Process	This host-level alert is triggered if the History Server process cannot be established to be up and listening on the network.	5
利用服务器进程	历史服务器进程不能被启动或从网络监听时会触发此主机级告警。	

Hive

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL) 默认值
WebHcat Server Status	This host-level alert is triggered if the templeton server status is not healthy.	5
WebHcat 服务器状态	templeton 服务器状态不健康时触发此主机级告警。	
HiveServer2 Interactive Process	This host-level alert is triggered if the HiveServerInteractive cannot be determined to be up and responding to client requests.	60
HiveServer2 交互过程	Hive服务器交互不能确认已启动和响应客户端时会触发此主机级告警。	
Hive MetaStore Process	This host-level alert is triggered if the Hive Metastore process cannot be determined to be up and listening on the network.	60
Hive元数据过程	Hive元数据过程不能确定已启动和从网络进行监听时触发此主机级告警。	
LLAP Application	This alert is triggered if the LLAP Application cannot be determined to be up and responding to requests.	120
LLAP应用	LLAP应用不能确定已启动和响应客户端时触发此主机级告警。	
HiveServer2 Process	This host-level alert is triggered if the HiveServer cannot be determined to be up and responding to client requests.	60
HiveServer2 进程	Hive服务器不能确定已启动和响应客户端时触发此主机级告警。	

HBase

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL) 默认值
HBase RegionServer Process	This host-level alert is triggered if the RegionServer processes cannot be confirmed to be up and listening on the network for the configured critical threshold, given in seconds.	5
HBase RegionServer 进程	HBase RegionServer 进程不能确认已启动和在给定阈值 (秒级) 下从网络监听时触发此主机级告警。	
HBase Master Process	This alert is triggered if the HBase master processes cannot be confirmed to be up and listening on the network for the configured critical threshold, given in seconds.	5
HBase 主节点进程	HBase 主节点进程不能确认已启动和在给定阈值 (秒级) 下从网络监听时触发此主机级告警。	
Percent RegionServers Available	This service-level alert is triggered if the configured percentage of RegionServer processes cannot be determined to be up and listening on the network for the configured warning and critical thresholds. It aggregates the results of RegionServer process down checks.	30%
RegionServer可用百分比	所配置一定百分比的RegionServer进程不能确认已启动或从网络监听时触发此服务级告警。这会聚合RegionServer进程失败检查结果。	
HBase Master CPU Utilization	This host-level alert is triggered if CPU utilization of the HBase Master exceeds certain warning and critical thresholds. It checks the HBase Master JMX Servlet for the SystemCPULoad property. The threshold values are in percent.	250%
HBase主节点CPU使用情况	HBase主节点上CPU使用超过所设置的警告及告警阈值会触发此主机级告警。这会检查HBase 主节点JMX中的系统CPU负荷情况。阈值以百分比形式展示。	

Oozie

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL) 默认值
Oozie Server Web UI	This host-level alert is triggered if the Oozie server Web UI is unreachable.	Connection failed to {1} ({3})
Oozie Server Web UI	不能访问此Oozie服务器Web UI时触发此告警。	
Oozie Server Status	This host-level alert is triggered if the Oozie server cannot be determined to be up and responding to client requests.	1
Oozie Server状态	Oozie Server不能确定已启动和不响应客户端请求时触发此主机级告警。	

ZooKeeper

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL) 默认值
--------------------------------	------------------	-------------------

ZooKeeper Server Process ZooKeeper服务器进程	This host-level alert is triggered if the ZooKeeper server process cannot be determined to be up and listening on the network. ZooKeeper服务器进程不能确认已启动和从网络监听时触发此主机级告警。	5
Percent ZooKeeper Servers Available ZooKeeper服务器可用百分比	This alert is triggered if the number of down ZooKeeper servers in the cluster is greater than the configured critical threshold. It aggregates the results of ZooKeeper process checks. 集群中所挂掉ZooKeeper服务器数量大于所配置阈值时触发此告警。会对ZooKeeper进程检查结果进行聚合。	70%

Storm

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL)	默认值
Supervisor Process		5	
Percent Supervisors Available		30%	
Storm Web UI		Connection failed to {1} ({3})	
Nimbus Process		5	
DRPC Server Process		5	

Kafka

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL)	默认值
Kafka Broker Process Kafka Broker进程	This host-level alert is triggered if the Kafka Broker cannot be determined to be up. Kafka Broker进程不能确定是否已启动时触发此主机级告警。	5	

spark2

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL)	默认值
Spark2 Livy Server Spark2 Livy服务器	This host-level alert is triggered if the Livy2 Server cannot be determined to be up. Livy2 服务器不能确认已启动时触发此主机级告警。	60	
Spark2 History Server Spark2 历史服务器	This host-level alert is triggered if the Spark2 History Server cannot be determined to be up. Spark2 历史服务不能确认已启动时触发此主机级告警。	5	
Spark2 Thrift Server Spark2 Thrift 服务器	This host-level alert is triggered if the Spark2 Thrift Server cannot be determined to be up. Spark2 Thrift 服务不能确认已启动时触发此主机级告警。	60	

