# 实验二 **数据血缘Apache Altas API开发**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **地 点：** | B10楼 | 303房； | **实验台号：** |  |
| **实验日期与时间：** | 2023/4/11 | | **评 分：** |  |
| **预习检查纪录：** |  | | **实验教师：** | 蔡文学 |

1. **实验目的**

在当今大数据的应用越来越广泛的情况下，数据治理一直是企业面临的巨大问题，大部分只是单纯的对数据进行了处理，而数据的血缘、分类等等却很难实现，市场上也急需要一个专注于数据治理的技术框架，这时Atlas应运而生。

Atlas是一套可扩展、可延伸的核心基础治理服务–使企业能够在Hadoop内有效、高效地满足其合规性要求，并允许与整个企业数据生态系统进行整合。

Apache Atlas为企业提供了开放的元数据管理和治理能力，可以为企业建立数据资产目录，对这些资产进行分类和治理，并为数据科学家、分析师和数据治理团队提供围绕这些数据资产的协作能力。

而atlas api分为2类，一类是atlas本身的api，一类是restful api。

1. **实验平台**

(1).操作系统：ubuntu

(2).Apache Altas版本：2.0.0

1. **实验步骤（包括实验结果/截图）**

**Atlas Type System**

**Atlas 类型系统，Atlas 允许用户为他们想要管理的元数据对象定义一个模型。该模型由称为 “类型” 的定义组成。被称为 “实体” 的 “类型” 实例表示被管理的实际元数据对象。类型系统是一个组件，允许用户定义和管理类型和实体。由 Atlas 管理的所有元数据对象（例如Hive表）都使用类型进行建模，并表示为实体。如果要在Atlas中存储新类型的元数据，需要了解类型系统组件的概念。**

**Types**

**Atlas中的 “类型” 定义了如何存储和访问特定类型的元数据对象。类型表示了所定义元数据对象的一个或多个属性集合。具有开发背景的用户可以将 “类型” 理解成面向对象的编程语言的 “类” 定义的或关系数据库的 “表模式”。**

**可以通过该API获取Atlas的所有类型：http://atlas:21000/api/atlas/v2/types/typedefs**

**下面通过该API获取hive\_table类型的定义：http://atlas:21000/api/atlas/v2/types/typedef/name/hive\_table**

**从上面的例子可以注意到以下几点：**

**Atlas中的类型由 “name” 唯一标识，**

**attributeDefs表示该类型中属性的定义**

**类型具有元类型。元类型表示 Atlas 中此模型的类型。 Atlas 有以下几种类型：**

**基本元类型： Int，String，Boolean等。**

**枚举元类型**

**集合元类型：例如Array，Map**

**复合元类型：Class，Struct，Trait**

**类型可以从称为 “superTypes” 的父类型 “extend” - 通过这种方式，它将拥有在 “supertype” 中定义的属性。这允许模型在一组相关类型等之间定义公共属性。这再次类似于面向对象语言如何定义类的超类的概念。 Atlas 中的类型也可以从多个超类型扩展。同时subTypes表示该类型的子类型**

**在该示例中，每个 hive\_table 类型从预定义的超类型（称为 “DataSet”）扩展。**

**具有 “Class”，“Struct” 或 “Trait” 的元类型的类型可以具有属性集合。每个属性都有一个名称（例如 “name”）和一些其他关联的属性。可以使用表达式 type\_name.attribute\_name 来引用属性。还要注意，属性本身是使用 Atlas 元类型定义的。**

**在这个例子中，hive\_table.name 是一个字符串，hive\_table.aliases 是一个字符串数组，hive\_table.db 引用一个类型的实例称为 hive\_db 等等。**

**在属性中键入引用（如hive\_table.db）。使用这样的属性，我们可以在 Atlas 中定义的两种类型之间的任意关系，从而构建丰富的模型。注意，也可以收集一个引用列表作为属性类型（例如 hive\_table.columns，它表示从 hive\_table 到 hive\_column 类型的引用列表）**

