

Solr安装部署

www.huawei.com





前言

- 本章主要介绍华为**FusionInsight HD Solr**服务部署规则、方法，以及相关注意事项。



目标

- 学完本课程后，您将：
 - 熟悉**FusionInsigh HD Solr**服务部署原则；
 - 掌握**FusionInsigh HD Solr**服务部署方法；
 - 掌握**FusionInsigh HD Solr**服务安装后配置；
 - 能够处理常见安装问题；



目录

1. **Solr**部署基本概念
2. **Solr**服务部署规则
3. **Solr**服务部署方法
4. **Solr**服务安装后参数配置
5. **Solr**服务健康指标说明

Solr部署基本概念

- 角色概念说明：
 - **SolrServerAdmin**：主要提供在**Manager**界面能够跳转到**Solr Admin**页面的链接；业务方面与其他**SolrServerN**（**N=1~5**）功能一致。
 - **SolrServer1~5**：提供**Solr**业务的角色，主要为索引与查询。
 - **HBaseIndexer**：将**HBase**中数据导入**Solr**的角色。



目录

1. **Solr**部署基本概念
2. **Solr**服务部署规则
3. **Solr**服务部署方法
4. **Solr**服务安装后参数配置
5. **Solr**服务健康指标说明

Solr部署规则

- **Solr**数据存储在**HDFS**相比存储在本地磁盘**性能下降30%~50%**，优先选择部署到**HDFS**，每个节点部署三个**Solr**实例；
- 单节点实时索引速度要求达到**2MB/s**以上的，建议部署到本地磁盘，每个节点**Solr**实例满配，每个**Solr**实例单独挂盘。

Index 存储位置	单节点 写索引速度	膨胀率
Local Disk	4MB/s	50%
HDFS	2MB/s	50%

Solr部署规则

服务名称	角色名称	内存最小要求	依赖关系	角色业务部署原则
Solr	SolrServerN (N=1~5)	2GB	依赖于 ZooKeeper 说明： Solr 数据选择存 放在 HDFS 上时，还依 赖于 HDFS	<ul style="list-style-type: none">每个节点可以配置5个实例。每个节点上的实例个数均匀分布。索引数据存放在HDFS上时，建议Solr实例与DataNode合布，且建议配置3个SolrServer实例内存资源充足情况下，建议内存配置为16G。
	SolrServer Admin	2GB		分别部署在 2 个数据节点上，非主备。 内存资源充足情况下，建议内存配置为 16G 。
	HBase Indexer	2GB	依赖 HBase 、 HDFS 和 ZooKeeper	建议每个 SolrServer 实例所在节点均部署一个。



目录

1. Solr部署基本概念
2. Solr服务部署规则
3. Solr服务部署方法
4. Solr服务安装后参数配置
5. Solr服务健康指标说明

Solr新建部署

LLD工具进行模板部署

- 在**LLD**工具中选择**Solr**服务
- 配置浮动**IP**，需要在业务平面，保证无冲突
- 选择索引数据是否存储在**HDFS**上，如果存储本地，需要对**Solr**进行磁盘规划

手动部署

- 在**Manager**界面上选择**Solr**服务
- 配置浮动**IP**，需要在业务平面，保证无冲突
- 选择索引数据是否存储在**HDFS**上，如果存储本地，需要采用**LLD**工具对**Solr**进行磁盘规划

Solr扩容部署

扩容场景：

- 物理资源消耗过大，即**Solr**服务节点的**CPU**、内存占用率过高、磁盘空间不足场景。具体操作步骤：进入**FusionInsight Manager**后，选择“服务管理>**Solr**>实例”，单击所要查看的实例，即可看到当前一段时间内，**CPU**、内存使用情况；同时，进入服务端后台查看磁盘使用情况。
- 服务扩容，即**Solr**单实例的索引数据太大，索引的数目达到**10**亿条或者是数据大小达到**1T**。

扩容方式：

- 增加节点，当**Solr**出现物理资源消耗过大，或者是无法在已经有的节点上添加**solr**实例，则需要新的节点上增加实例。
- 增加**SolrServer**实例，可直接在**FusionInsight Manager**上添加**solr**实例。

Solr减容部署

减容场景：

- 节点需要重新安装操作系统。
- 当Solr创建的shard下的replica过多，需要减少replica数目

。

减容方式：

- 在FusionInsight Manager上删除solr实例。



目录

1. **Solr**部署基本概念
2. **Solr**服务部署规则
3. **Solr**服务部署方法
4. **Solr**服务安装后参数配置
5. **Solr**服务健康指标说明