ElasticSearch

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL)	默认值
ElasticSearch Process Check ElasticSearch 进程检查	This host-level alert is triggered if the ElasticSearch Master cannot be determined to be up. ElasticSearch 主节点不能确定已启动时触发此主机级告警。	5	

Hue

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL)	默认值
Hue Web UI	This host-level alert is triggered if the Hue Web UI is unreachable. 不能访问Hue Web UI时触发此主机级告警。	5	

Ambari

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL)	默认值
Host Disk Usage 主机硬盘使用情况	This host-level alert is triggered if the amount of disk space used goes above specific thresholds. The default threshold values are 50% for WARNING and 80% for CRITICAL. 主机硬盘使用率超过阈值时触发此主机级告警。阈值默认为警告: 50%, 告警: 80%。	80%	
Ambari Agent Distro/conf Select Versions Ambari 客户端Distro/conf版本选择	This host-level alert is triggered if the distro selector such as hdp-select cannot calculate versions available on this host. This may indicate that /usr/\$stack/ directory has links/dirs that do not belong inside of it. 主机上没有选择distro版本 (如hdp选择) 时触发此主机级告警。这可能时因为在 /usr/\$stack/目录下含不属于其链接的目录。	5	
Host Disk Usage For Dir '/' 主机硬盘使用目录 '/'	This host-level alert is triggered if the amount of disk space used goes above specific thresholds. The default threshold values are 80% for WARNING and 90% for CRITICAL. 硬盘使用量超过阈值时触发此主机级告警。阈值默认为警告: 80%, 告警: 90%。	5.0E9 bytes	
Host Disk Usage For Dir '/mnt' 主机硬盘使用目录 '/mnt'	This host-level alert is triggered if the amount of disk space used goes above specific thresholds. The default threshold values are 80% for WARNING and 90% for CRITICAL. 硬盘使用量超过阈值时触发此主机级告警。阈值默认为警告: 80%, 告警: 90%。	5.0E9 bytes	
Ambari Agent Heartbeat Ambari客户端心跳	This alert is triggered if the server has lost contact with an agent. 服务器未收到客户端心跳时发出此告警。	2	
Ambari Server Alerts Ambari 服务器告警	This alert is triggered if the server detects that there are alerts which have not run in a timely manner. 如果服务器检测到没有及时运行的警报时触发此警报。	2	
Ambari Server Performance Ambari服务器性能	This alert is triggered if the Ambari Server detects that there is a potential performance problem with Ambari. This type of issue can arise for many reasons, but is typically attributed to slow database queries and host resource exhaustion. Ambari服务器检测Ambari有潜在运行问题时触发此告警。有很多因素都有可能导致这个问题, 但最常见的是由于数据库查询缓慢以及主机资源耗尽。	5000	
component Version 组件版本	This alert is triggered if the server detects that there is a problem with the expected and reported version of a component. The alert is suppressed automatically during an upgrade. 服务器检测到组件有版本问题时触发此告警。在组件升级时常会触发此告警。	5	

Ambari Metrics

告警定义名称 (Alert Definition Name)	描述 (Description)	危险 (CRITICAL)	默认值
Metrics Monitor Status Metrics 监控状态	This alert indicates the status of the Metrics Monitor process as determined by the monitor status script. 此告警指示监控器状态脚本所确定的Metrics监控进程状态。	1	
Metrics Collector Process Metrics收集器进程	This alert is triggered if the Metrics Collector cannot be confirmed to be up and listening on the configured port for number of seconds equal to threshold. Metrics收集器不能确认已启动或在数秒内监听到的次数没有达到所配置阈值次数时触发此告警。	5	
Metrics collector - HBase CPU Utilization Metrics收集器的HBase CPU使用	This host-level alert is triggered if CPU utilization of the Metrics Collector's HBase Master exceeds certain warning and critical thresholds. It checks the HBase Master JMX Servlet for the SystemCPULoad property. The threshold values are in percent. Metrics收集器的HBase主节点的CPU使用超过警告和告警阈值。回检查HBase主节点JMX的系统CPU加载情况。阈值以百分比形式展示。	250%	
Metrics Collector - Auto-Restart Status Metrics 收集器自动重启状态	This alert is triggered if the Metrics Collector has been restarted automatically too frequently in last one hour. By default, a Warning alert is triggered if restarted twice in one hour and a Critical alert is triggered if restarted 4 or more times in one hour. 在最后一小时中Metrics收集器自动重启太频繁会触发此告警。一小时中两次重启系统警告, 4次重启系统告警。		Metrics Collector has been auto-started {1} times{0}
Percent Metrics Monitors Available Metrics监控器可用百分比	This alert is triggered if a percentage of Metrics Monitor processes are not up and listening on the network for the configured warning and critical thresholds. 一定比例 (所配置的警告和告警阈值) 的Metrics 监控器未启动或从不能从网络监听到。	30%	
Metrics Collector - HBase Master Processes Metrics收集器的HBase主节点进程	This alert is triggered if the Metrics Collector's HBase master processes cannot be confirmed to be up and listening on the network for the configured critical threshold. 在给定的时间内 (秒) Metrics收集器的HBase主节点进程不能确认已启动或从网络监听到触发此报警。	5	