**可以看到DataSet有很多的子类型，部分Atlas自带的类型都继承自DataSet。**

**同时DataSet继承自Asset，Asset表示资产的意思，其中定义了一些通用的属性。Asset类型中定义了3个属性**

**name：名称**

**owner：所属人**

**description：描述**

**它也有不少子类型，表示资产的意义，比如说hive数据库hive\_db，hbase命名空间hbase\_namespace。它继承自Referenceable类型**

**该类型定义了一个非常重要的属性，qualifiedName， 该类型中唯一限定名，可以通过该属性配合类型名在Atlas中查找对应的唯一实体内容，注意与guid的区别，guid是全局唯一的**

**例如:**

**一个hive数据库的qualifiedName：test@primary**

**该数据库test下表test\_table的qualifiedName：test.test\_table@primary**

**该表test\_table中字段name的qualifiedName：test.test\_table.name@primary**

**@primary表示集群默认名字，通过如下配置集群名称，通过加上集群名称，在不同集群间唯一标识一个实体**

**​ atlas.cluster.name**

**​ primary**

**Process类型继承自Asset类型，所以自带有name，owner，description，quailifiedName四种属性。它自己特有的inputs和outputs表示该过程的输入输出，它是Atlas血缘管理中所有类型的超类**

**在概念上，它可以用于表示任何数据变换操作。例如，将原始数据的 hive 表转换为存储某个聚合的另一个 hive 表的 ETL 过程可以是扩展过程类型的特定类型。流程类型有两个特定的属性，输入和输出。输入和输出都是 DataSet 实体的数组。因此，Process 类型的实例可以使用这些输入和输出来捕获 DataSet 的 lineage 如何演变。**

**Entities**

**表示实体，前面的Type理解为Java的类，则Entity可以理解为类对应的一个实例，该类的一个对象**

**表名：student 包含3个列：name，age，phone**

**开头referredEntities中包含了该table的引用的实体对象，即3个列实体和一个hive\_storagedesc实体（表示该表的存储信息），实际存储时存储的为它们的guid**

**entity中包含了该表student的一些信息，name，owner，guid，status，attribute等等**

**作为 Class Type 实例的每个实体都由唯一标识符 GUID 标识。此 GUID 由 Atlas 服务器在定义对象时生成，并在实体的整个生命周期内保持不变。在任何时间点，可以使用其 GUID 来访问该特定实体。**

**Attributes**

**在扩展时常用的一种类型，表示某个类型的附加属性，该类型具有以下属性：**

**此标志与索引相关。如果指定为唯一，这意味着为 JanusGraph 中的此属性创建一个特殊索引，允许基于等式的查找。**

**具有此标志的真实值的任何属性都被视为主键，以将此实体与其他实体区分开。因此，应注意确保此属性在现实世界中模拟独特的属性。**

**例如，考虑 hive\_table 的 name 属性。孤立地，名称不是 hive\_table 的唯一属性，因为具有相同名称的表可以存在于多个数据库中。如果 Atlas 在多个集群中存储 hive 表的元数据，即使一对（数据库名称，表名称）也不是唯一的。只有集群位置，数据库名称和表名称可以在物理世界中被视为唯一。**

**isIndexable：此标志指示此属性是否应该索引，以便可以使用属性值作为谓词来执行查找，并且可以高效地执行查找。**

**constraints：限制类型，该属性的限制类型，猜测可以通过该值来实现类似于MySQL中外键的功能，默认值有如下3个**

**Atlas创建类型与更新类型**

**Atlas创建access\_info类型**

1. **思考讨论题或体会或对改进实验的建议**

**Restful api是**REST（英文：Representational State Transfer，简称REST，直译过来表现层状态转换）是一种软件架构风格、设计风格，而不是标准，只是提供了一组设计原则和约束条件。它主要用于客户端和服务器交互类的软件。基于这个风格设计的软件可以更简洁，更有层次，更易于实现缓存等机制。它首次出现在 2000 年 Roy Thomas Fielding 的博士论文中，这篇论文定义并详细介绍了表述性状态转移（Representational State Transfer，REST）的架构风格，并且描述了 如何使用 REST 来指导现代 Web 架构的设计和开发。

