



高级Hive

谭唐华

课程大纲

- Hive数据压缩
- Hive数据存储
- Hive企业优化

Hive数据压缩

- 压缩格式: bzip2, gzip, lzo, **snappy**等
- 压缩比 : bzip2>gzip>lzo bzip2最节省存储空间
- 解压速度 : lzo>gzip>bzip2 lzo解压速度是最快的



- Reduces Storage Need
- Less Disk I/O
- Speeds up Network Transfer



- Consumes CPU



Compression Speed



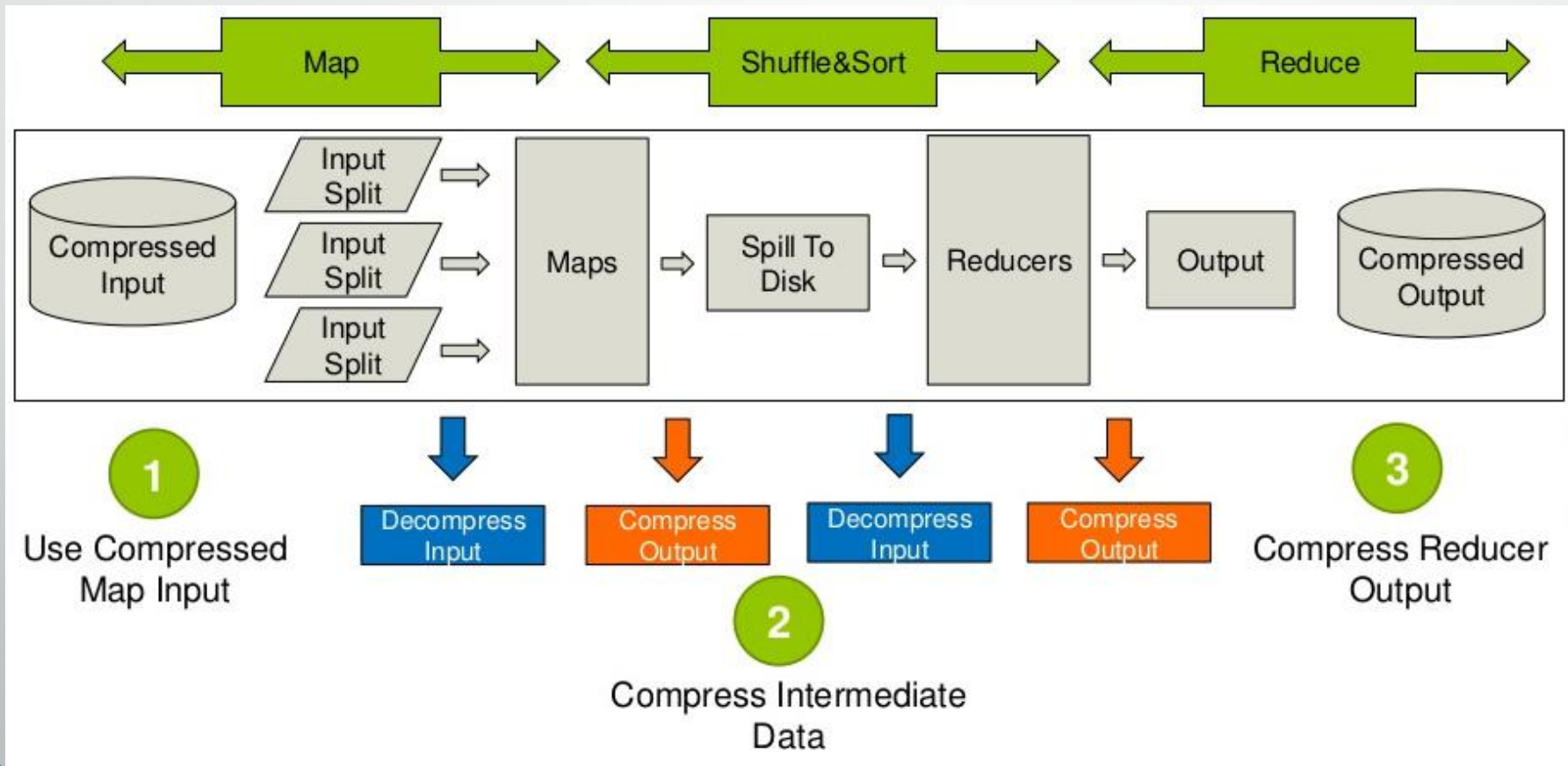
Compression Ratio



压缩特点

- Hadoop jobs are usually IO bound;
- Compression reduces the size of data transferred across network;
- Overall job performance may be increased by simple enabling compression;
- Splittability must be taken into account;