Grafana Web UI This host-level alert is triggered if the Grafana Web UI is unreachable. 不能访问Grafana Web UI时触发此告警。

5

去除Ambari页面的token认证

KMR集群为了保证Ambari链接访问的安全性，默认情况下无法直接访问ambari控制台，kmr集群会对访问进行token认证，只有带有有效token的session才能访问ambari控制台，token过期后的访问会被拒绝。

如果删除token认证，请注意管理帐号及访问权限（仅开放限定IP的访问权限）。

如果需要删除token认证机制，需要进行如下操作：

1. 登入到KMR集群master1节点中。
2. 编辑nginx.conf。

```
vim /opt/nginx/conf/nginx.conf
```

210行附近

```
...
location / {
    # secure_link $arg_kmr_token,$arg_kmr_exp;
    # secure_link_md5 4d145a527973363cf5f96da451727bf2$arg_kmr_exp;
    # if ( $secure_link = "" ) {
    #     return 402;
    # }
    # if ( $secure_link = "0" ) {
    #     return 405;
    # }
...

```

注释掉secure_link相关的几行，210-217行。

3. 重新加载nginx配置，重启nginx服务。

```
service nginx reload
```

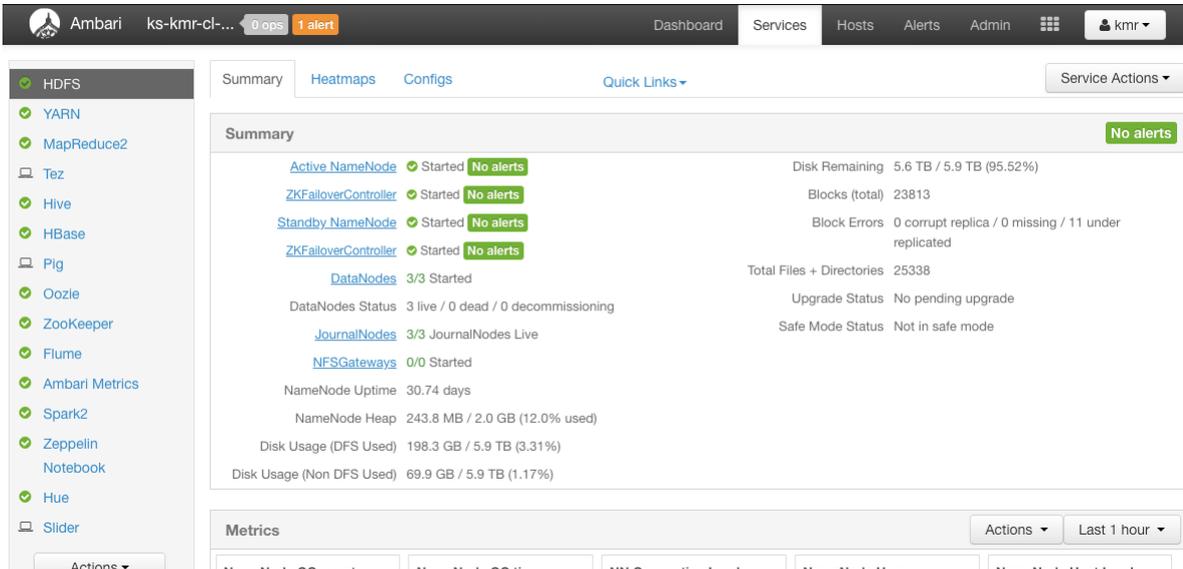
```
service nginx restart
```

Ambari集群配置的实践指南

KMR使用Ambari Web UI來管理主机、服务、告警、设置等。

管理集群配置

通过Ambari可以设定特定组件的配置。如果要调整组件的配置，可以进入 service 标签，或者点击左右的具体组件，点击 config 进行配置。



修改 NameNode Java 堆栈大小

NameNode Java 堆栈大小取决于许多因素，例如集群的负载、文件数和分块数。1GB的预设大小适用于大部分的集群，但是某些工作负载需要更多或更少的内存。

修改 NameNode Java 堆栈大小：

1. 点击 service 标签中的 HDFS，进入 config 标签。



2. 找到 NameNode Java heap size 设置。您也可以使用 filter 来搜索特定配置。选取右侧的画笔。



3. 在文本框中输入新的值，回车保存修改。

NameNode Java heap size

 MB

4. 点击 config 顶部的绿色的 save 保存修改。



Hive配置优化实践指南

下面说明一下Hive性能调优的配置选项。

1. 要修改Hive组件的参数，点击 **service** 标签中的 **Hive**。
2. 进入 **config** 标签。

设定Hive集群引擎

Hive提供两个执行引擎：MapReduce 和 Tez。Tez 比 MapReduce 更快。KMR集群使用Tez作为预设的执行引擎。更改引擎：

1. 在 **Hive** 的 **config** 标签中，**filter** 中筛选 **execution engine**。



2. 最佳的预设值是 **Tez**。

Optimization

Tez

Execution Engine

TEZ

MapReduce

TEZ

优化分组参数

Hadoop会将单个文件拆分（对应）为多个文件，并且并行处理产生的文件。对应程序的数目取决于分割的数目。下面两个组件的参数会影响Tez执行引擎的分割数量：

