1. **集群有8个节点，每个节点有8块硬盘（默认3副本）。如果某个节点有3块盘损坏，是否可能存在数据块丢失情况；如果有3个节点发生故障，是否可能存在数据块丢失情况；并简述原因。**

第一个不会，因为副本放在三个不同的节点上

第二个有可能，如果三个副本正好是在这三个节点上，则会丢失。

*–第1个副本放在客户端所在节点*

*• 如果是远程客户端，block会随机选择节点*

*• 系统会首先选择空闲的DataNode节点*

*– 第2个副本放在不同的机架节点上*

*– 第3个副本放在与第2个副本同一机架的不同机器上*

1. **请描述一个100GB文件写入HDFS的整个过程：**

– 客户端发送创建文件指令给分布式文件系统

– 文件系统告知namenode

• 检查权限，查看文件是否存在

• EditLog增加记录

• 返回输出流对象

– 客户端往输出流中写入数据，分成一个个数据包

– 根据namenode分配，输出流往datanode写数据

• 多个datanode构成一个管道pipeline，输出流写第一个，后面的转发

– 每个datanode写完一个块后，返回确认信息

– 写完数据，关闭输出流

– 发送完成信号给namenode

1. **假设集群的每个节点初始有6块硬盘，运行一段时间后，每个节点又加了4块新硬盘，为了使数据在所有硬盘上分布均匀，能否通过hdfs balancer达到效果，为什么？并列出能达到效果的两种措施。**

不能，旧版本的hdfs仅支持节点间的数据平衡，新版本可通过balancer实现

1. 手动重写所有数据
2. 将数据全部移到几个节点上，再在节点间数据平衡
3. **请描述HDFS的高可用性实现机制：**
4. 支持NameNode HA。集群可以运行两个冗余的NameNode，一个作为active节点提供服务，一个standby节点作为热备份，集群只能有一个activeNameNode为了保证standby节点与active节点之间数据的同步，两个节点会与另一组服务集群“journalNodes” (JNs)进行通信

• 当active节点对命名空间做任何改动时，操作日志会被记录到超过半数的JournalNode上

• standby节点会去JournalNode上读取这些日志

1. 利用2N+1台JournalNode存储EditLog
2. 最多容忍N台服务器挂掉
3. 基于Paxos的一致性算法算法

如果要实现主备切换的自动化，还需要额外添加两个服务

1. – 额外一个ZooKeeper集群

• 实现NameNode故障检测

• active NameNode的选举

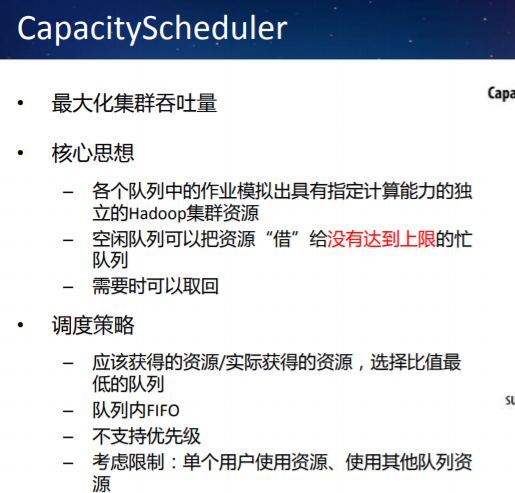
1. – 每个NameNode上运行ZKFailoverController (ZKFC)服务

