# 目录

入门

架构

简单教程

安装

配置

链接数据源

ODBC/JDBC接口

查询数据

性能调整

日记与调试

SQL参考

数据源和文件格式

开发自定义函数

故障排除

开发者信息

发行说明

# 入门简介

# 架构

# 简单教程

# 安装

# 配置

# 连接数据源

## 连接数据源简介：

存储插件是Drill用来连接数据源的一个软件模块。存储插件通常优化Drill查询的执行，提供数据的位置，并配置工作区和文件格式以读取数据。Drill安装了多个存储插件，您可以根据自己的环境进行配置。Drill通过存储插件连接到数据源，例如数据库，本地或分布式文件系统上的文件或Hive Metastore。

您可以修改存储插件的默认配置X，并为新配置指定一个唯一的名称Y.本文档将Y称为不同的存储插件，但它实际上只是原始接口的重新配置。在执行查询时，Drill通过以下几种方式之一获取存储插件配置名称：

查询的FROM子句可以标识要使用的插件。

使用 命令可以在查询之前。

您可以在启动Drill取时指定存储插件。

Drill 插件处于apache drill 和data Source 之间。

除了前面提到的功能之外，存储插件还执行扫描程序和写入程序功能，并向执行引擎通知任何本机功能，例如谓词下推，连接和SQL。

## 存储插件注册：

通过存储插件将Drill连接到文件系统，Hive，HBase或其他数据源。在“drill Web控制台”的“存储”选项卡上，您可以查看并重新配置存储插件。如果未启用HTTPS支持（默认），请http://<IP address>:8047/storage转至查看并配置存储插件。IP地址是分布式系统或localhost嵌入式系统中安装的Drillbits之一的主机名或IP地址。如果启用了HTTPS支持，请转到https://<IP address>:8047/storage。

drill安装注册cp，dfs，hbase，hive，和mongo默认存储插件的配置。

cp

指向Drill类路径中的JAR文件，例如employee.json您可以查询的文件。

dfs

指向本地文件系统，但可以将此存储插件配置为指向任何分布式文件系统，例如Hadoop或S3文件系统。

hbase

提供与HBase的连接。

hive

将Drill与Hive文件，HBase和库的元数据抽象相集成，以读取数据并在SerDes和UDF上进行操作。

mongo

提供到MongoDB数据的连接。

在Drill(Mapr-sandbox,这个是某个公司的发行的一个hadoop 版本)沙箱中，dfs存储插件配置将您连接到预先配置了Drill的Hadoop环境。如果您安装了Drill，则将您dfs连接到文件系统的根目录。

注册存储插件配置

要注册新的存储插件配置，请输入存储名称，单击CREATE，以JSON格式提供配置，然后单击UPDATE。

存储插件配置持久性

Drill将存储插件配置保存在临时目录（嵌入模式）或ZooKeeper（分布式模式）中。例如，在Mac OS X上，Drill使用/tmp/drill/sys.storage\_plugins存储插件配置。临时目录在您重新引导时清除。在嵌入式模式下运行钻取时，将sys.store.provider.local.path选项添加到drill-override.conf文件以及用于存储插件配置的路径。例如：

drill.exec: {

cluster-id: "drillbits1",

zk.connect: "localhost:2181",

sys.store.provider.local.path="/mypath"

}

## 存储插件配置（分为如下几个方面）

插件配置基础

工作区

Drill默认输入格式

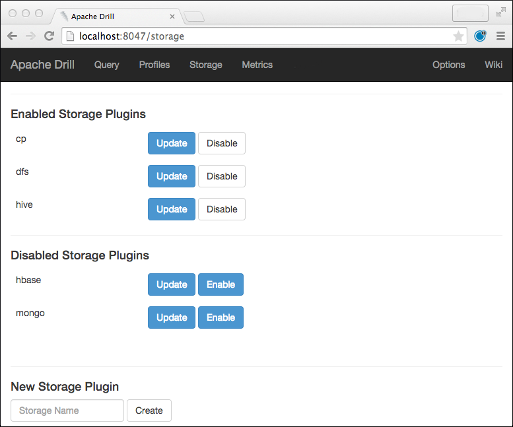
### 插件基础配置

当您在具有多个Drill安装的群集中的一个Drill节点上添加或更新存储插件配置时，Drill会将信息广播到其他Drill节点以同步存储插件配置。添加或更新存储插件配置时，不需要重新启动任何Drillbits。

使用drill Web控制台

你可以使用Dril web控制台更新或添加新的存储插件配置。drillbits需要运行以启动Web控制台。

要创建一个名称和新的配置：

1. [启动钻壳](http://drill.apache.org/docs/starting-drill-on-linux-and-mac-os-x/)。
2. [启动Web控制台](http://drill.apache.org/docs/starting-the-web-console/)。
3. 在存储选项卡上，在**新建存储插件中**输入名称。使用Drill注册的每个配置都必须具有不同的名称。名称区分大小写。  
   

**注意**

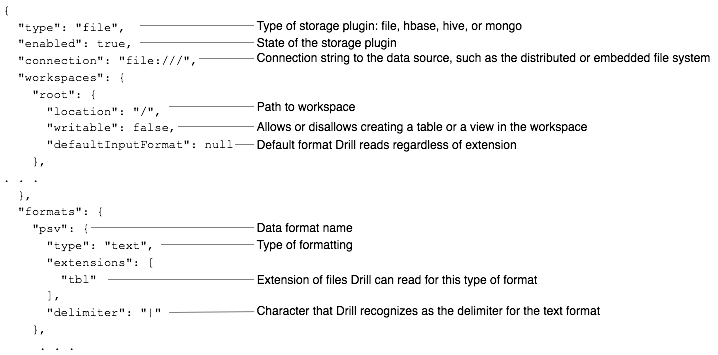
（Web）URL根据您的安装和配置而有所不同。

4，点击**创建**。

5，在配置中，如果可能，请使用JSON格式来修改现有配置的副本。  
使用现有配置的副本可以降低JSON编码错误的风险。使用下一节中的“存储插件属性”表作为进行典型修改的指南。

6，点击**创建**。

存储插件属性

下图显示了典型dfs存储插件配置的关键属性：  


**属性和定义列表**

**List of Attributes and Definitions**

The following table describes the attributes you configure for storage plugins installed with Drill.

|  |  |  |  |
| --- | --- | --- | --- |
| **Attribute** | **Example Values** | **Required** | **Description** |
| "type" | "file" "hbase" "hive" "mongo" | yes | A valid storage plugin type name. |
| "enabled" | true false | yes | State of the storage plugin. |
| "connection" | "classpath:///" "file:///" "mongodb://localhost:27017/" "hdfs://" | implementation-dependent | The type of distributed file system, such as HDFS, Amazon S3, or files in your file system, and an address/path name. |
| "workspaces" | null "logs" | no | One or more unique workspace names. If a workspace name is used more than once, only the last definition is effective. |
| "workspaces". . . "location" | "location": "/Users/johndoe/mydata" "location": "/tmp" | no | Full path to a directory on the file system. |
| "workspaces". . . "writable" | true false | no | One or more unique workspace names. If defined more than once, the last workspace name overrides the others. |
| "workspaces". . . "defaultInputFormat" | null "parquet" "csv" "json" | no | Format for reading data, regardless of extension. Default = "parquet" |
| "formats" | "pcap" "psv" "csv" "tsv" "parquet" "json" "avro" "maprdb" "sequencefile" | yes | One or more valid file formats for reading. Drill detects formats of some files; others require configuration. The maprdb format is in installations of the mapr-drill package. |
| "formats" . . . "type" | "pcap" "text" "parquet" "json" "maprdb" "avro" "sequencefile" | yes | Format type. You can define two formats, csv and psv, as type "Text", but having different delimiters. |
| formats . . . "extensions" | ["csv"] | format-dependent | File name extensions that Drill can read. |
| "formats" . . . "delimiter" | "\n" "\r" "\t" "\r\n" "," | format-dependent | Sequence of one or more characters that signifies the end of a line of text and the start of a new line in a delimited text file, such as CSV. Drill treats \n as the standard line delimiter. As of Drill 1.8, Drill supports multi-byte delimiters, such as \r\n. Use a 4-digit hex code syntax \uXXXX for a non-printable delimiter. |
| "formats" . . . "quote" | """ | no | A single character that starts/ends a value in a delimited text file. |
| "formats" . . . "escape" | "`" | no | A single character that escapes a quotation mark inside a value. |
| "formats" . . . "comment" | "#" | no | The line decoration that starts a comment line in the delimited text file. |
| "formats" . . . "skipFirstLine" | true | no | To include or omit the header when reading a delimited text file. Set to true to avoid reading headers as data. |
| "formats" . . . "extractHeader" | true | no | Set to true to extract and use headers as column names when reading a delimited text file, false otherwise. Ensure skipFirstLine is not true when extractHeader=false. |

使用格式属性

您可以formats在存储插件配置的区域中设置格式属性，例如skipFirstLine 。设置文本文件（如CSV）的sys.options属性时，还需要将该属性设置exec.storage.enable\_new\_text\_reader为true（默认值）。有关使用文本文件格式的更多信息和示例，请参阅[“文本文件：CSV，TSV，PSV”](http://drill.apache.org/docs/text-files-csv-tsv-psv/)。

**使用格式属性作为表函数参数**

在Drill版本1.4及更高版本中，您还可以在每个查询的基础上设置上面定义的格式属性。要将参数传递给格式插件，请使用表函数语法：

select a, b from table({table function name}(parameters))

的table function name是表名，类型参数是格式名称，其它参数是所述格式插件配置接受，如表中所定义以上（除字段extensions，其在这种情况下不适用）。

例如，要读取CSV文件并解析标题，请执行以下操作：  
select a, b from table(dfs.`path/to/data.csv`(type => 'text',

fieldDelimiter => ',', extractHeader => true))

有关格式插件配置的更多信息，请参阅[“文本文件：CSV，TSV，PSV”](http://drill.apache.org/docs/text-files-csv-tsv-psv/)。

**使用其他属性**

其它属性，如的配置size.calculator.enabled在hbase插件并configProps在hive插件，是实现相关的，并且超出了本文档的范围。

区分大小写的名称

如前所述，工作区和存储插件名称区分大小写。例如，以下查询使用存储插件名称dfs和工作区名称clicks。当您dfs.clicks在SQL语句中引用时，请使用定义的情况：

0: jdbc:drill:> USE dfs.clicks;

例如，使用小写字母定义存储插件和工作空间名称后，在查询中使用大写字母不起作用。

存储插件REST API

如果您需要将存储插件配置添加到Drill并且不想使用Web浏览器，则可以使用[Drill REST API](http://drill.apache.org/docs/rest-api/#get-status-threads)创建存储插件配置。使用POST请求并传递两个属性：

* 名称  
  存储插件配置名称。
* config   
  在Web控制台中输入的属性设置。

例如，该命令创建一个名为myplugin的存储插件，用于读取位于文件系统根目录下的未知类型的文件：

curl -X POST -H "Content-Type: application/json" -d '{"name":"myplugin", "config": {"type": "file", "enabled": false, "connection": "file:///", "workspaces": { "root": { "location": "/", "writable": false, "defaultInputFormat": null}}, "formats": null}}' http://localhost:8047/storage/myplugin.json

此示例假定HTTPS尚未启用。

**引导存储插件**

建议使用REST API以编程方式向Drill添加存储插件配置。仅在分布式环境中使用的替代方法是自举。您可以创建[bootstrap-storage-plugins.json](https://github.com/apache/drill/blob/master/contrib/storage-hbase/src/main/resources/bootstrap-storage-plugins.json)文件，并在启动Drill时将其包含在类路径中。Drill启动时加载存储插件配置。

目前，引导存储插件配置仅在集群中的第一个Drillbit首次启动时起作用。配置存储在ZooKeeper中，防止Drill再次获取bootstrap-storage-plugins.json。

集群启动后，您必须使用REST API或钻取Web控制台来添加存储插件配置。或者，您可以通过将该插件的json文件上传到zookeeper安装的/ drill目录来修改ZooKeeper中的条目，或者只要删除/ drill目录（如果没有要保留的配置属性）即可修改该条目。

如果使用bootstrap-storage-plugins.json文件加载HBase存储插件配置，并且未安装HBase，则在执行查询时可能会遇到延迟。在HBase插件配置的配置块中配置[HBase客户端超时](http://hbase.apache.org/book.html/#config.files)和重试设置。

### 工作区

您可以在[存储插件配置中](http://drill.apache.org/docs/plugin-configuration-basics/)定义一个或多个工作区。工作空间定义本地或分布式文件系统的子目录中文件的位置。当您运行查询时，Drill将搜索工作区以查找数据。隐藏的默认工作区dfs.default指向文件系统的根目录。