tez.grouping.min-size: 分组分割大小下限，预设值是 16 MB (16777216 位元组)。 tez.grouping.max-size: 分组分割大小上限，预设值为 1 GB (1,073,741,824 位元组)。

根据效能的经验法则，降低这两个参数可以改善延迟扩大吞吐。

例如，要设定128MB对应四个程序任务，可以将这两个参数都设置为32MB。

要修改限制参数，进入Tez服务的 **config** 标签。展开 **general** 选项，找到 **tez.grouping.max-size** 和 **tez.grouping.min-size** 参数。

将这两个参数设置为 33,554,432 位元组 (32 MB)。

tez.grouping.max-size 33554432

tez.grouping.min-size 33554432

这些变更会影响服务器的所有Tez任务。若要得到最佳效果，请适当调节参数值。

调优reducer

ORC 和 Snappy 均可以达到高性能。不过，Hive预设的reducer数量可能太少，因而导致瓶颈。

举例来说，如果有50GB的数据。数据是ORC格式，经过Snappy压缩处理后为1GB。Hive估计需要的reducer数量为：(对应程序分组的数据大小) / hive.exec.reducers.bytes.per.reducer

根据预设设定，这个例子需要4个reducer。

hive.exec.reducers.bytes.per.reducer 参数指定每个reducer处理的数据大小。预设值为64MB。降低这个值会增加并行度，可能会改善性能。过度降低也会产生过多的reducer，严重的可能会影响性能。此参数根据你的数据需求、压缩设置和其他环境设置决定。

1. 要修改参数，请进入 **Hive** 标签下的 **config** 标签，**filter** 搜索 **reducer** 参数。

Data per Reducer



2. 点击**编辑**将值修改为128MB (134,217,728)，然后回车保存。

Data per Reducer

 B

假设输入大小为 1024 MB，每个化简器的数据为 128 MB，则有 8 个化简器 (1024/128)。

为“Data per Reducer”参数提供错误的值可能导致生成大量的reducer，从而对查询性能产生负面影响。若要限制reducer的最大数目，请将 `hive.exec.reducers.max` 设置为适当的值。默认值为 1009。

启用并行执行

一个 Hive 查询是在一个或多个阶段中执行的。如果可以并行运行各个独立阶段，则会提高查询性能。

若要启用并行查询执行，请导航到 Hive 的「config」选项卡并搜索 `hive.exec.parallel` 属性。默认值为 `false`。将该值更改为 `true`，然后按 Enter 保存该值。

若要限制并行运行的作业数，请修改 `hive.exec.parallel.thread.number` 属性。默认值为 8。

启用矢量化

Hive 逐行处理数据。矢量化指导 Hive 以块（一个块包含 1,024 行）的方式处理数据，而不是以一次一行的方式处理数据。矢量化只适用于 ORC 文件格式。

- 若要启用矢量化查询执行，请导航到 Hive 的 `config` 选项卡并搜索 `hive.vectorized.execution.enabled` 参数。Hive 0.13.0 或更高版本的默认值为 `true`。
- 若要为查询的reducer启用矢量化执行，请将 `hive.vectorized.execution.reduce.enabled` 参数设置为 `true`。默认值为 `false`。