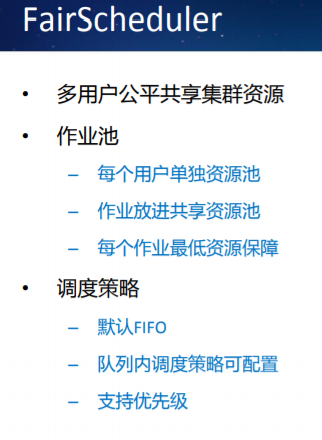
• NameNode健康状态的检查

• ZooKeeper会话的管理

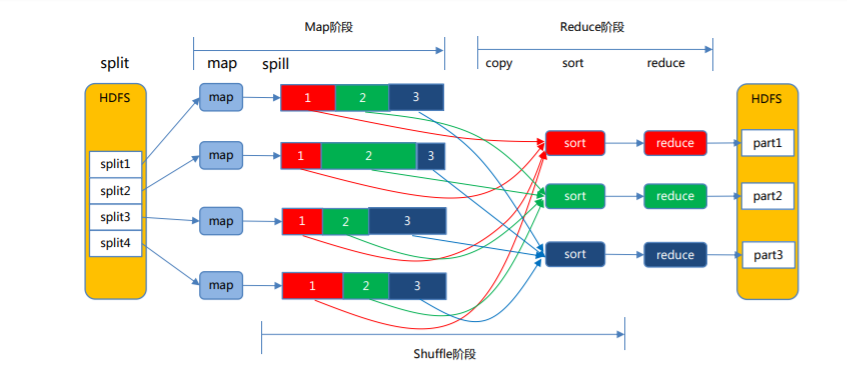
• ZooKeeper-based选举

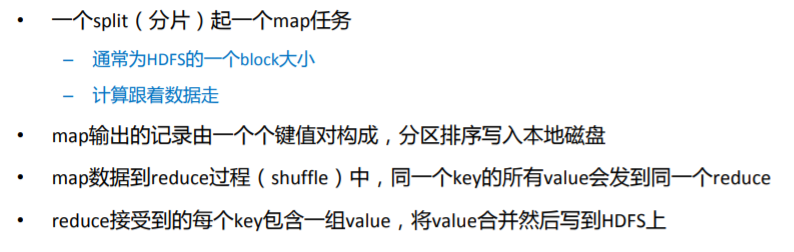
1. **请描述TDH平台中在Yarn上可以使用哪几种调度策略，并分别阐述各调度策略的特点。**

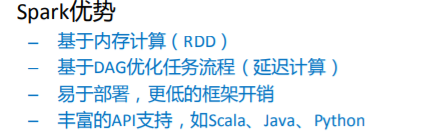




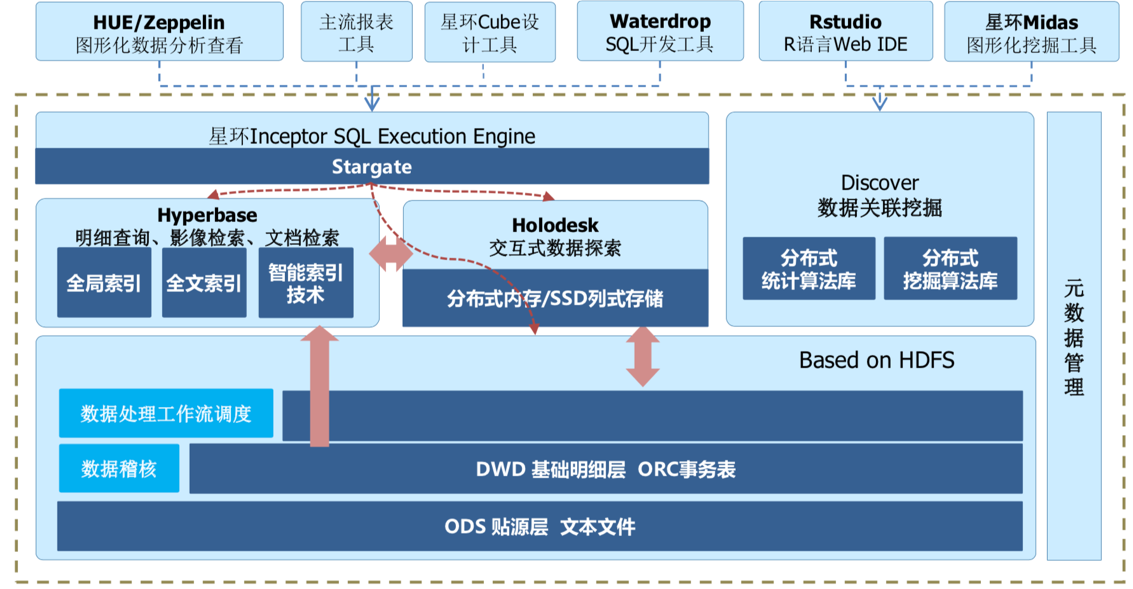
1. **请以WordCount为例描述MapReduce的运行过程，并列出Spark相较MapReduce的优势**：