以下dfs存储插件配置显示了定义的工作区的一些示例：

{

"type": "file",

"enabled": true,

"connection": "file:///",

"workspaces": {

"root": {

"location": "/",

"writable": false,

"defaultInputFormat": null

},

"tmp": {

"location": "/tmp",

"writable": true,

"defaultInputFormat": null

},

"emp": {

"location": "/Users/user1/emp",

"writable": true,

"defaultInputFormat": null

},

"donuts": {

"location": "/Users/user1/donuts",

"writable": true,

"defaultInputFormat": null

},

"sales": {

"location": "/Users/user1/sales",

"writable": true,

"defaultInputFormat": null

}

},

配置工作区以包含子目录可简化查询，这在重复查询

配置工作区以包含子目录可简化查询，这在重复查询相同文件时非常重要。在工作空间location属性中配置长路径名称之后，而不是使用数据源的完整路径名称，在FROM子句中使用点符号。

<workspace name>.`<location>`

哪里<location>是一个子目录的路径名称，如[“查询甜甜圈示例”中](http://drill.apache.org/docs/file-system-storage-plugin/#querying-donuts-example)/users/max/drill/json所示的双引号[。](http://drill.apache.org/docs/file-system-storage-plugin/#querying-donuts-example)

要在未将缺省模式名称设置为存储插件配置时查询数据源，请包含插件名称。此语法假定您没有发出USE语句来连接到定义数据位置的存储插件：

<plugin>.<workspace name>.`<location>`

重写 **dfs.default**

您可能希望在用户无权访问根目录的情况下覆盖隐藏的默认工作区。将以下工作空间条目添加到dfs存储插件配置以覆盖默认工作空间：

"default": {

"location": "</directory/path>",

"writable": true,

"defaultInputFormat": null

}

没有工作区的Hive和HBase

虽然Hive数据库在Drill中显示为工作空间，但不能在使用Apache Drill安装的插件hive和hbase插件的配置中包含工作 空间。每个hive存储插件配置都包含一个default指向Hive Metastore 的工作空间。在hive default工作区中查询文件和表格时，可以从查询中省略工作区名称。

例如，您可以default workspace 使用以下任一查询在Hive表上发出查询，并获得相同的结果：

**例**

SELECT \* FROM hive.customers LIMIT 10;

SELECT \* FROM hive.`default`.customers LIMIT 10;

**注意**

默认是保留字。在作为后面标记中的标识符时，必须附上保留字。

由于HBase存储插件不适合工作区，因此可以使用以下查询：

SELECT \* FROM hbase.customers LIMIT 10;

Drill 默认的输入格式：

### 钻取默认输入格式

您可以定义默认的输入格式来告诉Drill在文件系统中的工作区中存在哪种文件类型。

通常，Drill 在搜索工作区时根据文件扩展名和幻数确定文件类型。魔术数字是Drill用来识别Parquet文件的文件签名。如果Drill无法根据文件扩展名或幻数识别文件类型，则查询失败。定义默认的输入格式可以防止查询在Drill无法确定文件类型的情况下失败。

如果您没有在工作区中定义默认的文件类型，或者错误地定义了默认的文件类型，并且Drill无法确定没有此信息的文件类型，则查询失败。您可以为每个工作区定义一个默认的输入格式。您可以为Drill支持的任何文件类型定义默认的输入格式。目前，Drill支持以下输入类型：

* Avro公司
* CSV，TSV或PSV
* 实木复合地板
* JSON
* Hadoop序列文件

您必须先定义一个[工作区，](http://drill.apache.org/docs/workspaces)然后才能定义默认的输入格式。

为工作区定义默认的输入格式：

1. 导航到[钻取Web控制台](http://drill.apache.org/docs/plugin-configuration-basics/#using-the-drill-web-console)。在连接到钻取Web控制台之前，Drillbit进程必须在节点上运行。
2. 在工具栏中选择**存储**。
3. 单击要为其定义工作空间默认输入格式的存储插件配置旁边的**更新**。
4. 在“配置”区域中找到工作区，然后将该defaultInputFormat属性更改为任何受支持的文件类型。

**定义默认输入格式的示例**

{

"type": "file",

"enabled": true,

"connection": "hdfs://",

"workspaces": {

"root": {

"location": "/drill/testdata",

"writable": false,

"defaultInputFormat": "csv"

},

"local" : {

"location" : "/max/proddata",

"writable" : true,

"defaultInputFormat" : "json"

}

## 文件存储插件

### 文件系统存储插件

您可以注册将Drill连接到本地文件系统或Hadoop中注册的分布式文件系统core-site.xml（例如S3或HDFS）的存储插件配置。默认情况下，Apache Drill包含一个存储插件配置dfs，默认情况下，该插件配置指向您计算机上的本地文件系统。

### 将Drill连接到文件系统

在Drill群集中，通常不会查询本地文件系统，而是将文件放在分布式文件系统上。目前，在连接多个Drillbits时，您需要使用分布式文件系统以获取完整一致的查询结果，通过将文件复制到每个节点来模拟分布式文件系统，或者使用Amazon卷（例如Amazon Elastic File System）。

您可以配置存储插件工作空间的连接属性，以将Drill连接到分布式文件系统。例如，以下连接属性将Drill连接到来自客户端的HDFS集群：

"connection": "hdfs://<IP Address>:<Port>/"

要从群集上的节点查询HDFS上的文件，只需将存储插件中的连接更改为file:///即可。hdfs://dfs

要将dfs存储插件配置更改为指向不同的本地或分布式文件系统，请使用connection以下示例中所示的属性。

* 本地文件系统示例：

{

"type": "file",

"enabled": true,

"connection": "file:///",

"workspaces": {

"root": {

"location": "/user/max/donuts",

"writable": false,

"defaultInputFormat": null

}

},

"formats" : {

"json" : {

"type" : "json"

}

}

}

* 分布式文件系统示例：

{

"type" : "file",

"enabled" : true,

"connection" : "hdfs://10.10.30.156:8020/",

"workspaces" : {

"root" : {

"location" : "/user/root/drill",

"writable" : true,

"defaultInputFormat" : null

}

},

"formats" : {

"json" : {

"type" : "json"

}

}

}

要连接到Hadoop文件系统，需要包含名称节点的IP地址和端口号。

### 查询甜甜圈示例

以下示例显示了一个名为的工作空间的文件类型存储插件配置json\_files。配置点钻取到/users/max/drill/json/本地文件系统中的 目录(dfs)：

{

"type" : "file",

"enabled" : true,

"connection" : "file:///",

"workspaces" : {

"json\_files" : {

"location" : "/users/max/drill/json/",

"writable" : false,

"defaultInputFormat" : json

}

},

该connection配置中的参数是“ file:///”，将Drill连接到本地文件系统。

要查询示例json\_files工作区中的文件，可以发出以下USE 命令来告诉Drill使用json\_files工作空间，该工作空间包含在dfs 您发出的每个查询的配置中：(没有试验成功)

USE dfs.json\_files;

SELECT \* FROM `donuts.json` WHERE type='frosted'

如果json\_files工作空间不存在，查询将不得不包含文件的完整文件路径名称donuts.json：

SELECT \* FROM dfs.`/users/max/drill/json/donuts.json` WHERE type='frosted';

## HBase存储插件

当使用安装了Drill的HBase存储插件将Drill连接到HBase数据源时，您需要指定一个ZooKeeper quorum。Drill支持HBase版本0.98。

要查看或更改HBase存储插件配置，请使用[钻取Web控制台](http://drill.apache.org/docs/plugin-configuration-basics/#using-the-drill-web-console)。在Web控制台中，选择**存储**选项卡，然后单击存储插件配置的**更新**按钮hbase。以下示例显示了一个典型的HBase存储插件：

{

"type": "hbase",

"config": {

"hbase.zookeeper.quorum": "10.10.100.62,10.10.10.52,10.10.10.53",

"hbase.zookeeper.property.clientPort": "2181"

},

"size.calculator.enabled": false,

"enabled": true

}

## Hive存储插件

### 2016年1月8日

Drill 1.1和更高版本支持Hive 1.0。要使用自定义SerDes或InputFormat / OutputFormat访问Hive表，所有运行Drillbits的节点必须JAR在<drill\_installation\_directory>/jars/3rdparty文件夹中具有SerDes或InputFormat / OutputFormat 文件 。

您可以按照以下方式通过配置Hive存储插件来运行Hive查询：

* [将Drill连接到Hive远程Metastore](http://drill.apache.org/docs/hive-storage-plugin/#connect-drill-to-the-hive-remote-metastore-directly)
* [连接到Hive嵌入式Metastore](http://drill.apache.org/docs/hive-storage-plugin/#connect-to-the-hive-embedded-metastore)

您通过选择[钻取Web控制台](http://drill.apache.org/docs/plugin-configuration-basics/#using-the-drill-web-console)上的**存储选项卡**来更新Hive存储插件。从Drill Web控制台的已禁用存储插件列表中，单击旁边的**更新**。默认的Hive存储插件配置如下所示：hive

{

"type": "hive",

"enabled": false,

"configProps": {

"hive.metastore.uris": "",

"javax.jdo.option.ConnectionURL": "jdbc:derby:;databaseName=../sample-data/drill\_hive\_db;create=true",

"hive.metastore.warehouse.dir": "/tmp/drill\_hive\_wh",

"fs.default.name": "file:///",

"hive.metastore.sasl.enabled": "false"

}

}

### 将Drill连接到Hive Remote Metastore

Hive Metastore作为Hive之外的单独服务运行。钻子可以通过Thrift查询蜂房Metore。Metastore服务通过JDBC与Hive数据库进行通信。

按照下一节中的步骤将Drill指向Hive Metastore服务地址。在Hive存储插件配置中提供连接参数以配置到Drill的连接。此时，如果您没有使用HBaseStorageHandler查询Hive表，您完成了配置Hive存储插件; 否则，如果使用HBaseStorageHandler查询Hive表，则需要添加ZooKeeper仲裁和端口属性。HBaseStorageHandler需要这些属性。Drill使用这些属性发现HBase服务。如果您使用HBase存储插件，则假设您希望Hive查询访问相同的HBase源，请使用Hive存储插件和HBase存储插件中相同的ZooKeeper仲裁和端口属性。

**注意**

在注册Hive Metastore之前，验证Hive Metastore服务是否正在运行。

#### Hive远程Metastore配置

将Drill连接到远程蜂房Metastore：

1. 发出以下命令以在以下网址中指定的系统上启动Hive Metastore服务hive.metastore.uris：  
   hive --service metastore
2. 在“ [钻取Web控制台”中](http://drill.apache.org/docs/plugin-configuration-basics/#using-the-drill-web-console)，选择“ **存储”**选项卡。
3. 在“钻取Web控制台”中禁用的存储插件列表中，单击**“**旁边的**更新”**hive。
4. 在配置窗口中，添加Thrift URI和端口hive.metastore.uris。例如：
5. ...
6. "configProps": {
7. "hive.metastore.uris": "thrift://<host>:<port>",
8. ...
9. 更改文件的默认位置以适合您的环境; 例如，将"fs.default.name"属性从更改"file:///"为hdfs://或hdfs://<host name>:<port>。在"fs.default.name"包括主机名和端口必须指向主要控制节点。例如：
10. {
11. "type": "hive",
12. "enabled": false,
13. "configProps": {
14. "hive.metastore.uris": "thrift://hdfs41:9083",
15. "hive.metastore.sasl.enabled": "false",
16. "fs.default.name": "hdfs://10.10.10.41/"
17. }
18. }
19. 如果没有查询使用HBaseStorageHandler蜂房表，请跳过此步骤; 否则，请添加ZooKeeper仲裁主机和ZooKeeper端口的名称，例如2181。
20. {
21. "type": "hive",
22. "enabled": false,
23. "configProps": {
24. .
25. .
26. .
27. "hbase.zookeeper.quorum": "zkhost1,zkhost2,zkhost3",
28. "hbase.zookeeper.property.clientPort:" "2181"
29. }
30. }
31. 点击**启用**。

### 连接到Hive嵌入式Metastore

Hive Metastore配置嵌入在Drill过程中。仅在运行单个Drillbit的群集中配置嵌入式元数据，仅用于测试目的。不要将Hive Metastore嵌入到生产系统中。

在“钻取Web控制台”中提供Metastore数据库配置设置。在配置嵌入式Hive Metastore之前，请验证用于连接到Hive Metastore /<drill installation directory>/lib/.的驱动程序是否位于Drill classpath中。如果驱动程序不在该驱动程序中，请将驱动程序复制到/<drill

installation directory>/libDrill节点上。有关存储类型和配置的更多信息，请参阅[“Hive Metastore管理”](https://cwiki.apache.org/confluence/display/Hive/AdminManual+MetastoreAdmin)。