### REST架构特征

### 既然知道REST和RESTful的联系和区别，现在就要开始好好了解RESTful的一些约束条件和规则，RESTful是一种风格而不是标准，而这个风格大致有以下几个主要**特征**：

**以资源为基础** ：资源可以是一个图片、音乐、一个XML格式、HTML格式或者JSON格式等网络上的一个实体，除了一些二进制的资源外普通的文本资源更多以JSON为载体、面向用户的一组数据(通常从数据库中查询而得到)。  
**统一接口**: 对资源的操作包括获取、创建、修改和删除，这些操作正好对应HTTP协议提供的GET、POST、PUT和DELETE方法。换言而知，使用RESTful风格的接口但从接口上你可能只能定位其资源，但是无法知晓它具体进行了什么操作，需要具体了解其发生了什么操作动作要从其HTTP请求方法类型上进行判断。具体的HTTP方法和方法含义如下：

* GET（SELECT）：从服务器取出资源（一项或多项）。
* POST（CREATE）：在服务器新建一个资源。
* PUT（UPDATE）：在服务器更新资源（客户端提供完整资源数据）。
* PATCH（UPDATE）：在服务器更新资源（客户端提供需要修改的资源数据）。
* DELETE（DELETE）：从服务器删除资源。

当然也有很多在具体使用的时候使用PUT表示更新。

### REST架构限制条件

Fielding在论文中提出REST架构的6个**限制条件**，也可称为RESTful 6大原则， 标准的REST约束应满足以下6个原则：

**客户端-服务端（Client-Server）**: 这个更专注客户端和服务端的分离，服务端独立可更好服务于前端、安卓、IOS等客户端设备。

**无状态（Stateless）**：服务端不保存客户端状态，客户端保存状态信息每次请求携带状态信息。

**可缓存性（Cacheability）** ：服务端需回复是否可以缓存以让客户端甄别是否缓存提高效率。

**统一接口（Uniform Interface）**：通过一定原则设计接口降低耦合，简化系统架构，这是RESTful设计的基本出发点。

**分层系统（Layered System）**：客户端无法直接知道连接的到终端还是中间设备，分层允许你灵活的部署服务端项目。

**按需代码（Code-On-Demand，可选）**：按需代码允许我们灵活的发送一些看似特殊的代码给客户端例如JavaScript代码。

1. **实验源代码**

public static AtlasTypesDef getAtlasTypesDef() {

//获取Atlas类型定义，在这里面可以定义所有类型

AtlasTypesDef def = new AtlasTypesDef();

//创建一个Entity定义类型，名称为access\_info

AtlasEntityDef entityDef = new AtlasEntityDef("access\_info");

//设置继承自DataSet类型

entityDef.setSuperTypes(Collections.singleton("DataSet"));

//创建人为

entityDef.setCreatedBy("wangjian");

//设置该类型版本号

entityDef.setVersion((long) 1.0);

//创建一个属性定义数组

List<AtlasStructDef.AtlasAttributeDef> attributeDefs = new ArrayList<AtlasStructDef.AtlasAttributeDef>();

//创建一个accessTime属性

AtlasStructDef.AtlasAttributeDef accessTime = new AtlasStructDef.AtlasAttributeDef();

//设置属性名称为accessTime ，访问时间

accessTime.setName("accessTime");

//类型为date

accessTime.setTypeName("date");

//设置为可选择的，表明没有该项，entity也能创建

accessTime.setIsOptional(true);

//设置为索引，后期可以通过该属性进行有效查找

accessTime.setIsIndexable(true);

//设置为不唯一的

accessTime.setIsUnique(false);

//单值的属性

accessTime.setCardinality(AtlasStructDef.AtlasAttributeDef.Cardinality.SINGLE);

//值最小数量为0，最大为1，必须指定，不指定默认为-1

accessTime.setValuesMinCount(0);

accessTime.setValuesMaxCount(1);