1. **请列举出平台支持的5种存储格式/引擎的表，并详细描述各自的存储特点、使用场景、支持的操作以及是否支持分区分桶。**

****

Text表：TEXT表即文本格式的表，是Inceptor默认的表格式。在数据量大的情况下，TEXT表的统计和

查询性能都比较低；TEXT表也不支持事务处理，所以通常用于将文本文件中的原始数据导入

Inceptor中。针对不同的使用场景，用户可以将其中的数据放入ORC表或Holodesk表中。

ORC表：ORC表即ORC格式的表。优化的列式存储，轻量级索引，压缩比高，只支持insert

– Inceptor中数仓离线分析的主要表类型，可由Text表生成

– 在Inceptor中，ORC表还分为ORC事务表和非事务表。

事务表：– 支持事务处理和更多增删改语法（INSERT VALUES/UPDATE/DELETE/MERGE），所以如果

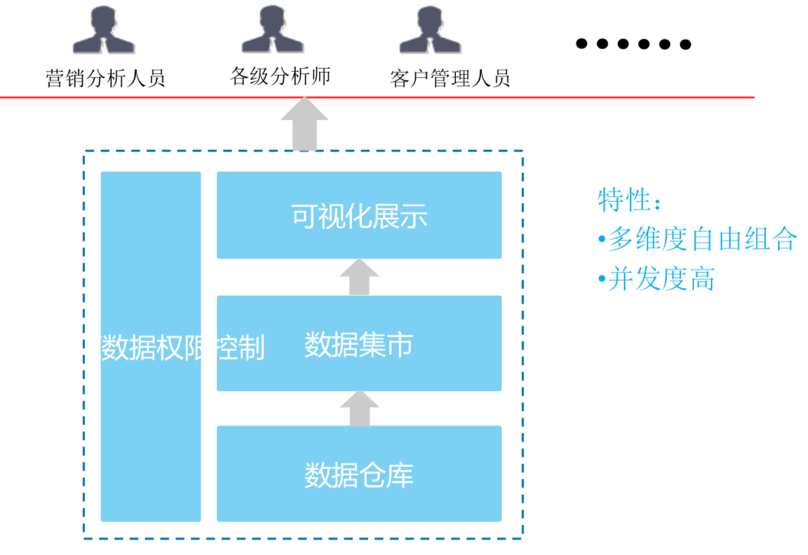
您需要对表进行事务处理，应该选择使用ORC事务表。

– ORC事务表是ORC分桶表的一个特例

– 表属性必须有TBLPROPERTIES ("transactional"="true")