#### Hive嵌入式Metastore配置

要配置嵌入式Hive Metastore，请完成以下步骤：

1. 在“ [钻取Web控制台”中](http://drill.apache.org/docs/plugin-configuration-basics/#using-the-drill-web-console)，选择“ **存储”**选项卡。
2. 在已禁用的存储插件配置部分中，单击旁边的**更新**hive。
3. 在配置窗口中，添加数据库配置设置。

**例**

{

"type": "hive",

"enabled": false,

"configProps": {

"hive.metastore.uris": "",

"javax.jdo.option.ConnectionURL": "jdbc:<database>://<host:port>/<metastore database>",

"hive.metastore.warehouse.dir": "/tmp/drill\_hive\_wh",

"fs.default.name": "file:///",

"hive.metastore.sasl.enabled": "false"

}

}

1. 更改"fs.default.name"属性以指定文件的默认位置。该值需要是一个可用并能够处理文件系统请求的URI。例如，将本地文件系统URI "file:///"更改为HDFS URI：hdfs://或使用namenode将其更改为HDFS 上的路径：hdfs://<hostname>:<port>
2. 点击**启用**。

“javax.jdo.option.ConnectionURL”可能是一个存储元数据的MySQL数据库。Hive Metastore服务提供对像MySQL这样的物理数据库的访问。MySQL存储元数据。

配置Hive存储插件后，可以[查询Hive表](http://drill.apache.org/docs/querying-hive/)。

# 数据查询

查询数据

* [查询数据介绍](http://drill.apache.org/docs/query-data-introduction/)
* [查询文件系统](http://drill.apache.org/docs/querying-a-file-system/)
  + [查询文件系统介绍](http://drill.apache.org/docs/querying-a-file-system-introduction/)
  + [查询Avro文件](http://drill.apache.org/docs/querying-avro-files/)
  + [查询JSON文件](http://drill.apache.org/docs/querying-json-files/)
  + [查询镶木地板文件](http://drill.apache.org/docs/querying-parquet-files/)
  + [查询纯文本文件](http://drill.apache.org/docs/querying-plain-text-files/)
  + [查询目录](http://drill.apache.org/docs/querying-directories/)
  + [查询序列文件](http://drill.apache.org/docs/querying-sequence-files/)
* [查询HBase](http://drill.apache.org/docs/querying-hbase/)
* [查询复杂数据](http://drill.apache.org/docs/querying-complex-data/)
  + [查询复杂数据介绍](http://drill.apache.org/docs/querying-complex-data-introduction/)
  + [示例数据：甜甜圈](http://drill.apache.org/docs/sample-data-donuts/)
  + [选择平面数据](http://drill.apache.org/docs/selecting-flat-data/)
  + [使用SQL函数，子句和联接](http://drill.apache.org/docs/using-sql-functions-clauses-and-joins/)
  + [为列选择嵌套数据](http://drill.apache.org/docs/selecting-nested-data-for-a-column/)
  + [在嵌套数据中选择多列](http://drill.apache.org/docs/selecting-multiple-columns-within-nested-data/)
* [查询蜂巢](http://drill.apache.org/docs/querying-hive/)
* [查询信息模式](http://drill.apache.org/docs/querying-the-information-schema/)
* [查询系统表](http://drill.apache.org/docs/querying-system-tables/)
* [在“钻取Web控制台”中监视和取消查询](http://drill.apache.org/docs/monitoring-and-canceling-queries-in-the-drill-web-console/)

## 查询数据介绍

您可以查询本节的子主题中所述的本地和分布式文件系统，Hive，HBase数据，复杂数据，INFORMATION SCHEMA和系统表。

查询指定数据源位置并包含数据转换。

### 指定数据源位置

可选的[USE语句](http://drill.apache.org/docs/use)针对特定的[存储插件](http://drill.apache.org/docs/connect-a-data-source-introduction/)运行后续查询。USE语句通常保存在FROM语句中输入一些存储插件信息。如果您省略USE语句，请指定一个存储插件（如dfs），以及可选的工作空间（如缺省值）以及使用点符号和反标记的数据源路径。例如：

dfs.`default`.`/Users/drill-user/apache-drill-1.1.0/log/sqlline\_queries.json`;

### 投射数据

在某些情况下，Drill将模式较少的数据隐式转换为正确类型的数据。在这种情况下，您不需要[将数据转换](http://drill.apache.org/docs/supported-data-types/#casting-and-converting-data-types)为其他类型。数据的文件格式和查询的性质决定了转换或转换的需求。铸造中的差异取决于数据源。

例如，您必须将"100"JSON文件中的字符串转换为整数才能应用数学函数或聚合函数。

要使用Drill查询HBase数据，请按照[“查询HBase”](http://drill.apache.org/docs/querying-hbase/)一节所述，将HBase表的每一列转换为/从字节数组转换为SQL数据类型。使用[CONVERT\_TO或CONVERT\_FROM](http://drill.apache.org/docs/data-type-conversion/#convert_to-and-convert_from)函数来执行HBase数据的转换。

### 疑难解答查询

除了在Drill Shell中以交互方式测试查询，并检查错误消息之外，还可以使用[EXPLAIN命令](http://drill.apache.org/docs/explain/)分析错误并对不运行的查询进行疑难解答。例如，如果遇到转换错误，则查询计划文本可能会帮助您隔离问题。

0: jdbc:drill:zk=local> !set maxwidth 10000

0: jdbc:drill:zk=local> explain plan for select ... ;

[钻壳命令](http://drill.apache.org/docs/configuring-the-drill-shell/)包括!set <set variable> <value>增加默认的文本显示（字符数）。默认情况下，大多数计划输出是隐藏的。

### 查询语法提示

使用Drill查询数据时请记住以下提示：

* 在SQL语句的末尾添加分号，除非发出[Drill shell命令](http://drill.apache.org/docs/configuring-the-drill-shell/)。  
  Example:！set maxwidth 10000`
* 在[关键字](http://drill.apache.org/docs/reserved-keywords)，特殊字符以及SQL无法解析的[标识符](http://drill.apache.org/docs/lexical-structure/#identifier)（如关键字default和包含正斜杠字符的路径）中使用反引号：示例：SELECT \* FROM dfs.`default`.`/Users/drilluser/apache-drill-1.1.0/sample-data/nation.parquet`;
* 选择全部（SELECT \*）无模式数据时，返回列的顺序可能与存储的顺序不同，并且可能因查询而异。

## 查询文件系统

* [查询文件系统介绍](http://drill.apache.org/docs/querying-a-file-system-introduction/)
* [查询Avro文件](http://drill.apache.org/docs/querying-avro-files/)
* [查询JSON文件](http://drill.apache.org/docs/querying-json-files/)
* [查询镶木地板文件](http://drill.apache.org/docs/querying-parquet-files/)
* [查询纯文本文件](http://drill.apache.org/docs/querying-plain-text-files/)
* [查询目录](http://drill.apache.org/docs/querying-directories/)
* [查询序列文件](http://drill.apache.org/docs/querying-sequence-files/)

### 查询镶木地板文件

2016年11月21日

Drill安装包括一个sample-data可以查询的具有Parquet文件的目录。使用SQL查询region.parquet和nation.parquet文件的sample-data目录。

**注意**

钻孔安装位置可能与此处使用的示例不同。

这些例子假设Drill是以[嵌入模式安装的](http://drill.apache.org/docs/installing-drill-in-embedded-mode)。如果您以分布模式安装Drill，或者您的sample-data目录与示例中使用的位置不同。sample-data在运行查询之前将目录更改为正确的位置。

#### 区域文件

要查看region.parquet文件中的数据，请发出以下查询：

SELECT \* FROM dfs.`<path-to-installation>/apache-drill-<version>/sample-data/region.parquet`;

该查询返回以下结果：

+--------------+--------------+-----------------------+

| R\_REGIONKEY | R\_NAME | R\_COMMENT |

+--------------+--------------+-----------------------+

| 0 | AFRICA | lar deposits. blithe |

| 1 | AMERICA | hs use ironic, even |

| 2 | ASIA | ges. thinly even pin |

| 3 | EUROPE | ly final courts cajo |

| 4 | MIDDLE EAST | uickly special accou |

+--------------+--------------+-----------------------+

5 rows selected (0.272 seconds)

#### 民族文件

如果您按照10分钟内的Apache Drill说明在嵌入模式下安装Drill，则不同操作系统之间的parquet文件路径会有所不同。

要查看nation.parquet文件中的数据，请发出适合您的操作系统的查询：

SELECT \* FROM dfs.`<path-to-installation>/sample-data/nation.parquet`;

该查询返回以下结果：

+--------------+-----------------+--------------+-----------------------+

| N\_NATIONKEY | N\_NAME | N\_REGIONKEY | N\_COMMENT |

+--------------+-----------------+--------------+-----------------------+

| 0 | ALGERIA | 0 | haggle. carefully f |

| 1 | ARGENTINA | 1 | al foxes promise sly |

| 2 | BRAZIL | 1 | y alongside of the p |

| 3 | CANADA | 1 | eas hang ironic, sil |

| 4 | EGYPT | 4 | y above the carefull |

| 5 | ETHIOPIA | 0 | ven packages wake qu |

| 6 | FRANCE | 3 | refully final reques |

| 7 | GERMANY | 3 | l platelets. regular |

| 8 | INDIA | 2 | ss excuses cajole sl |

| 9 | INDONESIA | 2 | slyly express asymp |

| 10 | IRAN | 4 | efully alongside of |

| 11 | IRAQ | 4 | nic deposits boost a |

| 12 | JAPAN | 2 | ously. final, expres |

| 13 | JORDAN | 4 | ic deposits are blit |

| 14 | KENYA | 0 | pending excuses hag |

| 15 | MOROCCO | 0 | rns. blithely bold c |

| 16 | MOZAMBIQUE | 0 | s. ironic, unusual a |

| 17 | PERU | 1 | platelets. blithely |

| 18 | CHINA | 2 | c dependencies. furi |

| 19 | ROMANIA | 3 | ular asymptotes are |

| 20 | SAUDI ARABIA | 4 | ts. silent requests |

| 21 | VIETNAM | 2 | hely enticingly expr |

| 22 | RUSSIA | 3 | requests against th |

| 23 | UNITED KINGDOM | 3 | eans boost carefully |

| 24 | UNITED STATES | 1 | y final packages. sl |

+--------------+-----------------+--------------+-----------------------+

25 rows selected (0.102 seconds)

### 查询目录

2016年11月21日

您可以将多个文件存储在一个目录中，并将它们查询为单个实体。您不必显式加入文件。这些文件必须是兼容的，因为它们必须具有相同的数据类型和列。没有可比数据类型的隐藏文件可能会导致[表未找到](http://drill.apache.org/docs/troubleshooting/#table-not-found)错误。您可以查询具有Drill支持格式的文件的目录，例如JSON，Parquet或文本文件。

例如，假设一个testdata目录包含两个具有相同结构的文件：plays.csv和moreplays.csv。第一个文件包含7个记录，第二个文件包含3个记录。以下查询返回由第一列排序的两个文件的“联合”：

0: jdbc:drill:zk=local> SELECT COLUMNS[0] AS `Year`, COLUMNS[1] AS Play

FROM dfs.`/Users/brumsby/drill/testdata` order by 1;

+------------+------------------------+

| Year | Play |

+------------+------------------------+

| 1594 | Comedy of Errors |

| 1595 | Romeo and Juliet |

| 1596 | The Merchant of Venice |

| 1599 | As You Like It |

| 1599 | Hamlet |

| 1601 | Twelfth Night |

| 1606 | Macbeth |

| 1606 | King Lear |

| 1609 | The Winter's Tale |

| 1610 | The Tempest |

+------------+------------------------+

10 rows selected (0.296 seconds)

您可以深入查看并自动查询子目录。例如，假设您有一个日志目录，其中包含每个年份的子目录和每个月份（1到12）的子目录。月份目录包含JSON文件。

[root@ip-172-16-1-200 logs]# pwd

/mapr/drilldemo/labs/clicks/logs

[root@ip-172-16-1-200 logs]# ls

2012 2013 2014

[root@ip-172-16-1-200 logs]# cd 2013

[root@ip-172-16-1-200 2013]# ls

1 10 11 12 2 3 4 5 6 7 8 9

您可以通过在查询中引用文件系统来查询所有这些文件或子集。例如，以下查询计算目录中所有文件中的记录数2013：

0: jdbc:drill:> SELECT COUNT(\*) FROM MFS.`/mapr/drilldemo/labs/clicks/logs/2013` ;