//设置访问对象引用，accessObject，类型为DataSet（hive\_table,kafka\_topic都继承自此类）

AtlasStructDef.AtlasAttributeDef accessObject = new AtlasStructDef.AtlasAttributeDef();

accessObject.setName("accessObject");

accessObject.setTypeName("DataSet");

accessObject.setIsOptional(true);

accessObject.setIsIndexable(true);

accessObject.setIsUnique(false);

accessObject.setCardinality(AtlasStructDef.AtlasAttributeDef.Cardinality.SINGLE);

accessObject.setValuesMinCount(0);

accessObject.setValuesMaxCount(1);

//添加属性到List

attributeDefs.add(accessTime);

attributeDefs.add(accessObject);

//设置属性定义

entityDef.setAttributeDefs(attributeDefs);

//设置类型定义到AtlasTypeDef

def.setEntityDefs(Collections.singletonList(entityDef));

return def;

}

Atlas更新kafka\_topic类型,添加属性

/\*\*

\* 给kafka\_topic类型添加sampleMessage属性

\*/

public void updateKafkaTopic() throws AtlasServiceException {

//根据实体类型定义名获取实体类型定义

AtlasEntityDef kafkaTopic = clientV2.getEntityDefByName("kafka\_topic");

//添加一个新的属性sample\_message,表示kafka消息示例

kafkaTopic.getAttributeDefs().add(new AtlasStructDef.AtlasAttributeDef("sampleMessage",

"array<string>",

true,

AtlasStructDef.AtlasAttributeDef.Cardinality.LIST,

0,

10,

false,

false,

null

));

//将实体类型定义设置到一个新的AtlasTypeDef中

AtlasTypesDef atlasTypesDef = new AtlasTypesDef();

atlasTypesDef.getEntityDefs().add(kafkaTopic);

clientV2.updateAtlasTypeDefs(atlasTypesDef);

}

**hive\_table类型示例：**

**{**

**"category": "ENTITY",**

**"guid": "30a12b7c-faed-4ead-ad83-868893ebed93",**

**"createdBy": "cloudera-scm",**

**"updatedBy": "cloudera-scm",**

**"createTime": 1536203750750,**

**"updateTime": 1536203750750,**

**"version": 1,**

**"name": "hive\_table",**

**"description": "hive\_table",**

**"typeVersion": "1.1",**

**"options": {**

**"schemaElementsAttribute": "columns"**

**},**

**"attributeDefs": [**

**{**

**"name": "db",**

**"typeName": "hive\_db",**

**"isOptional": false,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 1,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "createTime",**

**"typeName": "date",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "lastAccessTime",**

**"typeName": "date",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "comment",**

**"typeName": "string",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "retention",**

**"typeName": "int",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "sd",**

**"typeName": "hive\_storagedesc",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false,**

**"constraints": [**

**{**

**"type": "ownedRef"**

**}**

**]**

**},**

**{**

**"name": "partitionKeys",**

**"typeName": "array<hive\_column>",**

**"isOptional": true,**

**"cardinality": "SET",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 2147483647,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false,**

**"constraints": [**

**{**

**"type": "ownedRef"**

**}**

**]**

**},**

**{**

**"name": "aliases",**

**"typeName": "array<string>",**

**"isOptional": true,**

**"cardinality": "SET",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 2147483647,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "columns",**

**"typeName": "array<hive\_column>",**

**"isOptional": true,**

**"cardinality": "SET",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 2147483647,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false,**