MapReduce中的压缩



MapReduce支持压缩格式

- Zlib → `org.apache.hadoop.io.compress.DefaultCodec`
- Gzip → `org.apache.hadoop.io.compress.GzipCodec`
- Bzip2 → `org.apache.hadoop.io.compress.BZip2Codec`
- Lzo → `com.hadoop.compression.lzo.LzoCodec`
- Lz4 → `org.apache.hadoop.io.compress.Lz4Codec`
- Snappy → `org.apache.hadoop.io.compress.SnappyCodec`

MapReduce压缩设置

Compressed Input Usage	File format is auto recognized with extension. Codec must be defined in core-site.xml.
Compress Intermediate Data (Map Output)	<code>mapreduce.map.output.compress = True;</code> <code>mapreduce.map.output.compress.codec = CodecName;</code>
Compress Job Output (Reducer Output)	<code>mapreduce.output.fileoutputformat.compress = True;</code> <code>mapreduce.output.fileoutputformat.compress.codec = CodecName;</code>

Hive压缩设置

Compressed Input Usage	<p>Can be defined in table definition</p> <pre>STORED AS INPUTFORMAT \"com.hadoop.mapred.DeprecatedLzoTextInputFormat\"</pre>
Compress Intermediate Data (Map Output)	<pre>SET hive.exec.compress.intermediate = True; SET mapred.map.output.compression.codec = CodecName; SET mapred.map.output.compression.type = BLOCK / RECORD;</pre> <p>Use faster codecs such as Snappy, Lzo, LZ4 Useful for chained mapreduce jobs with lots of intermediate data such as joins.</p>
Compress Job Output (Reducer Output)	<pre>SET hive.exec.compress.output = True; SET mapred.output.compression.codec = CodecName; SET mapred.output.compression.type = BLOCK / RECORD;</pre>

课程大纲

- Hive数据压缩
- Hive数据存储
- Hive企业优化

Hive数据存储

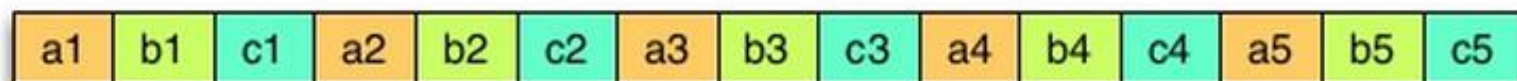
- Hive支持文件格式
 - TextFile
 - RCFILE
 - ORC
 - Parquet
 - AVRO