+------------+

| EXPR$0 |

+------------+

| 24000 |

+------------+

1 row selected (2.607 seconds)

您还可以使用变量dir0，dir1等，指的子目录在您的工作空间路径。例如，假定bob.logdata 是一个指向工作空间logs目录，它包含多个子目录：2012，2013，和2014。以下查询约束名为的子目录内的文件2013。变量dir0是指从日志dir1到下一个级别的第一个级别，依此类推。

0: jdbc:drill:> USE bob.logdata;

+------------+-----------------------------------------+

| ok | summary |

+------------+-----------------------------------------+

| true | Default schema changed to 'bob.logdata' |

+------------+-----------------------------------------+

1 row selected (0.305 seconds)

0: jdbc:drill:> SELECT \* FROM logs WHERE dir0='2013' LIMIT 10;

+------------+------------+------------+------------+------------+------------+------------+------------+------------+-------------+

| dir0 | dir1 | trans\_id | date | time | cust\_id | device | state | camp\_id | keywords |

+------------+------------+------------+------------+------------+------------+------------+------------+------------+-------------+

| 2013 | 2 | 12115 | 02/23/2013 | 19:48:24 | 3 | IOS5 | az | 5 | who's |

| 2013 | 2 | 12127 | 02/26/2013 | 19:42:03 | 11459 | IOS5 | wa | 10 | for |

| 2013 | 2 | 12138 | 02/09/2013 | 05:49:01 | 1 | IOS6 | ca | 7 | minutes |

| 2013 | 2 | 12139 | 02/23/2013 | 06:58:20 | 1 | AOS4.4 | ms | 7 | i |

| 2013 | 2 | 12145 | 02/10/2013 | 10:14:56 | 10 | IOS5 | mi | 6 | wrong |

| 2013 | 2 | 12157 | 02/15/2013 | 02:49:22 | 102 | IOS5 | ny | 5 | want |

| 2013 | 2 | 12176 | 02/19/2013 | 08:39:02 | 28 | IOS5 | or | 0 | and |

| 2013 | 2 | 12194 | 02/24/2013 | 08:26:17 | 125445 | IOS5 | ar | 0 | say |

| 2013 | 2 | 12236 | 02/05/2013 | 01:40:05 | 10 | IOS5 | nj | 2 | sir |

| 2013 | 2 | 12249 | 02/03/2013 | 04:45:47 | 21725 | IOS5 | nj | 5 | no |

+------------+------------+------------+------------+------------+------------+------------+------------+------------+-------------+

10 rows selected (0.583 seconds)

## 查询信息模式

2017年4月5日

当您使用Drill连接到多个数据源时，您需要一个简单的机制来发现每个数据源包含的内容。信息模式是ANSI标准的元数据表集，您可以查询这些元数据表以返回有关所有Drill数据源（或模式）的信息。数据源可能是数据库或文件系统; 在这方面他们都被称为“模式”。您可以查询以下INFORMATION\_SCHEMA表：

* SCHEMATA
* CATALOGS
* TABLES
* 列
* VIEWS

### SCHEMATA

SCHEMATA表包含CATALOG\_NAME和SCHEMA\_NAME列。为了在BI工具中提供最大的灵活性，Drill支持的唯一目录是 DRILL。

SELECT CATALOG\_NAME, SCHEMA\_NAME as all\_my\_data\_sources FROM INFORMATION\_SCHEMA.SCHEMATA ORDER BY SCHEMA\_NAME;

+--------------+---------------------+

| CATALOG\_NAME | all\_my\_data\_sources |

+--------------+---------------------+

| DRILL | INFORMATION\_SCHEMA |

| DRILL | cp.default |

| DRILL | dfs.default |

| DRILL | dfs.root |

| DRILL | dfs.tmp |

| DRILL | HiveTest.SalesDB |

| DRILL | maprfs.logs |

| DRILL | sys |

+--------------+---------------------+

INFORMATION\_SCHEMA名称和关联的关键字区分大小写。您还可以通过运行SHOW DATABASES命令来返回模式列表：

SHOW DATABASES;

+-------------+

| SCHEMA\_NAME |

+-------------+

| dfs.default |

| dfs.root |

| dfs.tmp |

...

### CATALOGS

CATALOGS表只返回一行，硬编码的DRILL目录名称和描述。

### TABLES

TABLES表返回数据库中每个表或视图的表名和类型。（类型表示TABLE或VIEW。）请注意，Drill不会返回可用于在基于文件的数据源中查询的文件。而是使用SHOW FILES来浏览这些数据源。

### 列

COLUMNS表返回每个表或视图中每列的列名称和其他元数据（如数据类型）。

### VIEWS

VIEWS表返回数据库中每个视图的名称和定义。请注意，文件模式是Drill中视图的规范存储库。根据您创建视图的方式，只能在使用后才显示在“钻取”中。

### 有用的查询

运行INFORMATION\_SCHEMA.`TABLES`查询以查看数据库中的所有表和视图。TABLES是Drill中的一个保留字，需要反标（`）。

例如，以下查询标识Drill可以访问的所有表和视图：

SELECT TABLE\_SCHEMA, TABLE\_NAME, TABLE\_TYPE

FROM INFORMATION\_SCHEMA.`TABLES`

ORDER BY TABLE\_NAME DESC;

----------------------------------------------------------------

TABLE\_SCHEMA TABLE\_NAME TABLE\_TYPE

----------------------------------------------------------------

HiveTest.CustomersDB Customers TABLE

HiveTest.SalesDB Orders TABLE

HiveTest.SalesDB OrderLines TABLE

HiveTest.SalesDB USOrders VIEW

dfs.default CustomerSocialProfile VIEW

----------------------------------------------------------------

**注意**

目前，Drill只支持查询Drill视图; Hive视图还不被支持。

您可以运行一个类似的查询来标识表中的列以及这些列的数据类型：

SELECT COLUMN\_NAME, DATA\_TYPE

FROM INFORMATION\_SCHEMA.COLUMNS

WHERE TABLE\_NAME = 'Orders' AND TABLE\_SCHEMA = 'HiveTest.SalesDB' AND COLUMN\_NAME LIKE '%Total';

+-------------+------------+

| COLUMN\_NAME | DATA\_TYPE |

+-------------+------------+

| OrderTotal | Decimal |

+-------------+------------+

在此版本中，Drill将禁用DECIMAL数据类型，包括转换为DECIMAL并从Parquet和Hive读取DECIMAL类型。您可以[启用DECIMAL类型](http://drill.apache.org/docs/supported-data-types/#enabling-the-decimal-type)，但不建议这样做。

## 查询HBase

2016年7月22日

本节涵盖以下主题：

* [教程 - 查询HBase数据](http://drill.apache.org/docs/querying-hbase/#tutorial-querying-hbase-data)  
  一个简单的教程，演示如何使用Drill查询HBase数据。
* [使用HBase字节数组](http://drill.apache.org/docs/querying-hbase/#working-with-hbase-byte-arrays)  
  如何使用HBase字节数组来处理严重的应用程序
* [查询Big Endian编码数据](http://drill.apache.org/docs/querying-hbase/#querying-big-endian-encoded-data)  
  如何使用Drill 1.2及更高版本中的优化功能
* [利用HBase有序字节编码](http://drill.apache.org/docs/querying-hbase/#leveraging-hbase-ordered-byte-encoding)  
  如何使用Drill 1.2来利用[HBASE-8201 Jira](https://issues.apache.org/jira/browse/HBASE-8201)引入的新功能

### 教程 - 查询HBase数据

本教程将演示如何将Drill连接到HBase数据源，创建简单的HBase表以及使用Drill查询数据。

#### 配置HBase存储插件

要使用Drill查询HBase数据源，请首先为您的环境配置[HBase存储插件](http://drill.apache.org/docs/hbase-storage-plugin/)。

#### 创建HBase表

您可以在HBase中创建两个表格，学生和点击，您可以使用Drill进行查询。您可以使用CONVERT\_TO和CONVERT\_FROM函数将二进制文本转换为类型化数据。在[查询HBase表的](http://drill.apache.org/docs/querying-hbase/#query-hbase-tables)第4步中，使用CAST函数将二进制数据转换为INT 。在转换INT或BIGINT号码时，如果目标/源中的字节数与二进制源/目标中的数字字节数不匹配，请使用CAST。

要创建HBase表，请完成以下步骤：

1. 将以下命令传递给HBase shell来创建学生并单击HBase中的表：
2. echo "create 'students','account','address'" | hbase shell
3. echo "create 'clicks','clickinfo','iteminfo'" | hbase shell
4. 发出以下命令来创建一个testdata.txt文件：

cat > testdata.txt

1. 将以下put命令复制并粘贴到**cat**命令下面的行中。按Return键，然后按CTRL Z关闭文件。
2. put 'students','student1','account:name','Alice'
3. put 'students','student1','address:street','123 Ballmer Av'
4. put 'students','student1','address:zipcode','12345'
5. put 'students','student1','address:state','CA'
6. put 'students','student2','account:name','Bob'
7. put 'students','student2','address:street','1 Infinite Loop'
8. put 'students','student2','address:zipcode','12345'
9. put 'students','student2','address:state','CA'
10. put 'students','student3','account:name','Frank'
11. put 'students','student3','address:street','435 Walker Ct'
12. put 'students','student3','address:zipcode','12345'
13. put 'students','student3','address:state','CA'
14. put 'students','student4','account:name','Mary'
15. put 'students','student4','address:street','56 Southern Pkwy'
16. put 'students','student4','address:zipcode','12345'
17. put 'students','student4','address:state','CA'
18. put 'clicks','click1','clickinfo:studentid','student1'
19. put 'clicks','click1','clickinfo:url','http://www.google.com'
20. put 'clicks','click1','clickinfo:time','2014-01-01 12:01:01.0001'
21. put 'clicks','click1','iteminfo:itemtype','image'
22. put 'clicks','click1','iteminfo:quantity','1'
23. put 'clicks','click2','clickinfo:studentid','student1'
24. put 'clicks','click2','clickinfo:url','http://www.amazon.com'
25. put 'clicks','click2','clickinfo:time','2014-01-01 01:01:01.0001'
26. put 'clicks','click2','iteminfo:itemtype','image'
27. put 'clicks','click2','iteminfo:quantity','1'
28. put 'clicks','click3','clickinfo:studentid','student2'
29. put 'clicks','click3','clickinfo:url','http://www.google.com'
30. put 'clicks','click3','clickinfo:time','2014-01-01 01:02:01.0001'
31. put 'clicks','click3','iteminfo:itemtype','text'
32. put 'clicks','click3','iteminfo:quantity','2'
33. put 'clicks','click4','clickinfo:studentid','student2'
34. put 'clicks','click4','clickinfo:url','http://www.ask.com'
35. put 'clicks','click4','clickinfo:time','2013-02-01 12:01:01.0001'
36. put 'clicks','click4','iteminfo:itemtype','text'
37. put 'clicks','click4','iteminfo:quantity','5'
38. put 'clicks','click5','clickinfo:studentid','student2'
39. put 'clicks','click5','clickinfo:url','http://www.reuters.com'
40. put 'clicks','click5','clickinfo:time','2013-02-01 12:01:01.0001'
41. put 'clicks','click5','iteminfo:itemtype','text'
42. put 'clicks','click5','iteminfo:quantity','100'
43. put 'clicks','click6','clickinfo:studentid','student3'
44. put 'clicks','click6','clickinfo:url','http://www.google.com'
45. put 'clicks','click6','clickinfo:time','2013-02-01 12:01:01.0001'
46. put 'clicks','click6','iteminfo:itemtype','image'
47. put 'clicks','click6','iteminfo:quantity','1'
48. put 'clicks','click7','clickinfo:studentid','student3'
49. put 'clicks','click7','clickinfo:url','http://www.ask.com'
50. put 'clicks','click7','clickinfo:time','2013-02-01 12:45:01.0001'
51. put 'clicks','click7','iteminfo:itemtype','image'
52. put 'clicks','click7','iteminfo:quantity','10'
53. put 'clicks','click8','clickinfo:studentid','student4'
54. put 'clicks','click8','clickinfo:url','http://www.amazon.com'
55. put 'clicks','click8','clickinfo:time','2013-02-01 22:01:01.0001'
56. put 'clicks','click8','iteminfo:itemtype','image'
57. put 'clicks','click8','iteminfo:quantity','1'
58. put 'clicks','click9','clickinfo:studentid','student4'
59. put 'clicks','click9','clickinfo:url','http://www.amazon.com'
60. put 'clicks','click9','clickinfo:time','2013-02-01 22:01:01.0001'
61. put 'clicks','click9','iteminfo:itemtype','image'
62. put 'clicks','click9','iteminfo:quantity','10'
63. 发出以下命令将数据放入HBase：
64. cat testdata.txt | hbase shell