**"constraints": [**

**{**

**"type": "ownedRef"**

**}**

**]**

**},**

**{**

**"name": "parameters",**

**"typeName": "map<string,string>",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "viewOriginalText",**

**"typeName": "string",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "viewExpandedText",**

**"typeName": "string",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "tableType",**

**"typeName": "string",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "temporary",**

**"typeName": "boolean",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": true,**

**"includeInNotification": false**

**}**

**],**

**"superTypes": [**

**"DataSet"**

**],**

**"subTypes": []**

**}**

**DataSet类型定义：**

**{**

**"category": "ENTITY",**

**"guid": "d31c0a02-6999-4f81-a62a-07d7654aec84",**

**"createdBy": "cloudera-scm",**

**"updatedBy": "cloudera-scm",**

**"createTime": 1536203676149,**

**"updateTime": 1536203676149,**

**"version": 1,**

**"name": "DataSet",**

**"description": "DataSet",**

**"typeVersion": "1.1",**

**"attributeDefs": [],**

**"superTypes": [**

**"Asset"**

**],**

**"subTypes": [**

**"rdbms\_foreign\_key",**

**"rdbms\_db",**

**"kafka\_topic",**

**"hive\_table",**

**"sqoop\_dbdatastore",**

**"hbase\_column",**

**"rdbms\_instance",**

**"falcon\_feed",**

**"jms\_topic",**

**"hbase\_table",**

**"rdbms\_table",**

**"rdbms\_column",**

**"rdbms\_index",**

**"hbase\_column\_family",**

**"access\_info",**

**"hive\_column",**

**"avro\_type",**

**"fs\_path"**

**]**

**}**

**{**

**"category": "ENTITY",**

**"guid": "349a5c61-47c3-4f4b-9a79-7fd59454a73a",**

**"createdBy": "cloudera-scm",**

**"updatedBy": "cloudera-scm",**

**"createTime": 1536203676083,**

**"updateTime": 1536203676083,**

**"version": 1,**

**"name": "Asset",**

**"description": "Asset",**

**"typeVersion": "1.1",**

**"attributeDefs": [**

**{**

**"name": "name",**

**"typeName": "string",**

**"isOptional": false,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 1,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": true,**

**"includeInNotification": false**

**},**

**{**

**"name": "description",**

**"typeName": "string",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "owner",**

**"typeName": "string",**

**"isOptional": true,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 1,**

**"isUnique": false,**

**"isIndexable": true,**

**"includeInNotification": false**

**}**

**],**

**"superTypes": [**

**"Referenceable"**

**],**

**"subTypes": [**

**"rdbms\_foreign\_key",**

**"rdbms\_db",**

**"rdbms\_instance",**

**"DataSet",**

**"rdbms\_table",**

**"rdbms\_column",**

**"rdbms\_index",**

**"Infrastructure",**

**"Process",**

**"avro\_type",**

**"hbase\_namespace",**

**"hive\_db"**

**]**

**}**

**Referenceable类型定义：**

**{**

**"category": "ENTITY",**

**"guid": "34c72533-2e80-4e5c-9226-e15b163f98d1",**

**"createdBy": "cloudera-scm",**

**"updatedBy": "cloudera-scm",**

**"createTime": 1536203673540,**

**"updateTime": 1536203673540,**

**"version": 1,**

**"name": "Referenceable",**

**"description": "Referenceable",**

**"typeVersion": "1.0",**

**"attributeDefs": [**

**{**

**"name": "qualifiedName",**

**"typeName": "string",**

**"isOptional": false,**

**"cardinality": "SINGLE",**

**"valuesMinCount": 1,**

**"valuesMaxCount": 1,**

**"isUnique": true,**

**"isIndexable": true,**

**"includeInNotification": false**

**}**

**],**

**"superTypes": [],**

**"subTypes": [**

**"hive\_storagedesc",**

**"Asset"**

**]**

**}**

**Process类型定义:**

**{**

**"category": "ENTITY",**

**"guid": "7c03ccad-29aa-4c5f-8a27-19b536068f69",**

**"createdBy": "cloudera-scm",**

**"updatedBy": "cloudera-scm",**

**"createTime": 1536203677547,**

**"updateTime": 1536203677547,**

**"version": 1,**

**"name": "Process",**

**"description": "Process",**

**"typeVersion": "1.1",**

**"attributeDefs": [**

**{**

**"name": "inputs",**

**"typeName": "array<DataSet>",**

**"isOptional": true,**

**"cardinality": "SET",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 2147483647,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**},**