行存储和列存储

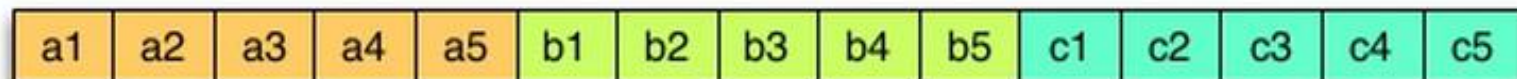
Logical table representation

a	b	c
a1	b1	c1
a2	b2	c2
a3	b3	c3
a4	b4	c4
a5	b5	c5

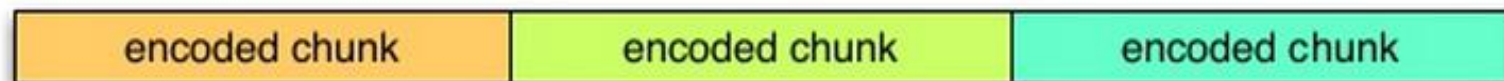
Row layout



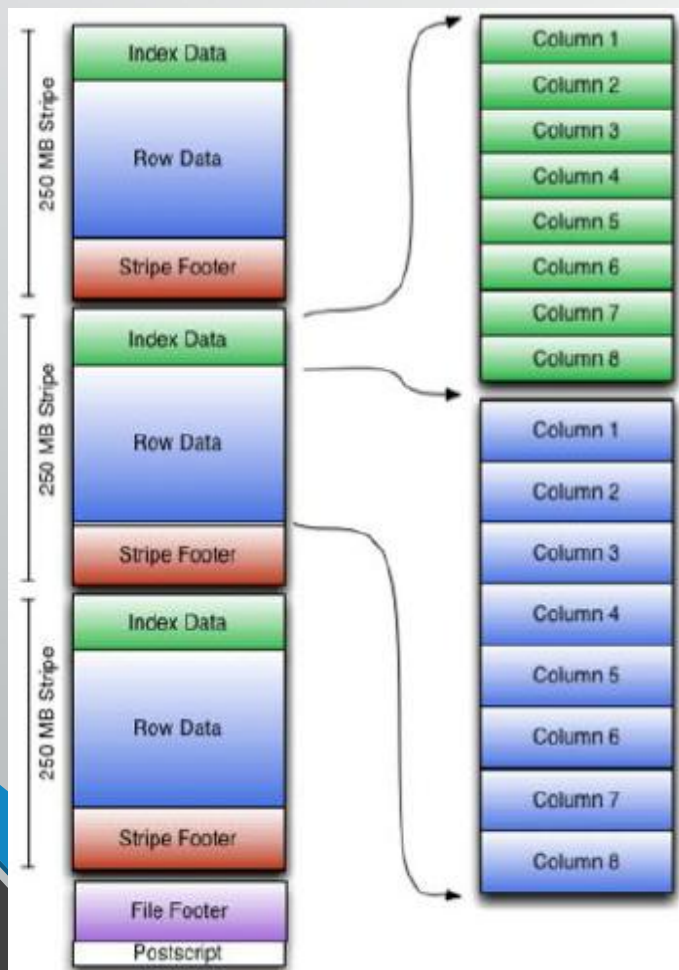
Column layout



encoding



ORC格式 (hive/shark/spark支持)



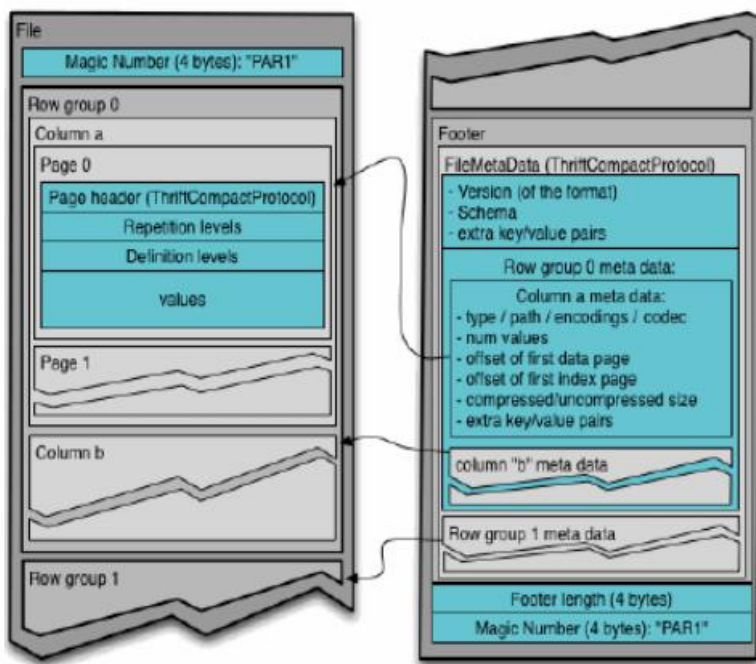
存储方式采用数据按行分块，每块按照 列存储

压缩快，快速列存取

效率比rcfile高，是rcfile的改良版本

PARQUET格式

■ PARQUET格式 (twitter+cloudera开源 , Hive、Spark、drill , Impala、Pig等支持)