启用基于成本的优化（CBO）

默认情况下，Hive 遵循一组规则来找到一个最佳的查询执行计划。基于成本的优化（CBO）可以评估多个查询执行计划并向每个计划分配一个成本，然后确定成本最低的查询执行计划。

若要启用 CBO，请导航到 Hive 的 `config` 选项卡并搜索 parameter `hive.cbo.enable`，然后将开关按钮切换到“on”。

启用 CBO 后，可使用以下附加配置参数提高 Hive 查询性能：

- `hive.compute.query.using.stats` 如果设置为 `true`，则 Hive 会使用其元存储中存储的统计信息来应答类似于 `count(*)` 的简单查询。

- `hive.stats.fetch.column.stats` 启用 CBO 时，会创建列统计信息。Hive 使用元存储中存储的列统计信息来优化查询。如果列数较多，则提取每个列的列统计信息需要花费很长时间。如果设置为 `false`，则会禁用从元存储中提取列统计信息。

- `hive.stats.fetch.partition.stats`

行数、数据大小和文件大小等基本分区统计信息存储在元存储中。如果设置为 `true`，则会从元存储中提取分区统计信息。如果为 `false`，则从文件系统中提取文件大小，并从行架构中提取行数。

▼ Advanced hive-site

Fetch partition stats at compiler	<input type="text" value="true"/>
-----------------------------------	-----------------------------------

启用中间压缩

映射任务将创建reducer任务使用的中间文件。中间压缩可以缩小中间文件大小。

Hadoop 作业通常会遇到 I/O 瓶颈。压缩数据能够加快 I/O 和总体网络传输速度。

可用的压缩类型包括：

格式	工具	算法	文件扩展名	是否可拆分？
Gzip	Gzip	DEFLATE	.gz	否
Bzip2	Bzip2	Bzip2	.bz2	是
LZO	Lzop	LZO	.bz2	是（如果已编制索引）
Snappy	不适用	Snappy	Snappy	否

一般规则是，尽量使用可拆分的压缩方法，否则会创建极少的mapper。如果输入数据为文本，则 bzip2 是最佳选项。对于 ORC 格式，Snappy 是最快的压缩选项。

- 若要启用中间压缩，请导航到 **Hive** 的 **config** 选项卡，并将 `hive.exec.compress.intermediate` 参数设置为 `true`。默认值为 `false`。

▼ Advanced hive-site

hive.exec.compress.intermediate	<input type="text" value="true"/>
---------------------------------	-----------------------------------

备注 若要压缩中间文件，请选择一个 CPU 开销较低的压缩编解码器，即使该编解码器不能提供较高的压缩输出。

- 若要设置中间压缩编解码器，请将自定义属性 `mapred.map.output.compression.codec` 添加到 `hive-site.xml` 或 `mapred-site.xml` 文件。
- 添加自定义设置：
 - 导航到 **Hive** 的 **config** 选项卡并选择 **advance** 选项卡。
 - 在 **advance** 选项卡下，找到并展开 **custom hive-site** 窗格。
 - 单击 **custom hive-site** 窗格底部的 **Add Property** 链接。
 - 在 **Add Property** 窗口中，输入 `mapred.map.output.compression.codec` 作为键，输入 `org.apache.hadoop.io.compress.SnappyCodec` 作为值。
 - 单击 **Add**。

这会使用Snappy压缩来压缩中间文件。添加该属性后，它会显示在 **custom hive-site** 窗格中。

备注 此过程会修改 `$HADOOP_HOME/conf/hive-site.xml` 文件。

压缩最终输出

还可以压缩最终的 Hive 输出。

- 若要压缩最终的 Hive 输出，请导航到 **Hive** 的 **config** 选项卡，并将 `hive.exec.compress.output` 参数设置为 `true`。默认值为 `false`。
- 若要选择输出压缩编解码器，请根据上一部分的步骤 3 所述，将 `mapred.output.compression.codec` 自定义属性添加到 **custom hive-site** 窗格。

Add Property

Type	<input type="text" value="hive-site.xml"/>
Key	<input type="text" value="mapred.output.compression.codec"/>
Value	<input type="text" value="org.apache.hadoop.io.compress.SnappyCodec"/>

启用推理执行

推理执行可以启动特定数量的重复任务以检测运行速度缓慢的任务跟踪程序并将其加入方块列表，同时通过优化各项任务结果来改善总体作业执行。

不应该对输入量较大的长时间运行的 MapReduce 任务启用推理执行。

- 若要启用推理执行，请导航到 **Hive** 的 **config** 选项卡，并将 `hive.mapred.reduce.tasks.speculative.execution` 参数设置为 `true`。默认值为 `false`。

▼ Advanced hive-site

hive.mapred.reduce.tasks.speculative.execution	<input type="text" value="true"/>
--	-----------------------------------