**{**

**"name": "outputs",**

**"typeName": "array<DataSet>",**

**"isOptional": true,**

**"cardinality": "SET",**

**"valuesMinCount": 0,**

**"valuesMaxCount": 2147483647,**

**"isUnique": false,**

**"isIndexable": false,**

**"includeInNotification": false**

**}**

**],**

**"superTypes": [**

**"Asset"**

**],**

**"subTypes": [**

**"falcon\_feed\_replication",**

**"falcon\_process",**

**"falcon\_feed\_creation",**

**"sqoop\_process",**

**"hive\_column\_lineage",**

**"storm\_topology",**

**"hive\_process"**

**]**

**}**

**hive\_table类型的实体对象示例：**

**{**

**"referredEntities" : {**

**"779734cc-9011-4066-9bb1-25df6f28ac72" : {**

**"typeName" : "hive\_column",**

**"attributes" : {**

**"owner" : "wangjian5185",**

**"qualifiedName" : "test.student.age@primary",**

**"name" : "age",**

**"description" : null,**

**"comment" : null,**

**"position" : 1,**

**"type" : "int",**

**"table" : {**

**"guid" : "75a0c17a-dfdb-4532-9aae-c87b64be958d",**

**"typeName" : "hive\_table"**

**}**

**},**

**"guid" : "779734cc-9011-4066-9bb1-25df6f28ac72",**

**"status" : "ACTIVE",**

**"createdBy" : "admin",**

**"updatedBy" : "admin",**

**"createTime" : 1536215508751,**

**"updateTime" : 1536215508751,**

**"version" : 0**

**},**

**"c47aed54-e4d2-4080-aa7a-5428075f5b20" : {**

**"typeName" : "hive\_column",**

**"attributes" : {**

**"owner" : "wangjian5185",**

**"qualifiedName" : "test.student.phone@primary",**

**"name" : "phone",**

**"description" : null,**

**"comment" : null,**

**"position" : 2,**

**"type" : "int",**

**"table" : {**

**"guid" : "75a0c17a-dfdb-4532-9aae-c87b64be958d",**

**"typeName" : "hive\_table"**

**}**

**},**

**"guid" : "c47aed54-e4d2-4080-aa7a-5428075f5b20",**

**"status" : "ACTIVE",**

**"createdBy" : "admin",**

**"updatedBy" : "admin",**

**"createTime" : 1536215508751,**

**"updateTime" : 1536215508751,**

**"version" : 0**

**},**

**"7c4b4cd5-841d-409b-b38a-77ec8779e252" : {**

**"typeName" : "hive\_column",**

**"attributes" : {**

**"owner" : "wangjian5185",**

**"qualifiedName" : "test.student.name@primary",**

**"name" : "name",**

**"description" : null,**

**"comment" : null,**

**"position" : 0,**

**"type" : "string",**

**"table" : {**

**"guid" : "75a0c17a-dfdb-4532-9aae-c87b64be958d",**

**"typeName" : "hive\_table"**

**}**

**},**

**"guid" : "7c4b4cd5-841d-409b-b38a-77ec8779e252",**

**"status" : "ACTIVE",**

**"createdBy" : "admin",**

**"updatedBy" : "admin",**

**"createTime" : 1536215508751,**

**"updateTime" : 1536215508751,**

**"version" : 0**

**},**

**"a6038b00-ce2d-4612-9436-d63092d09182" : {**

**"typeName" : "hive\_storagedesc",**

**"attributes" : {**

**"bucketCols" : null,**

**"qualifiedName" : "test.student@primary\_storage",**

**"sortCols" : null,**

**"storedAsSubDirectories" : false,**

**"location" : "hdfs://cdhtest/user/hive/warehouse/test.db/student",**

**"compressed" : false,**

**"inputFormat" : "org.apache.hadoop.mapred.TextInputFormat",**

**"outputFormat" : "org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat",**