➤ 列式存储格式

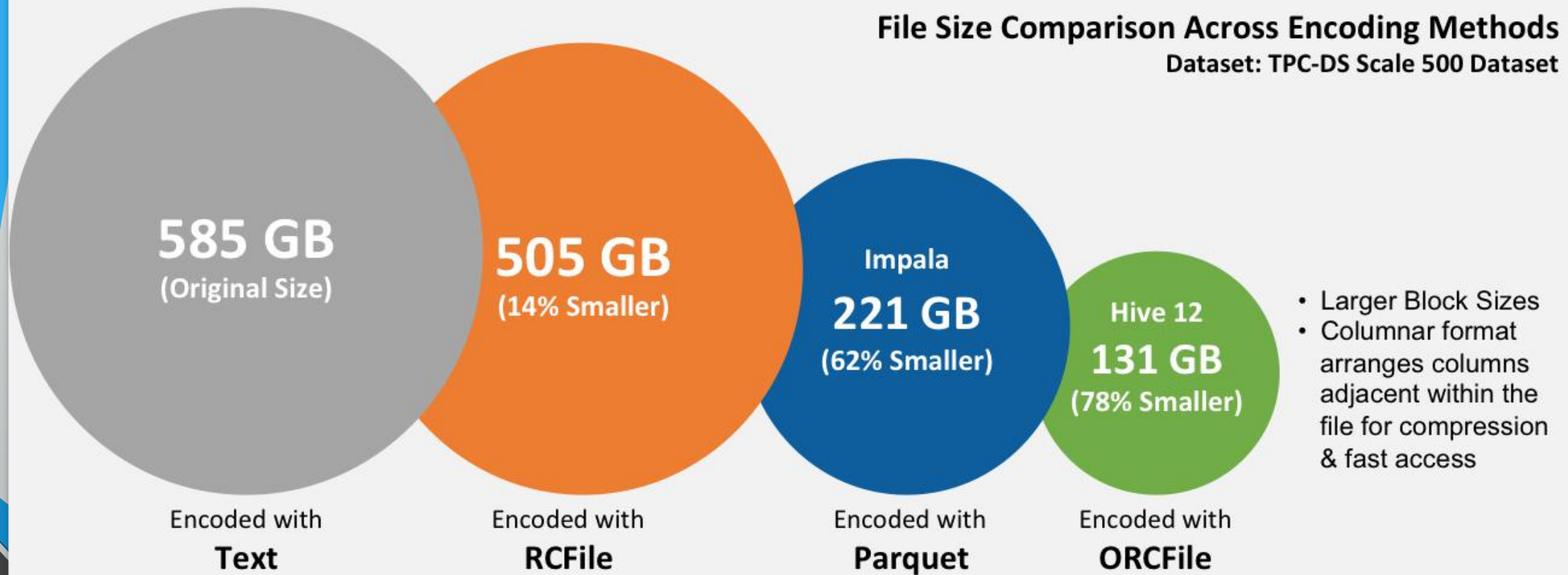
➤ 高效压缩和编码，且使用更少的IO 操作取出需要的数据。

➤ Parquet比较复杂，其灵感主要来自于dremel，Parquet存储结构的主要亮点是支持嵌套数据结构以及高效且种类丰富的算法（以应对不同值分布特征的压缩）

压缩对比

File Size Comparison Across Encoding Methods

Dataset: TPC-DS Scale 500 Dataset



课程大纲

- Hive数据压缩
- Hive数据存储
- Hive企业优化

Hive企业优化

- 大表【拆分】

- 子表

- 外部表、分区表

- 结合使用
 - 多级分区

- 数据

- 存储格式 (textfile、orcfile、parquet)
 - 数据压缩 (snappy)

- ◆ SQL

- 优化SQL语句
 - join , filter

- ◆ MapReduce

- Reduce Number
 - JVM重用
 - 推测执行

Hive企业优化

◆ 并行执行

hive.exec.parallel.thread.number	8
hive.exec.parallel	false

◆ Reduce数目

mapreduce.job.reduces	1
-----------------------	---

◆ JVM重用

mapreduce.job.jvm.numtasks	1
----------------------------	---

◆ 推测执行

mapreduce.map.speculative	true
hive.mapred.reduce.tasks.speculative.execution	true
mapreduce.reduce.speculative	true

◆ Map数目

hive.merge.size.per.task	256000000
--------------------------	-----------

严格模式

-- 对分区表进行查询，在 where 子句中没有加分区过滤的话，将禁止提交任务（默认：nonstrict）

```
set hive.mapred.mode=strict;
```

注：使用严格模式可以禁止 3 种类型的查询：

（1）对于分区表，不加分区字段过滤条件，不能执行

（2）对于 order by 语句，必须使用 limit 语句。

（3）限制笛卡尔积的查询（join 的时候不使用 on，而使用 where 的）。

Hive传递参数

hivevar和hiveconf传递变量的方法

hive -hivevar -f file	hive -hivevar tbnome='a' -hivevar count=10 -f filename.hql
hive -hivevar -e cmd	hive -hivevar tbnome='a' -hivevar count=10 -e 'select * from \${hivevar:tbnome} limit \${hivevar:count}'
hive -hiveconf -f file	hive -hiveconf tbnome='a' -hiveconf count=10 -f filename.hql
hive -hiveconf -e cmd	hive -hiveconf tbnome='a' -hiveconf count=10 -e 'select * from \${hivevar:tbnome} limit \${hivevar:count}'

元数据备份

- mysql备份：备份的数据库名称是metastore
- `$ mysqldump -uroot -p metastore > /path/metastore.sql`
- mysql还原：
- 方式1: `$ mysql -uroot -p metastore < /path/metastore.sql`
- 方式2: `$ mysql -uroot -p`
- `mysql> source /path/metastore.sql ;`

总结

- Hive数据压缩和文件存储格式
- Hive企业优化策略