优化动态分区

Hive 允许在表中插入记录时创建动态分区，且无需预定义每个分区。这是一项强大功能，尽管，它可能导致创建大量的分区并为每个分区创建大量的文件。

- 要让 Hive 执行动态分区，应将 `hive.exec.dynamic.partition` 参数值设置为 `true`（默认值）。
- 将动态分区模式更改为 `strict`。在 `strict`（严格）模式下，必须至少有一个分区是静态的。这可以阻止未在 `WHERE` 子句中包含分区筛选器的查询，即，`strict` 可阻止扫描所有分区的查询。进入 **Hive** 的 **config** 选项卡，并将 `hive.exec.dynamic.partition.mode` 设置为 `strict`。默认值为 `nonstrict`。
- 若要限制要创建的动态分区数，请修改 `hive.exec.max.dynamic.partitions` 参数。默认值为 `5000`。
- 若要限制每个节点的动态分区总数，请修改 `hive.exec.max.dynamic.partitions.pernode`。默认值为 `2000`。

启用本地模式

本地模式可让 Hive 在一台计算机上（有时是在单个进程中）执行某个作业的所有任务。如果输入数据较小，并且查询启动任务的开销会消耗总体查询执行资源的绝大部分，则此模式可以提高查询性能。

若要启用本地模式，请根据启用中间压缩部分的步骤 3 所述，将 `hive.exec.mode.local.auto` 参数添加到「custom hive-site」面板。

Add Property

Type	hive-site.xml
Key	hive.exec.mode.local.auto
Value	true

设置单个 MapReduce MultiGROUP BY

如果此属性设置为 true，则包含通用 `group-by` 键的 MultiGROUP BY 查询将生成单个 MapReduce 作业。

若要启用此行为，请根据启用中间压缩部分的步骤 3 所述，将 `hive.multigroupby.single reducer` 参数添加到 `custom hive-site` 面板。

Add Property

Type	hive-site.xml
Key	hive.multigroupby.single reducer
Value	true

其他 Hive 优化

以下部分介绍了可以设置的其他 Hive 相关优化。

联接优化

Hive 中的默认联接类型是“随机联接”。在 Hive 中，特殊的 mapping 会读取输入，并向中间文件发出联接键/值对。Hadoop 在随机阶段中排序与合并这些对。此随机阶段的系统开销较大。根据数据选择右联接可以显著提高性能。

联接类型	时间	方式	Hive 设置	注释
随机联接	默认选项 始终运行	从某个表的一部分内容中读取 根据联接键存储和排序 向每个化简器发送一个存储桶 在化简端执行联接	不需要过多的 Hive 设置	每次运行
映射联接	一个表可以装入内存	将小型表读入内存哈希表 通过大型文件的一部分流式处理 联接哈希表中的每条记录 只按映射器执行联接	<code>hive.auto.convert.join=true</code>	速度很快，但受限
排序合并存储桶	如果两个表： 排序方式相同 存储方式相同 每个进程：			
按排序/存储的列执行联接	从每个表中读取存储桶 处理值最小的行		<code>hive.auto.convert.sortmerge.join=true</code>	非常高效

执行引擎优化

有关优化 Hive 执行引擎的其他建议：

设置	建议	KMR默认值
<code>hive.mapjoin.hybridgrace.hashtable</code>	True = 更安全，但速度更慢；false = 速度更快	false
<code>tez.am.resource.memory.mb</code>	大多数引擎的上限为 4 GB	自动优化
<code>tez.session.am.dag.submit.timeout.secs</code>	300+	300
<code>tez.am.container.idle.release-timeout-min.millis</code>	20000+	10000
<code>tez.am.container.idle.release-timeout-max.millis</code>	40000+	20000

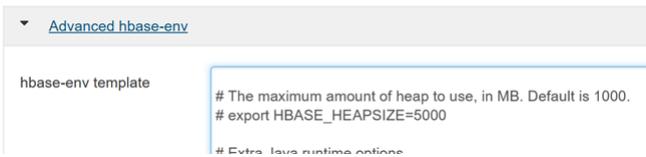
HBASE配置优化实践指南

可以通过 HBase 的 `config` 选项卡修改 HBase 配置。以下部分介绍了一些影响 HBase 性能的重要配置设置。

设置 HBASE_HEAPSIZE

HBase 堆大小指定 region server 和主服务器要使用的最大堆数量（以 MB 为单位）。默认值为 1,000 MB。优化群集工作负荷时应优化此项设置。

- 若要修改，请导航到 HBase 的 `config` 选项卡中的 `advanced hbase-env` 标签，然后找到 `HBASE_HEAPSIZE` 设置。
- 将默认值更改为 5,000 MB。