**"parameters" : null,**

**"table" : {**

**"guid" : "75a0c17a-dfdb-4532-9aae-c87b64be958d",**

**"typeName" : "hive\_table"**

**},**

**"serdeInfo" : {**

**"typeName" : "hive\_serde",**

**"attributes" : {**

**"serializationLib" : "org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe",**

**"name" : null,**

**"parameters" : null**

**}**

**},**

**"numBuckets" : 0**

**},**

**"guid" : "a6038b00-ce2d-4612-9436-d63092d09182",**

**"status" : "ACTIVE",**

**"createdBy" : "admin",**

**"updatedBy" : "admin",**

**"createTime" : 1536215508751,**

**"updateTime" : 1536215508751,**

**"version" : 0**

**}**

**},**

**"entity" : {**

**"typeName" : "hive\_table",**

**"attributes" : {**

**"owner" : "wangjian5185",**

**"temporary" : false,**

**"lastAccessTime" : 1533807252000,**

**"aliases" : null,**

**"qualifiedName" : "test.student@primary",**

**"columns" : [ {**

**"guid" : "7c4b4cd5-841d-409b-b38a-77ec8779e252",**

**"typeName" : "hive\_column"**

**}, {**

**"guid" : "779734cc-9011-4066-9bb1-25df6f28ac72",**

**"typeName" : "hive\_column"**

**}, {**

**"guid" : "c47aed54-e4d2-4080-aa7a-5428075f5b20",**

**"typeName" : "hive\_column"**

**} ],**

**"description" : null,**

**"viewExpandedText" : null,**

**"sd" : {**

**"guid" : "a6038b00-ce2d-4612-9436-d63092d09182",**

**"typeName" : "hive\_storagedesc"**

**},**

**"tableType" : "MANAGED\_TABLE",**

**"createTime" : 1533807252000,**

**"name" : "student",**

**"comment" : null,**

**"partitionKeys" : null,**

**"parameters" : {**

**"transient\_lastDdlTime" : "1533807252"**

**},**

**"db" : {**

**"guid" : "a804165e-77ff-4c60-9ee7-956760577a1e",**

**"typeName" : "hive\_db"**

**},**

**"retention" : 0,**

**"viewOriginalText" : null**

**},**

**"guid" : "75a0c17a-dfdb-4532-9aae-c87b64be958d",**

**"status" : "ACTIVE",**

**"createdBy" : "admin",**

**"updatedBy" : "admin",**

**"createTime" : 1536215508751,**

**"updateTime" : 1536281983181,**

**"version" : 0**

**}**

**}**

**此处以hive\_table类型的db，createTime，columns属性为例**

**"attributeDefs" : [ {**

**"name" : "db",**

**"typeName" : "hive\_db",**

**"isOptional" : false,**

**"cardinality" : "SINGLE",**

**"valuesMinCount" : 1,**

**"valuesMaxCount" : 1,**

**"isUnique" : false,**

**"isIndexable" : false**

**}, {**

**"name" : "createTime",**

**"typeName" : "date",**

**"isOptional" : true,**

**"cardinality" : "SINGLE",**

**"valuesMinCount" : 0,**

**"valuesMaxCount" : 1,**

**"isUnique" : false,**

**"isIndexable" : false**

**}, {**

**"name" : "columns",**

**"typeName" : "array<hive\_column>",**

**"isOptional" : true,**

**"cardinality" : "SET",**

**"valuesMinCount" : 0,**

**"valuesMaxCount" : 2147483647,**

**"isUnique" : false,**

**"isIndexable" : false,**

**"constraints" : [ {**

**"type" : "ownedRef"**

**} ]**

**}**

**name：该属性名**

**typeName：该属性类型，包括基本类型，以及date，各种type类型，和集合类型等等**

**isOptional：是否可选，false表示该属性必须指定**

**cardinality：如下图三种，SINGLE(单个)，LIST(可重复多个)，SET(不可重复多个)**

**valuesMinCount：该属性最小个数**

**valuesMaxCount：该属性最大个数**

**isUnique：是否为唯一属性**