



北京大学  
PEKING UNIVERSITY

## HBase中的空间数据存储组织

城市与环境学院尹懿行 (1401214663)



College of Urban and Environmental Sciences

---

---

---

---

---

---

---

---



北京大学  
PEKING UNIVERSITY

## 第一部分：Hadoop简介

### 1.Hadoop是什么：

- Hadoop是一个由Apache基金会所开发的分布式系统基础架构。

### 2.Hadoop名字的由来：

- Hadoop这个名字不是一个缩写，它是一个虚构的名字。该项目的创建者Doug Cutting解释了这一名称的来历：“这个名字是我孩子给一头吃饱了的棕黄色大象取得。我的命名标准是简短，容易发音和拼写，没有太多的含义，并且不会被用于别处。小孩子是这方面的高手，Google就是小孩子起的名字（10的100次方google的由来）”



Hadoop之父Doug Cutting

College of Urban and Environmental Sciences

---

---

---

---

---

---

---

---



北京大学  
PEKING UNIVERSITY

## 第一部分：Hadoop简介

### 1.Hadoop发展史：

- 2002年，Doug Cutting 启动Apache Lucene项目的一个子项目，名为Apache Nutch。目标是做一个网页爬取工具和搜索引擎。但开发者认为该项目扩展度不够，不能解决10亿网页的搜索问题
- 2003年，Google发表论文，提出GFS文件系统，GFS或者类似的架构可以解决他们在网络爬取和索引过程中产生的超大文件储存的问题。GFS可以节省系统管理所花的大量时间
- 2004年，开始着手实现一个GFS的开源实现，即Nutch分布式文件系统（NDFS）
- 2004年，Google发表论文，介绍他们的MapReduce系统
- 2005年，Nutch的开发人员在Nutch上实现了一个MapReduce系统，年中Nutch的所有主要算法均完成移植，用MapReduce和NDFS来运行
- 2006年，开发人员将NDFS和MapReduce移出Nutch形成Lucene的一个子项目。称为Hadoop
- 2008年，Hadoop已经成为了Apache的顶级项目

College of Urban and Environmental Sciences

---

---

---

---

---

---

---

---



## 第一部分：Hadoop简介

### 1.Hadoop生态圈：

- **Common**：一组分布式文件系统和通用的I/O的组件接口（序列化、JavaRPC、持久化）
- **Avro**：一种支持高效，跨语言的RPC以及永久存储的序列化系统。
- **MapReduce**：分布式数据处理模型和执行环境，运行于大型商用机集群。
- **HDFS**：分布式文件系统，运行于大型商用机集群。
- **Pig**：一种数据流语言和运行环境，用以检索非常大的数据集。Pig运行在MapReduce和HDFS的集群上。
- **Hive**：一个分布式，按列存储数据仓库。Hive管理HDFS中储存的数据，并提供基于SQL的查询语言（由运行时引擎翻译成MapReduce作业）用以查询数据。
- **HBase**：一个分布式、按列存储数据库。HBase使用HDFS作为底层储存，同时支持MapReduce的批量式计算和点查询（随机读取）。
- **Zookeeper**：一个分布式、可用性高的协调服务。Zookeeper提供分布式锁之类的基本服务用于构建分布式应用。
- **Sqoop**：在关系型数据库和HDFS之间高效传输数据的工具

College of Urban and Environmental Sciences

---

---

---

---

---

---

---

---



## 第二部分：HBase简介

### 1.Hadoop与HBase：

- **HBase**是Apache的Hadoop项目的子项目
- **HBase**是一个在HDFS上开发的分布式数据库，适合于储存大数据，支持单条记录的快速查询，在任意指定位置单条或批量添加、删除数据。Hbase是基于稀疏存储的行/列矩阵，不存储列值为空的表元素，大大节省了存储空间。HBase与HDFS一样，支持MapReduce框架做分布式计算操作。HBase的行键设计灵活，自动按照字典排序。这些特性，十分适合存储和处理空间数据

College of Urban and Environmental Sciences

---

---

---

---

---

---

---

---



## 第一部分：HBase简介

### 2.HBase应用：

- **Hbase**支持大数据的存储，又支持单条记录的快速查询等特性，所以其在一些企业应用比较广泛，比如FaceBook的实时消息存储处理系统、ODS（实时统计用户访问量）业务，eBay的主搜索业务，阿里的阿里旺旺，个性化推荐系统等，Yahoo拥有世界上最大的HBase集群。