#### 查询HBase表

[启动Drill](http://drill.apache.org/docs/installing-drill-in-embedded-mode/)并完成以下步骤来查询您创建的HBase表。

1. 使用HBase存储插件配置。  
   USE hbase;
2. 发出以下查询以查看学生表中的数据：  
   SELECT \* FROM students;

查询返回不可用的结果。在下一步中，将数据从字节数组转换为有意义的UTF8类型。

+-------------+-----------------------+---------------------------------------------------------------------------+

| row\_key | account | address |

+-------------+-----------------------+---------------------------------------------------------------------------+

| [B@e6d9eb7 | {"name":"QWxpY2U="} | {"state":"Q0E=","street":"MTIzIEJhbGxtZXIgQXY=","zipcode":"MTIzNDU="} |

| [B@2823a2b4 | {"name":"Qm9i"} | {"state":"Q0E=","street":"MSBJbmZpbml0ZSBMb29w","zipcode":"MTIzNDU="} |

| [B@3b8eec02 | {"name":"RnJhbms="} | {"state":"Q0E=","street":"NDM1IFdhbGtlciBDdA==","zipcode":"MTIzNDU="} |

| [B@242895da | {"name":"TWFyeQ=="} | {"state":"Q0E=","street":"NTYgU291dGhlcm4gUGt3eQ==","zipcode":"MTIzNDU="} |

+-------------+-----------------------+---------------------------------------------------------------------------+

4 rows selected (1.335 seconds)

1. 发出以下查询，其中包括CONVERT\_FROM函数，将students表转换为键入的数据：
2. SELECT CONVERT\_FROM(row\_key, 'UTF8') AS studentid,
3. CONVERT\_FROM(students.account.name, 'UTF8') AS name,
4. CONVERT\_FROM(students.address.state, 'UTF8') AS state,
5. CONVERT\_FROM(students.address.street, 'UTF8') AS street,
6. CONVERT\_FROM(students.address.zipcode, 'UTF8') AS zipcode
7. FROM students;

**注意**

使用点表示法向下钻取到HBase表中的列：tablename.columnfamilyname.columnnname

查询返回的结果看起来好多了：

+------------+------------+------------+------------------+------------+

| studentid | name | state | street | zipcode |

+------------+------------+------------+------------------+------------+

| student1 | Alice | CA | 123 Ballmer Av | 12345 |

| student2 | Bob | CA | 1 Infinite Loop | 12345 |

| student3 | Frank | CA | 435 Walker Ct | 12345 |

| student4 | Mary | CA | 56 Southern Pkwy | 12345 |

+------------+------------+------------+------------------+------------+

4 rows selected (0.504 seconds)

1. 查询点击表以查看哪些学生访问了google.com：
2. SELECT CONVERT\_FROM(row\_key, 'UTF8') AS clickid,
3. CONVERT\_FROM(clicks.clickinfo.studentid, 'UTF8') AS studentid,
4. CONVERT\_FROM(clicks.clickinfo.`time`, 'UTF8') AS `time`,
5. CONVERT\_FROM(clicks.clickinfo.url, 'UTF8') AS url
6. FROM clicks WHERE clicks.clickinfo.url LIKE '%google%';
7. +------------+------------+--------------------------+-----------------------+
8. | clickid | studentid | time | url |
9. +------------+------------+--------------------------+-----------------------+
10. | click1 | student1 | 2014-01-01 12:01:01.0001 | http://www.google.com |
11. | click3 | student2 | 2014-01-01 01:02:01.0001 | http://www.google.com |
12. | click6 | student3 | 2013-02-01 12:01:01.0001 | http://www.google.com |
13. +------------+------------+--------------------------+-----------------------+
14. 3 rows selected (0.294 seconds)
15. 查询点击表以获得具有100个项目的学生的studentid。使用CONVERT\_FROM转换文本studentid和itemtype数据，但使用CAST转换整数量。
16. SELECT CONVERT\_FROM(tbl.clickinfo.studentid, 'UTF8') AS studentid,
17. CONVERT\_FROM(tbl.iteminfo.itemtype, 'UTF8'),
18. CAST(tbl.iteminfo.quantity AS INT) AS items
19. FROM clicks tbl WHERE tbl.iteminfo.quantity=100;
20. +------------+------------+------------+
21. | studentid | EXPR$1 | items |
22. +------------+------------+------------+
23. | student2 | text | 100 |
24. +------------+------------+------------+
25. 1 row selected (0.656 seconds)

#### 使用HBase字节数组

上一节中的小例子查询HBase中的小端编码数据。对于严重的应用程序，您需要了解如何使用HBase字节数组。如果您想让Drill将底层HBase行键解释为除字节数组之外的其他字段，则需要知道HBase中数据的编码。默认情况下，HBase以小端存储数据，Drill认为数据是小端，这是未排序的。下表显示了典型的行密钥ID（以字节为单位）的排序，分别以little endian和big endian编码：

| **字节表示中的ID Little Endian排序** | **十进制符号中的ID** | **字节表示法中的ID Big Endian排序** | **十进制符号中的ID** |
| --- | --- | --- | --- |
| 0 x 010000。。。000 | 1 | 0 x 00000001 | 1 |
| 0×010100。。。000 | 17 | 0 x 00000002 | 2 |
| 0 x 020000。。。000 | 2 | 0 x 00000003 | 3 |
| 。。。 |  | 0 x 00000004 | 4 |
| 0 x 050000。。。000 | 五 | 0 x 00000005 | 五 |
| 。。。 |  | 。。。 |  |
| 0×0A0000。。。000 | 10 | 0×0000000A | 10 |
|  |  | 0 x 00000101 | 17 |

#### 查询大端编码的数据

当您在大端编码数据上使用[“CONVERT\_TO和CONVERT\_FROM数据类型”](http://drill.apache.org/docs/supported-data-types/#data-types-for-convert_to-and-convert_from-functions)时，Drill优化了HBase表的扫描。Drill提供用于CONVERT\_TO和CONVERT\_FROM的\* \_BE编码类型以利用这些优化。以下是一些\* \_BE类型的例子。

* DATE\_EPOCH\_BE
* TIME\_EPOCH\_BE
* TIMESTAMP\_EPOCH\_BE
* UINT8\_BE
* BIGINT\_BE

例如，当您在大端编码数据上使用以下查询时，Drill会以一定的方式返回结果：

SELECT

CONVERT\_FROM(BYTE\_SUBSTR(row\_key, 1, 8), 'DATE\_EPOCH\_BE') d,

CONVERT\_FROM(BYTE\_SUBSTR(row\_key, 9, 8), 'BIGINT\_BE') id,

CONVERT\_FROM(tableName.f.c, 'UTF8')

FROM hbase.`TestTableCompositeDate` tableName

WHERE

CONVERT\_FROM(BYTE\_SUBSTR(row\_key, 1, 8), 'DATE\_EPOCH\_BE') < DATE '2015-06-18' AND

CONVERT\_FROM(BYTE\_SUBSTR(row\_key, 1, 8), 'DATE\_EPOCH\_BE') > DATE '2015-06-13';

此查询假定表的行键表示以big-endian格式编码的DATE\_EPOCH类型。Drill HBase插件将能够修剪扫描范围，因为行键的大端编码前缀存在条件。有关更多示例，请参阅[测试代码](https://github.com/apache/drill/blob/95623912ebf348962fe8a8846c5f47c5fdcf2f78/contrib/storage-hbase/src/test/java/org/apache/drill/hbase/TestHBaseFilterPushDown.java)。

查询HBase数据：

1. 使用[HBase存储插件](http://drill.apache.org/docs/hbase-storage-plugin/)将数据源连接到Drill 。

USE hbase;

1. 确定要查询的HBase数据的编码。询问负责创建数据的人员。
2. 根据数据的编码类型，在查询数据时，使用[“CONVERT\_TO和CONVERT\_FROM数据类型”](http://drill.apache.org/docs/supported-data-types/#data-types-for-convert_to-and-convert_from-functions)将HBase二进制表示转换为SQL类型。  
   例如，在您的Drill查询中使用CONVERT\_FROM将大端编码的行键转换为SQL BIGINT类型：

SELECT CONVERT\_FROM(BYTE\_SUBSTR(row\_key, 1, 8),'BIGINT\_BE’) FROM my\_hbase\_table;

在这个例子中，[BYTE\_SUBSTR函数](http://drill.apache.org/docs/string-manipulation/#byte_substr)分离HBase复合键的各个部分。Drill优化基于Drill 1.2和更高版本的功能，当HBase数据采用big endian格式时，将条件筛选器推送到存储层。

如果只有组合键的第一个组件以big endian格式进行编码，那么Drill可以正常查询使用组合键的HBase数据，如上例所示。如果HBase行键没有以big endian存储，请不要使用\* \_BE类型。如果要将小端字节数组转换为整数，请使用BIGINT而不是BIGINT\_BE，例如，作为CONVERT\_FROM的参数。

#### 利用HBase有序字节编码

Drill 1.2利用[HBASE-8201 Jira](https://issues.apache.org/jira/browse/HBASE-8201)引入的新功能，允许对不同数据类型进行有序字节编码。当数据作为排序字节数组存储在磁盘上时，此编码方案保留本机数据类型的排序顺序。因此，如果行键已经以OrderedBytes格式编码，Drill将能够通过HBase存储插件处理数据。

为了执行下面的查询，Drill修剪扫描范围，只包含代表[-32,59]范围的行键，从而减少读取的数据量。

SELECT

CONVERT\_FROM(t.row\_key, 'INT\_OB') rk,

CONVERT\_FROM(t.`f`.`c`, 'UTF8') val

FROM

hbase.`TestTableIntOB` t

WHERE

CONVERT\_FROM(row\_key, 'INT\_OB') >= cast(-32 as INT) AND

CONVERT\_FROM(row\_key, 'INT\_OB') < cast(59 as INT);

有关更多示例，请参阅[测试代码](https://github.com/apache/drill/blob/95623912ebf348962fe8a8846c5f47c5fdcf2f78/contrib/storage-hbase/src/test/java/org/apache/drill/hbase/TestHBaseFilterPushDown.java)。

通过利用有序的字节编码，Drill 1.2和更高版本可以在HBase的big endian数据上执行没有二级索引的条件查询。

## Hive 查询

### 2016年8月4日

这是一个简单的练习，它提供了创建Hive表和插入可以使用Drill查询的数据的步骤。在执行这些步骤之前，请下载[customers.csv](http://doc.mapr.com/download/attachments/28868943/customers.csv?version=1&modificationDate=1426874930765&api=v2)文件。

**注意**

Drill 1.8实现了DROP TABLE和DROP VIEW命令的IF EXISTS参数，使IF成为Drill中的保留字。因此，当您在Hive表的查询中使用它时，必须在Hive的IF条件函数中包含反引号。或者，您可以使用CASE语句而不是IF函数。

要创建Hive表并使用Drill查询，请完成以下步骤：

1. 发出以下命令启动Hive shell：

hive

1. 从Hive shell发出以下命令创建一个表架构：
2. hive> create table customers(FirstName string, LastName string, Company string, Address string, City string, County string, State string, Zip string, Phone string, Fax string, Email string, Web string) row format delimited fields terminated by ',' stored as textfile;
3. 发出以下命令将客户数据加载到customers表中：
4. hive> load data local inpath '/<directory path>/customers.csv' overwrite into table customers;
5. 发出quit或exit离开Hive shell。
6. 启动钻壳。
7. 向Drill发出以下查询以获取Hive表中前十个客户的名字和姓氏：
8. 0: jdbc:drill:schema=hiveremote> SELECT firstname,lastname FROM hiveremote.`customers` limit 10;`

该查询返回以下结果：

+------------+------------+

| firstname | lastname |

+------------+------------+

| Essie | Vaill |

| Cruz | Roudabush |

| Billie | Tinnes |

| Zackary | Mockus |

| Rosemarie | Fifield |

| Bernard | Laboy |

| Sue | Haakinson |

| Valerie | Pou |

| Lashawn | Hasty |

| Marianne | Earman |

+------------+------------+

10 rows selected (1.5 seconds)

0: jdbc:drill:schema=hiveremote>

### 优化实木复合表的读取

使用该store.hive.optimize\_scan\_with\_native\_readers选项优化Hive对Parquet支持的外部表的读取。当设置为TRUE时，此选项使用Drill本地读取器而不是Hive Serde接口，从而导致Parquet支持的外部表的更高性能查询。（Drill 1.2及更高版本）

store.hive.optimize\_scan\_with\_native\_readers按照[“计划和执行选项”](http://drill.apache.org/docs/planning-and-execution-options/)部分所述设置[选项](http://drill.apache.org/docs/planning-and-execution-options/)。