优化读取密集型工作负荷

以下配置对于提高读取密集型工作负荷的性能非常重要。

块缓存大小

块缓存是读取缓存。其大小由 `hfile.block.cache.size` 参数控制。默认值为 0.4，即总区域服务器内存的 40%。块缓存大小越大，随机读取的速度越快。

- 若要修改此参数，请导航到 HBase 的 `config` 选项卡，然后找到 `% of Regionserver Allocated to Read Buffers`。



2. 若要更改此值，请选择 **edit** 图标。

Memstore 大小

所有编辑内容都存储在称作 Memstore 的内存缓冲区中。此机制增大了可在单个操作中写入磁盘的总数据量，并可加速以后对最近编辑内容的访问。Memstore 大小由以下两个参数定义：

- `hbase.regionserver.global.memstore.UpperLimit`: 定义 Memstore 总共可以使用的区域服务器最大内存百分比。
- `hbase.regionserver.global.memstore.LowerLimit`: 定义 Memstore 总共可以使用的区域服务器最小内存百分比。

若要优化随机读取，可以减小 Memstore 的上限和下限。

从磁盘扫描时提取的行数

`hbase.client.scanner.caching` 设置定义在扫描程序中调用 `next` 方法时，要从磁盘读取的行数。默认值为 100。该数字越大，从客户端向区域服务器发出的远程调用数就越少，因而扫描速度也就越快。但是，这也会增大客户端上的内存压力。

Number of Fetched Rows when Scanning from Disk rows

重要 设置此值时，请不要使扫描程序中的下一次方法调用间隔时间大于扫描程序的超时时间。扫描程序超时期限由 `hbase.regionserver.lease.period` 属性定义。

优化写入密集型工作负荷

以下配置对于提高写入密集型工作负荷的性能非常重要。

最大区域文件大小

HBase 使用称作 HFile 的内部文件格式存储数据。属性 `hbase.hregion.max.filesize` 定义区域的单个 HFile 的大小。如果区域 (region) 中的 HFiles 总数大于此设置，则会将该区域拆分为两个区域。

Disk



区域文件大小越大，拆分数目越小。可以增大文件大小，以确定可以最大程度地提高写入性能的值。

避免阻止更新

- 属性 `hbase.hregion.memstore.flush.size` 定义 Memstore 刷新到磁盘的增量大小。默认大小为 128 MB。
- `hbase.hregion.memstore.block.multiplier` 定义。默认值为 4。允许的最大值为 8。
- 如果 Memstore 为 (`hbase.hregion.memstore.flush.size * hbase.hregion.memstore.block.multiplier`) 字节，则 HBase 会阻止更新。

使用刷新大小和块乘数的默认值时，如果 Memstore 大小为 $128 * 4 = 512$ MB，则会阻止更新。若要减少更新阻止计数，请增大 `hbase.hregion.memstore.block.multiplier` 的值。



HBase Region Block Multiplier

定义 Memstore 大小

Memstore 大小由 `hbase.regionserver.global.memstore.UpperLimit` 和 `hbase.regionserver.global.memstore.LowerLimit` 参数定义。将这些值设置为相等可以减少写入期间的暂停次数（同时提高刷新频率），并可以提高写入性能。

设置 Memstore 本地分配缓冲区

Memstore 本地分配缓冲区使用率由 `hbase.hregion.memstore.mslab.enabled` 属性确定。如果已启用 (true)，则可以防止在执行写入密集型操作期间出现堆碎片。默认值为 true。

开启服务的自动启动

Ambari 提供了服务异常中止后的自动重启功能，位置为 **Admin > service Auto Start**，将需要自动重启的组件的状态设置为 `enable`，保存设置。

在服务出现异常中止的时候，Ambari 会尝试自动重启该服务。

[Stack and Versions](#)[Service Accounts](#)[Kerberos](#)[Service Auto Start](#)

Service Auto Start Configuration

Ambari services can be configured to start automatically on system boot. Each service can be configured to start or not start selectively.

Auto-Start Services

Enabled

Service

Component

YARN



App Timeline Server

HDFS



NodeManager

Flume



ResourceManager

HBase



E

MapReduce2



Hive



Hue



Spark2



Ambari Metrics



Oozie



Zeppelin Notebook



ZooKeeper



设置HDFS再平衡 (rebalance)

HDFS集群在新增加点后, 会出现已有节点与新节点存储空间不平衡的状态, 已有节点占用空间较多, 而新节点占用空间较少。

Actions ▾

🔍 Filter by host and component attributes or search by keyword ...