Solr安装后参数配置

- 登录**FusionInsight Manager**，点击“服务管理>**Solr**>服务配置”，选择“参数类别”为全部配置。

服务 > Solr 服务配置

服务状态 实例 **服务配置** 资源贡献排名

保存配置 导入服务配置 导出服务配置 参数类别: 全部配置 角色: 主机:

Solr

Default

HBaseIndexer

HBaseIndexer

日志

系统

参数	值
fs.defaultFS	hdfs://hacluster
INDEX_STORED_ON_HDFS	<input checked="" type="checkbox"/> TRUE <input type="checkbox"/> FALSE
* SOLRHASERVER_FLOAT_IP	

当前页: 1 总页数: 1 < > 跳转到 确定

常用参数说明-SolrServer

参数	说明	建议值
SOLRHASERVER_FLOAT_IP	HA 的浮动IP	N/A
INDEX_STORED_ON_HDFS	是否将索引存储在 HDFS 上。对于数据膨胀快，随时准备扩容的，并且对索引和查询性能要求不是特别苛刻的，推荐索引数据存储在 HDFS 。对于数据膨胀慢，长时间不需要进行扩容，并且对索引和查询性能要求苛刻的，就推荐使用索引数据存储在本地磁盘。	FALSE
SOLR_CORES_RECOVER_DELAY	Solr 自动恢复 core 的延迟时间，当 Collection 数较多时，可以适当增大。	30000
SOLR_MEM_USED_THRESHOLD	Solr 创建 Collection 时，当已使用的内存超过该阈值时，则无法创建 Collection 。	90
SOLR_OVERSEER_QUEUE_CLEAR_NUM	Overseer 任务列表缓冲的长度，当前的任务列表>(任务列表最大长度 - 任务列表缓冲长度)，开始执行清除 overseer 的 queue 下任务。	1000

常用参数说明-SolrServer（续）

参数	说明	建议值
SOLR_OVERSEER_QUEUE_MAX_NUMCHIL	Overseer 处理的任务列表的最大长度。	10000
SOLR_GC_OPTS	JVM的GC配置，根据节点实际内存，可适当调大Xms，Xmx，提升SolrServer性能。	-Xms2G -Xmx4G -DIgnoreReplayReq Detect -XX:+UseConcMark SweepGC -XX:+CMSParallelRem arkEnabled
SOLR_CONTROL_PORT	SolrServer的tomcat停止端口	21100-21149
SOLR_PORT	SolrServer业务的http的端口	21100-21149
SOLR_SEC_PORT	SolrServer业务的https的端口	21100-21149

常用参数说明-HBaseIndexer

参数名称	说明	建议值
GC_OPTS	JVM的GC配置。	-Xms4G -Xmx4G -XX:NewSize=512M -XX:MaxNewSize=512M -XX:MetaspaceSize=128M -XX:MaxMetaspaceSize=512M -XX:MaxDirectMemorySize=512M -XX:+UseConcMarkSweepGC -XX:+CMSParallelRemarkEnabled -XX:CMSInitiatingOccupancyFraction=65
hbaseindexer.indexer.threads	HBaseIndexer实例进行索引操作时启动的并发线程数量	20
hbaseindexer.sep.batch.size	提交到SolrServer的建立索引数据的batch大小	100



目录

1. Solr部署基本概念
2. Solr服务部署规则
3. Solr服务部署方法
4. Solr服务安装后参数配置
5. Solr服务健康指标说明

Solr服务健康指标说明

- 服务健康状态
 - 指标项名称：服务状态
 - 指标项含义：检查**Solr**服务状态是否正常。如果状态不正常，则认为不健康。
 - 恢复指导：如果该指标项异常，建议参见告警进行处理。
- 检查告警
 - 指标项名称：告警信息
 - 指标项含义：检查主机是否存在未清除的告警。如果存在，则认为不健康。
 - 恢复指导：如果该指标项异常，建议参考告警处理指南进行处理。



本章总结

- 本章主要介绍了**FusionInsight HD Solr**的安装与部署，分别从安装的部署原则和约束、部署的方法、组件的配置几个方面进行详细介绍。通过本章的学习，能够清楚**FusionInsight HD Solr**服务的安装部署流程，并能够自行安装**FusionInsight HD Solr**服务，同时能够进行简单问题的处理。

习题

1. Solr组件的角色有哪些（ ）。

- A. SolrServerAdmin B. SolrServerN (N=1~5)
C. HBaseIndexer D. RegionServer

思考题

1. **SolrServerAdmin**角色与**SolrServerN**（**N=1~5**）角色差别？
2. **Solr**组件根据什么进行判断数据是存储于**HDFS**还是本地磁盘？

Thank you

www.huawei.com