## 复杂查询

### 复杂查询简介

Apache Drill查询不需要事先知道您尝试访问的实际数据，无论其源系统，架构和数据类型如何。Apache Drill的甜蜜点是针对复杂数据的SQL查询工作负载 ：由各种类型的记录和字段组成的数据，而不是以可识别的关系形式（离散的行和列）的数据。在您提交查询时，Drill能够发现数据的形式。嵌套的数据格式，如JSON（JavaScript Object Notation）文件和Parquet文件不仅可以访问：Drill提供了特殊的操作符和函数，您可以使用它们深入到这些文件并提出有趣的分析问题。

这些运营商和功能包括：

* 引用嵌套的数据值
* 访问数组中的数组和数组中的重复值（数组索引）

SQL查询开发人员需要足够好地了解这些数据，以编写查询来识别目标文件中感兴趣的值。例如，编写者需要知道记录是由什么组成的，以及它的数据类型，以便可靠地请求选择列表中正确的“列”。虽然这些数据值不会在源文件中显示为列，但Drill会将它们返回到结果集中，就好像它们在表中具有可预测的列形式一样。Drill还通过将数据视为“柱状”而不是阅读和分析完整记录来优化查询。（Drill使用类似的并行执行和优化功能来处理商业化的MPP数据库。）

鉴于输入文件的基本知识，开发人员需要知道如何使用Drill提供的SQL扩展，以及如何使用它们“嵌入”嵌套的数据。下面的例子展示了如何编写针对JSON文件的简单查询和解压缩嵌套数据的有趣查询。这些示例显示了如何在标准SQL SELECT语句的上下文中使用Drill扩展。在大多数情况下，扩展使用标准JavaScript表示法来引用层次结构中的数据元素。

在你开始之前

本节中的示例使用JSON数据文件。为了编写自己的查询，您需要知道这些文件中的基本数据类型：

* 字符串（双引号内的所有数据），如"0001"or"Cake"
* 数字：整数和浮点数，如0.55或10
* 空值
* 布尔值：true，false

检查您在钻取Web控制台（dfs存储插件配置）中是否具有JSON文件的以下配置设置：

"json" : {

"type" : "json"

}

示例数据：

### 示例数据：甜甜圈

复杂的数据查询使用示例donuts.json文件。要下载此文件，请转到[钻取测试资源](https://github.com/apache/drill/blob/master/exec/java-exec/src/test/resources)页面，在文件列表中找到donuts.json，然后下载它。

这是文件中的第一个“记录”（0001）donuts.json。在查询处理方面，这个记录相当于表中的单个记录。

{

"id": "0001",

"type": "donut",

"name": "Cake",

"ppu": 0.55,

"sales": 35,

"batters":

{

"batter":

[

{ "id": "1001", "type": "Regular" },

{ "id": "1002", "type": "Chocolate" },

{ "id": "1003", "type": "Blueberry" },

{ "id": "1004", "type": "Devil's Food" }

]

},

"topping":

[

{ "id": "5001", "type": "None" },

{ "id": "5002", "type": "Glazed" },

{ "id": "5005", "type": "Sugar" },

{ "id": "5007", "type": "Powdered Sugar" },

{ "id": "5006", "type": "Chocolate with Sprinkles" },

{ "id": "5003", "type": "Chocolate" },

{ "id": "5004", "type": "Maple" }

]

}

数据由地图，数组和嵌套数组组成。名称 - 值对和嵌入的名称/值对定义每个记录的内容。例如， type: donut是一张地图。在下topping，这些对id和type值属于一个数组（在方括号内）。

### 选择平面数据

针对该donuts.json文件的一个非常简单的查询将返回五个“扁平”列的值（仅包含顶级数据的列：不包含嵌套数据）：

0: jdbc:drill:zk=local> select id, type, name, ppu

from dfs.`/Users/brumsby/drill/donuts.json`;

+-------+--------+----------------+-------+

| id | type | name | ppu |

+-------+--------+----------------+-------+

| 0001 | donut | Cake | 0.55 |

| 0002 | donut | Raised | 0.69 |

| 0003 | donut | Old Fashioned | 0.55 |

| 0004 | donut | Filled | 0.69 |

| 0005 | donut | Apple Fritter | 1.0 |

+-------+--------+----------------+-------+

5 rows selected (0.18 seconds)

请注意，dfs是模式名称，文件的路径由反引号括起来，查询必须以分号结尾。

### 使用SQL函数，子句和联接

您可以使用标准的SQL子句（如WHERE和ORDER BY）来详细说明这种简单的查询：

0: jdbc:drill:zk=local> select id, type from dfs.`/Users/brumsby/drill/donuts.json`

where id>0

order by id limit 1;

+------------+------------+

| id | type |

+------------+------------+

| 0001 | donut |

+------------+------------+

1 row selected (0.318 seconds)

您也可以使用标准语法来加入文件（或表格，或文件和表格）：

0: jdbc:drill:zk=local> select tbl1.id, tbl1.type from dfs.`/Users/brumsby/drill/donuts.json` as tbl1

join

dfs.`/Users/brumsby/drill/moredonuts.json` as tbl2

on tbl1.id=tbl2.id;

+------------+------------+

| id | type |

+------------+------------+

| 0001 | donut |

+------------+------------+

1 row selected (0.395 seconds)

也支持等同的USING语法和WHERE子句中的连接。

标准聚合函数针对JSON数据。例如：

0: jdbc:drill:zk=local> select type, avg(ppu) as ppu\_sum from dfs.`/Users/brumsby/drill/donuts.json` group by type;

+------------+------------+

| type | ppu\_sum |

+------------+------------+

| donut | 0.55 |

+------------+------------+

1 row selected (0.216 seconds)

0: jdbc:drill:zk=local> select type, sum(sales) as sum\_by\_type from dfs.`/Users/brumsby/drill/moredonuts.json` group by type;

+------------+-------------+

| type | sum\_by\_type |

+------------+-------------+

| donut | 1194 |

+------------+-------------+

1 row selected (0.389 seconds)

### 为列选择嵌套数据

以下查询显示如何访问记录中不平坦（如topping）的部分内的嵌套数据。要隔离并返回嵌套数据，请使用[n]符号，其中n是指向数组中特定位置的数字。数组使用基于0的索引，因此topping[3]指向数组中的第四个元素topping，而不是第三个元素。

0: jdbc:drill:zk=local> select topping[3] as top from dfs.`/Users/brumsby/drill/donuts.json`;

+------------+

| top |

+------------+

| {"id":"5007","type":"Powdered Sugar"} |

+------------

1 row selected (0.137 seconds)

请注意，此查询为嵌套在文件段内的所有数据生成一列topping。查询写入不解开id和type名称/值对。还要注意为列名使用别名。（如果没有别名，默认的列名就是EXPR$0。）

一些JSON文件将数组存储在数组中。如果您的数据具有此特征，则可以使用以下表示法探测内部数组：[n][n]

例如，假设JSON文件的一部分如下所示：

...

group:

[

[1,2,3],

[4,5,6],

[7,8,9]

]

...

以下查询将返回6（第二个内部数组的第三个值）。

select group[1][2]

### 在嵌套数据中选择多列

以下查询进一步提取JSON数据，从数组内部 选择特定值id和type数据值作为单独的列topping。此查询与前面的查询类似，但是它将值id和type值作为单独的列返回。

0: jdbc:drill:zk=local> select tbl.topping[3].id as record, tbl.topping[3].type as first\_topping

from dfs.`/Users/brumsby/drill/donuts.json` as tbl;

+------------+---------------+

| record | first\_topping |

+------------+---------------+

| 5007 | Powdered Sugar |

+------------+---------------+

1 row selected (0.133 seconds)

该查询还引入了对嵌套数据进行查询的典型要求：使用表别名（本例中名为tbl）。没有表别名，查询将返回一个错误，因为解析器会假定id是一个名为topping的表内的列。和所有标准SQL查询一样，选择tbl.col意味着tbl是现有表的名称（至少在查询期间），col是该表中存在的列。

## 查询系统表

Drill有一个包含系统表的sys数据库。您可以查询系统表以获取有关钻取的信息，包括钻取端口，系统上运行的钻取版本以及可用的钻取选项。在Drill中查看数据库以标识sys数据库，然后使用sys数据库查看可以查询的系统表。

### 查看钻取数据库

发出SHOW DATABASES命令以查看钻取数据库。

0: jdbc:drill:zk=10.10.100.113:5181> show databases;

+--------------------+

| SCHEMA\_NAME |

+--------------------+

| M7 |

| hive.default |

| dfs.default |

| dfs.root |

| dfs.views |

| dfs.tmp |

| dfs.tpcds |

| sys |

| cp.default |

| hbase |

| INFORMATION\_SCHEMA |

+--------------------+

11 rows selected (0.162 seconds)

钻取返回sys数据库结果。

### 使用系统数据库

发出USE命令为后续SQL请求选择sys数据库。

0: jdbc:drill:zk=10.10.100.113:5181> use sys;

+-------+----------------------------------+

| ok | summary |

+-------+----------------------------------+

| true | Default schema changed to [sys] |

+-------+----------------------------------+

1 row selected (0.101 seconds)

### 查看表格

发出SHOW TABLES命令以查看sys数据库中的表。

0: jdbc:drill:zk=10.10.100.113:5181> show tables;

+---------------+-------------+

| TABLE\_SCHEMA | TABLE\_NAME |

+---------------+-------------+

| sys | boot |

| sys | drillbits |

| sys | memory |

| sys | options |

| sys | threads |

| sys | version |

+---------------+-------------+

3 rows selected (0.934 seconds)

0: jdbc:drill:zk=10.10.100.113:5181>

### 查询系统表

查询sys数据库中的drillbits，版本，选项，启动，线程和内存表。

#### 查询钻头表。

0: jdbc:drill:zk=10.10.100.113:5181> select \* from drillbits;

+-------------------+------------+--------------+------------+---------+

| hostname | user\_port | control\_port | data\_port | current|

+-------------------+------------+--------------+------------+--------+

| qa-node115.qa.lab | 31010 | 31011 | 31012 | true |

| qa-node114.qa.lab | 31010 | 31011 | 31012 | false |

| qa-node116.qa.lab | 31010 | 31011 | 31012 | false |

+-------------------+------------+--------------+------------+---------+

3 rows selected (0.146 seconds)

* 主机  
  名称运行Drillbit服务的节点的名称。
* user\_port   
  用户端口地址，用于连接到外部客户端和钻取Web控制台的群集中的节点之间。
* control\_port   
  控制端口地址，用于Apache Drill的多节点安装的节点之间。
* data\_port   
  数据端口地址，用于Apache Drill多节点安装的节点之间。
* 当前  
  True表示Drillbit连接到运行查询的会话或客户端。这个Drillbit是本届会议的工头。

#### 查询版本表。

0: jdbc:drill:zk=10.10.100.113:5181> select \* from version;

+-------------------------------------------+--------------------------------------------------------------------+----------------------------+--------------+----------------------------+

| commit\_id | commit\_message | commit\_time | build\_email | build\_time |

+-------------------------------------------+--------------------------------------------------------------------+----------------------------+--------------+----------------------------+

| d8b19759657698581cc0d01d7038797952888123 | DRILL-3100: TestImpersonationDisabledWithMiniDFS fails on Windows | 15.05.2015 @ 05:18:03 UTC | Unknown | 15.05.2015 @ 06:52:32 UTC |

+-------------------------------------------+--------------------------------------------------------------------+----------------------------+--------------+----------------------------+

1 row selected (0.099 seconds)

* commit\_id   
  你正在运行的版本的github ID。例如，
* commit\_message   
  解释更改的消息。
* commit\_time更改  
  的日期和时间。
* build\_email   
  进行更改的人员的电子邮件地址，在本例中未知。
* build\_time   
  发布版本的时间。

#### 查询选项表。

Drill提供您可以查询的系统，会话和启动选项。

以下示例显示了有关系统选项的查询：

0: jdbc:drill:zk=10.10.100.113:5181> select \* from options where type='SYSTEM' limit 10;

+-------------------------------------------------+----------+---------+----------+-------------+-------------+-----------+------------+

| name | kind | type | status | num\_val | string\_val | bool\_val | float\_val |