<input type="checkbox"/>	Name ↕	IP Address ↕	Rack ↕	Cores ↕	RAM ↕	Disk
<input type="checkbox"/>	✓ kmr-b3e226fa-gn-43dd9a2b...	172.31.48.183	/default-rack/89962...	8 (8)	29.45GB	
<input type="checkbox"/>	✓ kmr-b3e226fa-gn-43dd9a2b...	172.31.48.184	/default-rack/9c44c...	8 (8)	29.45GB	
<input type="checkbox"/>	✓ kmr-b3e226fa-gn-43dd9a2b...	172.31.48.172	/default-rack/85d07...	8 (8)	29.45GB	
<input type="checkbox"/>	✓ kmr-b3e226fa-gn-43dd9a2b...	172.31.48.191	/default-rack/e9c7fc...	8 (8)	29.45GB	
<input type="checkbox"/>	✓ kmr-b3e226fa-gn-43dd9a2b...	172.31.48.190	/default-rack/d7d66...	8 (8)	29.45GB	
<input type="checkbox"/>	✓ kmr-b3e226fa-gn-43dd9a2b...	172.31.48.163	/default-rack/1f07c7...	8 (8)	29.45GB	
<input type="checkbox"/>	✓ kmr-b3e226fa-gn-43dd9a2b...	172.31.48.185	/default-rack/70c41...	8 (8)	29.45GB	
<input type="checkbox"/>	✓ kmr-b3e226fa-gn-43dd9a2b...	172.31.48.189	/default-rack/70c41...	8 (8)	29.45GB	
<input type="checkbox"/>	✓ kmr-b3e226fa-gn-43dd9a2b...	172.31.48.57	/default-rack/1f07c7...	8 (8)	29.45GB	
<input type="checkbox"/>	✓ kmr-b3e226fa-gn-43dd9a2b...	172.31.48.66	/default-rack/1f07c7...	8 (8)	29.45GB	

在这个时候可以通过hdfs的rebalance功能对hdfs文件进行再平衡。

rebalance的选项在HDFS>service actions>rebalance HDFS

Summary

Heatmaps

Configs

Quick Links ▾

Group

Default (0) ▾

Manage Config Groups

b

<
>

V12

admin
4 months ago
HDP-2.6

✓

V11

admin
4 months ago
HDP-2.6

V10

kmr
4 months ago
HDP-2.6

V9

kmr
4 months ago
HDP-2.6

V8

4 m

⌘ ▾

V12

✓

admin authored on Sun, Aug 12, 2018 12:04

Settings

Advanced

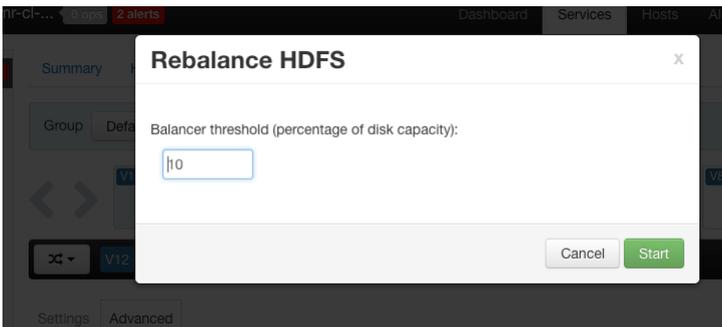
▾ Advanced hdfs-site

dfs.datanode.balance.bandwidthPerSec

6250000

Balancer threshold这个参数代表的含义是与HDFS磁盘使用率的偏差的阈值。

The dashboard shows a sidebar with service status: HDFS (1 alert), YARN, MapReduce2, Tez, Hive, HBase, Pig, and Sqoop. The main area has tabs for Metrics, Heatmaps, and Config History. Under Metrics, there are dropdowns for 'Metric Actions' and 'Last 1 hour'. Two cards are visible: 'HDFS Disk Usage' with a 33% gauge chart, and 'DataNodes Live' showing 10/10.



如上的设置，如果一个data node的使用率大于44%，多出的部分的数据就会rebalance到小于23%的节点上。

此外，rebalance过程会占用内部带宽，因此，这个参数会决定rebalance的所需的时间，不过也不建议将这个参数设置过大，以免影响线上的任务。

The configuration page shows 'bandwidthPerSec' set to 6250000. It includes a list of nodes (V12 to V7) with their last update times and user information. The 'Advanced hdfs-site' section is visible below.

默认值为每秒60M。

设置ES balance的阈值

ES集群中可以设置watermark的阈值来决定存shard rebalance的阈值

The configuration page shows 'water' set to 0.8. It includes a list of nodes (V8 to V1) with their last update times and user information. The 'Advanced elastic-site' section is visible below, showing 'watermark_high' and 'watermark_low' settings.

high: 表示超过这个阈值就会触发shard balance

low: 表示超过这个阈值就不会再向这个节点分配shard