– 数据量特别大，建议使用分区的ORC事务表

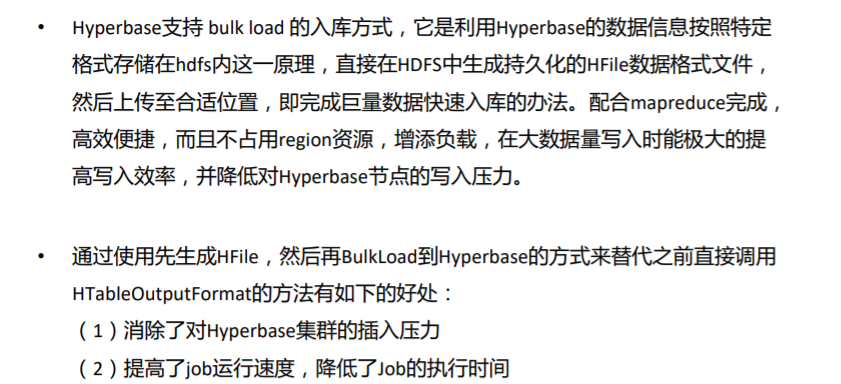
HoloDesk表：



Hyperbase表：



1. **请简述bulkload的作用和操作步骤（包括原理的步骤和使用sqlbulkload的步骤）**







1. **写出一下场景下的优化思路**

（1）、假设在Inceptor上执行任务，发现Map Task数量多、执行时间短，应采取哪种措施来提升性能？

对数据块进行合并：Automerge（碎片自动合并）

（2）、请简述在Inceptor中大表与大表做join、大表与小表做join时分别有哪些优化手段

大表与大表的普通JOIN（这里所说的普通JOIN并不单指语句中的JOIN操作，而是指它的实现方法）：实现普通JOIN的过程是这样的：扫描过滤两张表的数据（Map Stages），然后通过Shuffle将Key哈希值相同的数据分发到各个节点，在各节点内部执行JOIN（Reduce Stages）

MapJoin是一种针对大表与小表JOIN的特殊实现方式，在大小表数据量悬殊的情况下能有效的提升JOIN执行效率，一般受优化开关或者Hint控制启动。

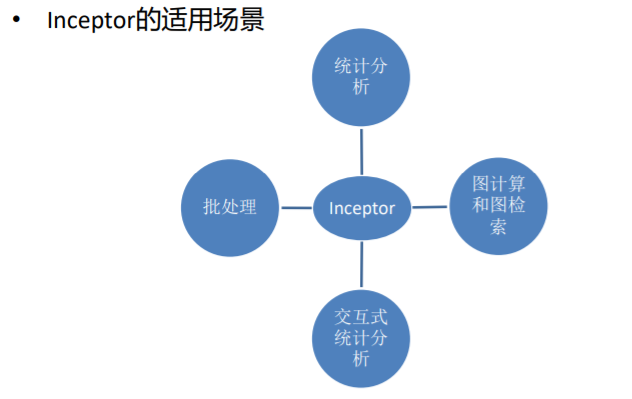
1. **请描述一个100GB文件写入Hyperbase表的整个过程（使用bulkload方式实现）**



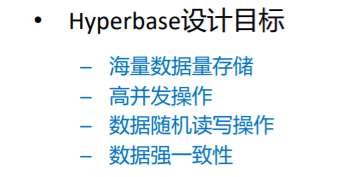
1. **请列出TDH下的4大组件（Inceptor、Hyperbase、StreamSQL、Discover）的特性以及适用场景。**

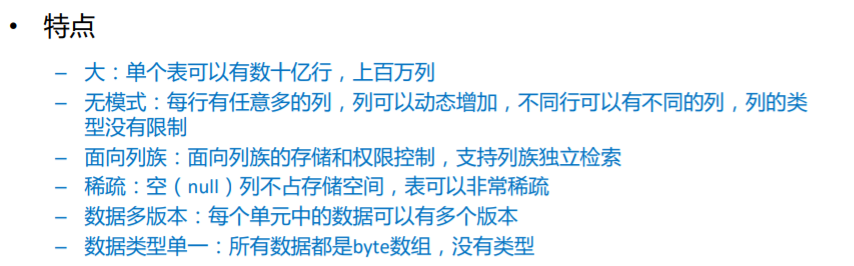
****

**Inceptor**



Hyperbase





Streamsql

流式sql，StreamSQL的计算运行于流计算引擎Transwarp Slipstream之上，该引擎混合了事件驱动和微批处理，因此既可以支持有低延迟需求的任务也可以处理高吞吐任务，能够应对不同类型业务。

Discover

分布式统计学习和机器学习算法库、同时提供R语言交互式数据挖掘和数据展现功能

1. **请描述高并发检索和综合搜索的场景特点，这两种场景应使用哪种技术来做支撑，并指出数据和索引各自的存储位置。**

高并发检索：海量数据存储，高并发操作，数据随机读写操作，数据强一致性

技术：使用hbase，数据可以存在hbase中，也可以存在hdfs中，索引表存在hbase中

综合搜索：分布式实时文档存储，每个字段都可以被索引与搜索，分布式实时分析搜索引擎，支持高扩展，支持结构化和非结构数据

技术：使用Elasticsearch，数据和索引都存储在ES中。