+-------------------------------------------------+----------+---------+----------+-------------+-------------+-----------+------------+

| drill.exec.functions.cast\_empty\_string\_to\_null | BOOLEAN | SYSTEM | DEFAULT | null | null | false | null |

| drill.exec.storage.file.partition.column.label | STRING | SYSTEM | DEFAULT | null | dir | null | null |

| exec.errors.verbose | BOOLEAN | SYSTEM | DEFAULT | null | null | false | null |

| exec.java\_compiler | STRING | SYSTEM | DEFAULT | null | DEFAULT | null | null |

| exec.java\_compiler\_debug | BOOLEAN | SYSTEM | DEFAULT | null | null | true | null |

| exec.java\_compiler\_janino\_maxsize | LONG | SYSTEM | DEFAULT | 262144 | null | null | null |

| exec.max\_hash\_table\_size | LONG | SYSTEM | DEFAULT | 1073741824 | null | null | null |

| exec.min\_hash\_table\_size | LONG | SYSTEM | DEFAULT | 65536 | null | null | null |

| exec.queue.enable | BOOLEAN | SYSTEM | DEFAULT | null | null | false | null |

| exec.queue.large | LONG | SYSTEM | DEFAULT | 10 | null | null | null |

+-------------------------------------------------+----------+---------+----------+-------------+-------------+-----------+------------+

10 rows selected (0.216 seconds)

* 名称  
  选项的名称。
* kind   
  选项值的数据类型。
* 键入  
  输出中的选项类型：系统或会话。
* 状态选项的状态：默认或更改。
* num\_val   
  默认值，它是long或int数据类型; 否则为空。
* string\_val   
  默认值，它是一个字符串; 否则为空。
* bool\_val   
  默认值，即true或false; 否则为空。
* float\_val   
  默认值，它是double，float或long double数据类型的默认值; 否则为空。

#### 查询启动表。

0: jdbc:drill:zk=10.10.100.113:5181> select \* from boot limit 10;

+--------------------------------------+----------+-------+---------+------------+-------------------------+-----------+------------+

| name | kind | type | status | num\_val | string\_val | bool\_val | float\_val |

+--------------------------------------+----------+-------+---------+------------+-------------------------+-----------+------------+

| awt.toolkit | STRING | BOOT | BOOT | null | "sun.awt.X11.XToolkit" | null | null |

| drill.client.supports-complex-types | BOOLEAN | BOOT | BOOT | null | null | true | null |

| drill.exec.buffer.size | STRING | BOOT | BOOT | null | "6" | null | null |

| drill.exec.buffer.spooling.delete | BOOLEAN | BOOT | BOOT | null | null | true | null |

| drill.exec.buffer.spooling.size | LONG | BOOT | BOOT | 100000000 | null | null | null |

| drill.exec.cluster-id | STRING | BOOT | BOOT | null | "SKCluster" | null | null |

| drill.exec.compile.cache\_max\_size | LONG | BOOT | BOOT | 1000 | null | null | null |

| drill.exec.compile.compiler | STRING | BOOT | BOOT | null | "DEFAULT" | null | null |

| drill.exec.compile.debug | BOOLEAN | BOOT | BOOT | null | null | true | null |

| drill.exec.compile.janino\_maxsize | LONG | BOOT | BOOT | 262144 | null | null | null |

+--------------------------------------+----------+-------+---------+------------+-------------------------+-----------+------------+

10 rows selected (0.192 seconds)

* 名称  
  启动选项的名称。
* kind   
  选项值的数据类型。
* 键入  
  这总是启动。
* 状态这总是启动。
* num\_val   
  默认值，它是long或int数据类型; 否则为空。
* string\_val   
  默认值，它是一个字符串; 否则为空。
* bool\_val   
  默认值，即true或false; 否则为空。
* float\_val   
  默认值，它是double，float或long double数据类型的默认值; 否则为空。

#### 查询线程表。

0: jdbc:drill:zk=10.10.100.113:5181> select \* from threads;

+--------------------+------------+----------------+---------------+

| hostname | user\_port | total\_threads | busy\_threads |

+--------------------+------------+----------------+---------------+

| qa-node115.qa.lab | 31010 | 33 | 33 |

| qa-node114.qa.lab | 31010 | 33 | 32 |

| qa-node116.qa.lab | 31010 | 29 | 29 |

+--------------------+------------+----------------+---------------+

3 rows selected (0.618 seconds)

* 主机  
  名称运行Drillbit服务的节点的名称。
* user\_port   
  用户端口地址，用于连接到外部客户端和钻取Web控制台的群集中的节点之间。
* total\_threads节点上的最高线程数。
* busy\_threads节点上的活动线程（守护进程和非守护进程）的当前数量。

#### 查询内存表。

0: jdbc:drill:zk=10.10.100.113:5181> select \* from memory;

+--------------------+------------+---------------+-------------+-----------------+---------------------+-------------+

| hostname | user\_port | heap\_current | heap\_max | direct\_current | jvm\_direct\_current | direct\_max |

+--------------------+------------+---------------+-------------+-----------------+---------------------+-------------+

| qa-node115.qa.lab | 31010 | 443549712 | 4294967296 | 11798941 | 167772974 | 8589934592 |

| qa-node114.qa.lab | 31010 | 149948432 | 4294967296 | 7750365 | 134218542 | 8589934592 |

| qa-node116.qa.lab | 31010 | 358612992 | 4294967296 | 7750365 | 83886894 | 8589934592 |

+--------------------+------------+---------------+-------------+-----------------+---------------------+-------------+

3 rows selected (0.172 seconds)

* 主机  
  名称运行Drillbit服务的节点的名称。
* user\_port   
  用户端口地址，用于连接到外部客户端和钻取Web控制台的群集中的节点之间。
* heap\_current在堆上使用的内存量，以字节为单位。
* heap\_max堆中可用的最大内存量，以字节为单位。
* direct\_current分配器使用的当前直接内存，以字节为单位。
* jvm\_direct\_current当前的JVM直接内存分配，以字节为单位。
* direct\_max分配器可用的最大直接内存，以字节为单位。

有关如何配置钻取系统和会话选项的信息，请参阅[计划和执行选项](http://drill.apache.org/docs/planning-and-execution-options)。

有关如何配置Drill启动选项的信息，请参阅[启动选项](http://drill.apache.org/docs/start-up-options)。

性能调优

# 性能调整

* [性能调整介绍](http://drill.apache.org/docs/performance-tuning-introduction/)
* [分区修剪](http://drill.apache.org/docs/partition-pruning/)
  + [分区修剪简介](http://drill.apache.org/docs/partition-pruning-introduction/)
  + [如何分区数据](http://drill.apache.org/docs/how-to-partition-data/)
* [异步拼花阅读器](http://drill.apache.org/docs/asynchronous-parquet-reader/)
* [优化拼花元数据阅读](http://drill.apache.org/docs/optimizing-parquet-metadata-reading/)
* [实木复合地板过滤器下推](http://drill.apache.org/docs/parquet-filter-pushdown/)
* [Hive元数据缓存](http://drill.apache.org/docs/hive-metadata-caching/)
* [选择存储格式](http://drill.apache.org/docs/choosing-a-storage-format/)
* [查询计划和调整](http://drill.apache.org/docs/query-plans-and-tuning/)
  + [查询计划和调整介绍](http://drill.apache.org/docs/query-plans-and-tuning-introduction/)
  + [加入计划指南](http://drill.apache.org/docs/join-planning-guidelines/)
  + [优化聚合指南](http://drill.apache.org/docs/guidelines-for-optimizing-aggregation/)
  + [修改查询计划选项](http://drill.apache.org/docs/modifying-query-planning-options/)
  + [基于排序和哈希的内存约束运算符](http://drill.apache.org/docs/sort-based-and-hash-based-memory-constrained-operators/)
  + [启用查询队列](http://drill.apache.org/docs/enabling-query-queuing/)
  + [节流](http://drill.apache.org/docs/throttling/)
  + [控制并行化以平衡性能与多租户](http://drill.apache.org/docs/controlling-parallelization-to-balance-performance-with-multi-tenancy/)
* [识别性能问题](http://drill.apache.org/docs/identifying-performance-issues/)
  + [查询计划](http://drill.apache.org/docs/query-plans/)
  + [查询配置文件](http://drill.apache.org/docs/query-profiles/)
* [性能调优参考](http://drill.apache.org/docs/performance-tuning-reference/)
  + [查询配置文件列说明](http://drill.apache.org/docs/query-profile-column-descriptions/)
  + [物理运营商](http://drill.apache.org/docs/physical-operators/)
* [监测指标](http://drill.apache.org/docs/monitoring-metrics/)

# 数据源和文件格式

* [数据源和文件格式介绍](http://drill.apache.org/docs/data-sources-and-file-formats-introduction/)
* [Hive-Drill数据类型映射](http://drill.apache.org/docs/hive-to-drill-data-type-mapping/)
* [部署和使用Hive UDF](http://drill.apache.org/docs/deploying-and-using-a-hive-udf/)
* [镶木地板格式](http://drill.apache.org/docs/parquet-format/)
* [JSON数据模型](http://drill.apache.org/docs/json-data-model/)
* [文本文件：CSV，TSV，PSV](http://drill.apache.org/docs/text-files-csv-tsv-psv/)
* [序列文件](http://drill.apache.org/docs/sequence-files/)

## 数据源和文件格式介绍：

### 2016年5月5日

包含在Drill支持的数据源中的是这些关键数据源：

* HBase的
* 蜂巢
* MAPR-DB
* 文件系统

Drill认为数据源具有强大的模式或弱的模式。

下表介绍了每种模式类型：

| **架构类型** | **描述** |
| --- | --- |
| 强大 | 除文本文件数据源外，Drill验证与强模式关联的数据源包含与查询中使用的数据类型兼容的数据类型。Drill还会验证查询中引用的列是否存在于基础数据源中。如果列不存在，CREATE VIEW将失败。 |
| 弱 | Drill不会验证与弱架构相关联的数据源是否包含与查询中使用的数据类型兼容的数据类型。Drill不验证在Parquet数据源的查询中引用的列是否存在，因此CREATE VIEW总是成功。在JSON文件的情况下，Drill不验证文件是否包含视图中指定的映射。 |

下表列出了模式的当前类别以及与其相关的数据源：

|  | **强大的架构** | **模式弱** |
| --- | --- | --- |
| 数据源 | 视图，配置单元表，hbase列族，文本 | json，mongodb，hbase列限定符，实木复合地板 |

### Drill支持以下数据输入格式：

* [Avro公司](http://avro.apache.org/docs/current/spec.html)
* CSV（逗号分隔值）
* TSV（制表符分隔值）
* PSV（管道分离值）
* 实木复合地板
* MAPR-DB \*
* Hadoop序列文件

\*仅在使用mapr-drill软件包在群集上安装Drill时才可用。

您将[存储插件](http://drill.apache.org/docs/storage-plugin-registration)定义的工作空间部分中来自数据源的数据的输入格式设置为Drill 。Drill中默认的输入格式是Parquet。

您更改[sys.options表](http://drill.apache.org/docs/configuration-options-introduction/)store中的一个属性以设置Drill数据的输出格式。Drill CREATE TABLE AS（CTAS）语句的默认存储格式是Parquet。

## Parquet 数据格式：

2017年8月9日

[Apache Parquet](http://parquet.incubator.apache.org/documentation/latest)具有以下特点：

* 自描述
* 列格式
* 与语言无关

自描述数据将数据本身嵌入到模式或结构中。Hadoop用例推动自描述数据格式（如Parquet和JSON）以及NoSQL数据库（如HBase）的增长。这些格式和数据库非常适合Hadoop应用程序和BI /分析的敏捷和迭代开发周期。为了处理大文件而进行了优化，Parquet将列中的数据进行排列，将相关值彼此靠近以优化查询性能，最大限度地减少I / O并促进压缩。实木复合地板使用节省资源的技术检测和编码相同或相似的数据。

Apache Drill包含以下对Parquet的支持：

* 查询文件或NoSQL数据库中的自描述数据，而无需在集中式转换中定义和管理模式覆盖定义
* 从其他文件格式（如JSON）创建Parquet文件，无需任何设置
* 生成具有演变或改变模式的Parquet文件，并实时查询数据
* 处理Parquet数据类型

### Parquet File

当读取Parquet数据时，Drill只加载必要的数据列，从而减少I / O。Drill可以从数据文件或表中只读取一小部分Parquet数据，可以检查和分析跨多个文件的列的所有值。您可以创建一个格式的钻表，并以另一种格式存储数据，包括Parquet。

### 写parquet文件

CREATE TABLE AS（CTAS）可以使用存储插件提供的任何数据源。要使用CTAS命令写入Parquet数据，请设置session store.format选项，如[配置Parquet存储格式中所示](http://drill.apache.org/docs/parquet-format/#configuring-the-parquet-storage-format)。或者，将存储插件配置为指向包含Parquet文件的目录。

尽管数据驻留在单个表中，但Parquet输出通常由多个文件组成，这些文件类似于具有编号文件名的MapReduce输出，如目录中的0\_0\_0.parquet。

### 日期值自动更正

从Drill 1.10开始，Drill写入标准Parquet日期值。Drill还具有自动修正功能，可自动检测并修正Drill 1.10之前Drill写入Parquet文件的损坏日期值。

默认情况下，自动更正功能已打开，并可在未来5000年的时间内工作。在极少的情况下，Drill需要将日期写入将来数千年，关闭自动更正功能。

要禁用自动更正功能，请导航到存储插件配置，autoCorrectCorruptDates并将Parquet配置中的选项更改为“false”，如下例所示：

"formats": {

"parquet": {

"type": "parquet",

"autoCorrectCorruptDates": false

}

或者，您可以在发出查询时将该选项设置为false，如以下示例所示：

SELECT l\_shipdate, l\_commitdate FROM table(dfs.`/drill/testdata/parquet\_date/dates\_nodrillversion/drillgen2\_lineitem`

(type => 'parquet', autoCorrectCorruptDates => false)) LIMIT 1;

### 配置Parquet存储格式

要读取或写入Parquet数据，您需要在存储插件格式定义中包含Parquet格式。该dfs插件定义包括镶木格式。

使用该store.format选项在会话或系统级别设置Parquet行组的CTAS输出格式。

使用ALTER命令设置store.format选项。

ALTER SYSTEM|SESSION SET `store.format` = 'parquet';

### 配置Parquet文件的大小

通过设置配置Parquet文件的大小store.parquet.block-size可以提高写入性能。块大小是MFS，HDFS或文件系统的大小。

Drill需要缓冲数据的块越大。包含单个块的镶木地板文件使得数据量最大化，连续地存储在磁盘上。给定每个文件的单个行组，Drill将整个Parquet文件存储到块中，从而避免网络I / O。

要使性能最大化，请使用该store.parquet.block-size选项将Parquet行组的目标大小设置为小于或等于MFS，HDFS块大小或文件系统的字节数，如下所示：

ALTER SYSTEM|SESSION SET `store.parquet.block-size` = 536870912;

默认的块大小是536870912字节。

### 配置Parquet文件的HDFS块大小

Drill 1.11引入了这个store.parquet.writer.use\_single\_fs\_block选项，它使Drill能够在不更改默认文件系统块大小的情况下将Parquet文件写入单个文件系统块。当Drill将Parquet文件作为文件系统上的一个块读取时，查询性能会提高。当store.parquet.writer.use\_single\_fs\_block启用该选项时，该store.parquet.block-size设置确定创建的Parquet文件的块大小。该store.parquet.writer.use\_single\_fs\_block选项的默认设置是“false”。使用SET命令启用或禁用该选项，如下所示：

ALTER SYSTEM|SESSION SET store.parquet.writer.use\_single\_fs\_block = 'true|false';

### 类型映射

Parquet和SQL数据类型之间的高度相关性使得在Drill中轻松阅读Parquet文件。写入Parquet文件需要比阅读更多的工作。由于SQL不支持所有的Parquet数据类型，为了防止Drill推断除你想要的类型以外的类型，使用转换[函数](http://drill.apache.org/docs/data-type-conversion/#cast)如果Parquet数据是逻辑类型，则Drill为Parquet转换提供比SQL更多的自由转换功能。

以下一般过程将文件从JSON转换为Parquet：

* 创建或使用现有的存储插件，指定Parquet文件的存储位置，数据的可变性和支持的文件格式。
* 看看JSON数据。
* 创建一个选择JSON文件的表。
* 在CTAS命令中，将JSON字符串数据转换为相应的[SQL类型](http://drill.apache.org/docs/json-data-model/#data-type-mapping)。

### 例如：读取JSON，写入实木复合地板

此示例演示存储插件定义，来自JSON文件的示例数据行以及将JSON输入写入Parquet输出的Drill查询。

### 存储插件定义

您可以使用与Drill一起安装的默认dfs存储插件来读写Parquet文件。存储插件需要将工作区的可写选项配置为true，所以Drill可以写入Parquet输出。dfs存储插件定义了tmp可写入的工作区，您可以在CTAS命令中使用该工作区来创建Parquet表。

### JSON数据示例行

名为sample.json的JSON文件包含由JSON数据的典型字符串组成的数据。以下示例显示了JSON文件的一行：

{"trans\_id":0,"date":"2013-07-26","time":"04:56:59","amount":80.5,"user\_info":

{"cust\_id":28,"device":"WEARABLE2","state":"mt"

},"marketing\_info":

{"camp\_id":4,"keywords": ["go","to","thing","watch","made","laughing","might","pay","in","your","hold"]

},

"trans\_info":

{"prod\_id":[16],

"purch\_flag":"false"

}

}

### CTAS查询

以下示例显示一个CTAS查询，该查询根据上一个示例中显示的JSON数据创建一个表。该命令将日期，时间和数量字符串转换为SQL类型DATE，TIME和DOUBLE。其他字符串的字符串到VARCHAR转换自动发生。

CREATE TABLE dfs.tmp.sampleparquet AS

(SELECT trans\_id,

cast(`date` AS date) transdate,

cast(`time` AS time) transtime,

cast(amount AS double) amountm,

user\_info, marketing\_info, trans\_info

FROM dfs.`/Users/drilluser/sample.json`);

CTAS查询不指定输出的文件扩展名。Drill默认创建一个parquet文件，如输出中的文件名所示：

+------------+---------------------------+

| Fragment | Number of records written |

+------------+---------------------------+

| 0\_0 | 5 |

+------------+---------------------------+

1 row selected (1.369 seconds)

您可以查询Parquet文件以验证Drill现在将转换的字符串解释为日期。

SELECT extract(year from transdate) AS `Year`, t.user\_info.cust\_id AS Customer

FROM dfs.tmp.`sampleparquet` t;

+------------+------------+

| Year | Customer |

+------------+------------+

| 2013 | 28 |

| 2013 | 86623 |

| 2013 | 11 |

| 2013 | 666 |

| 2013 | 999 |

+------------+------------+

5 rows selected (0.039 seconds)

有关使用Parquet数据的更多示例和信息，请参阅[“](https://www.mapr.com/blog/evolving-parquet-self-describing-data-format-new-paradigms-consumerization-hadoop-data#.VNeqQbDF_8f)将Parquet [演化为自描述数据格式 - Hadoop数据消费化的新范例”](https://www.mapr.com/blog/evolving-parquet-self-describing-data-format-new-paradigms-consumerization-hadoop-data#.VNeqQbDF_8f)。

### SQL数据类型实木复合地板

本节中的第一个表将SQL数据类型映射到Parquet数据类型，由Parquet创建者有意限制，以最大限度地减少对磁盘存储的影响：

| **SQL Type** | **Parquet Type** | **Description** |
| --- | --- | --- |
| BIGINT | INT64 | 8-byte signed integer |
| BOOLEAN | BOOLEAN | TRUE (1) or FALSE (0) |
| N/A | BYTE\_ARRAY | Arbitrarily long byte array |
| FLOAT | FLOAT | 4-byte single precision floating point number |
| DOUBLE | DOUBLE | 8-byte double precision floating point number |
| INTEGER | INT32 | 4-byte signed integer |
| VARBINARY(12)\* | INT96 | 12-byte signed int |

\* Drill 1.10及更高版本可以隐式地将Parquet INT96类型解释为TIMESTAMP（具有标准的8字节/毫秒精度），当store.parquet.reader.int96\_as\_timestamp启用该选项时。在Drill（1.2到1.9）的早期版本中或者store.parquet.reader.int96\_as\_timestamp禁用选项时，您必须使用Drill的CONVERT\_FROM函数将INT96值正确解释为TIMESTAMP值。

### 关于INT96支持

从Drill 1.10开始，当store.parquet.reader.int96\_as\_timestamp启用该选项时，Drill可以隐式地解释Parquet文件中的INT96时间戳数据类型。对于Drill的早期版本，或者store.parquet.reader.int96\_as\_timestamp禁用选项时，您必须使用CONVERT\_FROM函数，

该store.parquet.reader.int96\_as\_timestamp选项默认是禁用的。使用[ALTER SYSTEM | SESSION SET](http://drill.apache.org/docs/alter-system/)命令启用该选项。不必要地启用此选项可能会导致查询失败，因为CONVERT\_FROM（col，'TIMESTAMP\_IMPALA'）功能在store.parquet.reader.int96\_as\_timestamp启用时不起作用。

### 使用CONVERT\_FROM来解释INT96

在较早版本的Drill（1.2到1.9）中，必须使用CONVERT\_FROM函数进行Drill来解释Parquet INT96类型。例如，要解码来自Hive或Impala（INT96类型）的时间戳，请使用CONVERT\_FROM函数和[TIMESTAMP\_IMPALA](http://drill.apache.org/docs/supported-data-types/#data-types-for-convert_to-and-convert_from-functions)类型参数：

SELECT CONVERT\_FROM(timestamp\_field, 'TIMESTAMP\_IMPALA') as timestamp\_field FROM `dfs.file\_with\_timestamp.parquet`;

由于INT96仅支持只读，所以不能将TIMESTAMP\_IMPALA用作CONVERT\_TO的数据类型参数。您可以使用CAST函数将SQL TIMESTAMP转换为VARBINARY，但生成的VARBINARY与INT96不同。

例如，在读取INT96并将一些数据转换为时间戳后创建一个钻取表。

CREATE TABLE t2(c1) AS SELECT CONVERT\_FROM(created\_ts, 'TIMESTAMP\_IMPALA') FROM t1 ORDER BY 1 LIMIT 1;

t1.created\_ts是INT96（或Hive / Impala时间戳），t2.created\_ts是SQL时间戳。这些类型是不可比较的。你不能使用像t1.created\_ts = t2.created\_ts这样的条件。

### 配置时区

默认情况下，INT96时间戳值表示本地日期和时间，与Hive类似。要获得UTC的INT96时间戳值，请配置钻取[UTC时间](http://drill.apache.org/docs/data-type-conversion/#time-zone-limitation)。

### SQL类型实例逻辑类型

Parquet还支持逻辑类型，在[Apache Parquet站点](https://github.com/Parquet/parquet-format/blob/master/LogicalTypes.md)上进行了充分描述。嵌入式类型，JSON和BSON，注释表示JSON或BSON文档的二进制基本类型。逻辑类型及其对SQL类型的映射如下

| **SQL Type** | **Drill Description** | **Parquet Logical Type** | **Parquet Description** |
| --- | --- | --- | --- |
| DATE | Years months and days in the form in the form YYYY-­MM-­DD | DATE | Date, not including time of day. Uses the int32 annotation. Stores the number of days from the Unix epoch, 1 January 1970. |
| VARCHAR | Character string variable length | UTF8 (Strings) | Annotates the binary primitive type. The byte array is interpreted as a UTF-8 encoded character string. |
| None |  | INT\_8 | 8 bits, signed |
| None |  | INT\_16 | 16 bits, usigned |
| INT | 4-byte signed integer | INT\_32 | 32 bits, signed |
| DOUBLE | 8-byte double precision floating point number | INT\_64 | 64 bits, signed |
| None |  | UINT\_8 | 8 bits, unsigned |
| None |  | UINT\_16 | 16 bits, unsigned |
| None |  | UINT\_32 | 32 bits, unsigned |
| None |  | UINT\_64 | 64 bits, unsigned |
| DECIMAL\* | 38-digit precision | DECIMAL | Arbitrary-precision signed decimal numbers of the form unscaledValue \* 10^(-scale) |
| TIME | Hours, minutes, seconds, milliseconds; 24-hour basis | TIME\_MILLIS | Logical time, not including the date. Annotates int32. Number of milliseconds after midnight. |
| TIMESTAMP | Year, month, day, and seconds | TIMESTAMP\_MILLIS | Logical date and time. Annotates an int64 that stores the number of milliseconds from the Unix epoch, 00:00:00.000 on 1 January 1970, UTC. |
| INTERVAL | Integer fields representing a period of time depending on the type of interval | INTERVAL | An interval of time. Annotates a fixed\_len\_byte\_array of length 12. Months, days, and ms in unsigned little-endian encoding. |

\*在此版本中，Drill将禁用DECIMAL数据类型，包括转换为DECIMAL并从Parquet和Hive读取DECIMAL类型。要启用DECIMAL类型，请将该planner.enable\_decimal\_data\_type选项设置为true。

### 数据描述语言支持

Parquet支持以下数据描述语言：

* 阿帕奇Avro
* Apache Thrift
* Google协议缓冲区

实施自定义存储插件，为Thrift等格式创建Parquet读写器。