College of Urban and Environmental Sciences

---

---

---

---

---

---

---

---

## 第一部分：HBase简介

### 3.Hbase对空间数据的存储与管理：

- **Hbase**的出现，使得海量空间数据的存储、管理、查询、汇交变成了可能。国内外均有大量的相关研究资料与案例。总体说来，目前市面上还没有一款类似于ArcGIS SDE的基于NOSQL数据库的成熟空间数据库产品（ArcGIS SDE是基于关系数据库的空间数据库引擎）。
- 就不同科研课题或企业应用而言，基于HBase，充分发挥HBase在大数据管理上的优势，设计一套高效的集群架构、一套合理的空间数据模型，满足空间数据存储和管理的需要不失为一种值得推荐的方案。

## 第一部分：HBase简介

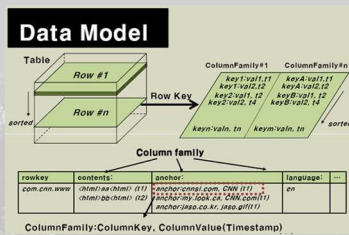
### 4.官网地址：

- **Hadoop**:<http://hadoop.apache.org/>
- **HBase**:<http://hbase.apache.org/>

## 第二部分：HBase的数据模型

### 1.HBase的逻辑模型：

- HBase以表的形式存储数据，每个表由行和列组成，每个列属于一个特定的列族（Column Family），行和列确定的存储单元为一个元素（Cell），每个元素保存同一份数据的多个版本，通过时间戳（Time Stamp）标示。由于HBase是稀疏存储数据，所以某些列可以为空。

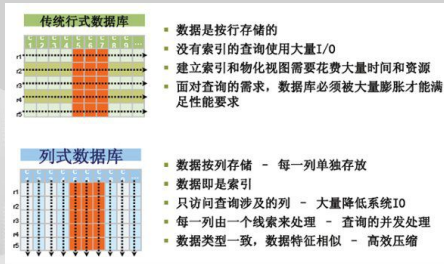


## 第二部分：HBase的数据模型

### 1.HBase的逻辑模型：

- **优点：**列存储不同于传统的关系型数据库，其数据在表中是按行存储的，列方式所带来的重要好处之一就是，由于查询中的选择规则是通过列来定义的，因此整个数据库是自动索引化的。按列存储每个字段的数据聚集存储，在查询只需要少数几个字段的时候，能大大减少读取的数据量，一个字段的的数据聚集存储，那就更容易为这种聚集存储设计更好的压缩/解压算法。下图讲述了传统的行存储和列存储的区别：

## 第二部分：HBase的数据模型



## 第二部分：HBase的数据模型

### 2.HBase的逻辑数据模型：

- HBase是稀疏存储数据，所以一些列可以为空。图1是一个表的逻辑数据模型，表中每一条数据对应一个时间戳，用数字表示，编号越大表示数据越旧。

表 1 HBase 逻辑数据模型 <sup>(1)</sup>					
Table 1 Logic data model of HBase					
Row Key	Time Stamp	Column Family <sup>1</sup> : c1		Column Family <sup>2</sup> : c2	
		列	值	列	值
r1	t7	c1:1	value1-1/1		
	t6	c1:2	value1-1/2		
	t5	c1:3	value1-1/3		
	t4			c2:1	value1-2/1
r2	t3			c2:2	value1-2/2
	t2	c1:1	value2-1/1		
	t1			c2:1	value2-1/1

图1

## 第二部分：HBase的数据模型

### 3.HBase的物理模型：

- **物理存储表示：**HBase是按照列存储的稀疏行 / 列矩阵，在物理存储方面，它实际就是把概念模型中的一行分割，并按照列族存储。图 1 中 r 1 行在物理存储时表现如图 2 所示。可见，空值未储存，节省了空间。

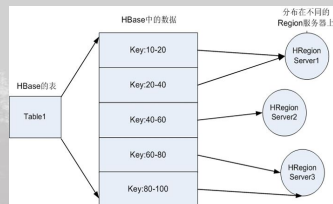
表 2 HBase 数据的物理视图 Table 2 Physical data model of HBase			
Row Key	Time Stamp	Column Family+cl	
		列	值
r1	t7	cf1+1	value1-1/1
	t6	cf1+2	value1-1/2
	t5	cf1+3	value1-1/3
	t4	cf2+1	value1-2/1
	t3	cf2+2	value1-2/2

图2

## 第二部分：HBase的数据模型

### 2.HBase的物理模型：

- **Table & Region：**当Table随着记录数不断增加而变大后，会逐渐分裂成多份splits，成为regions，一个region由[startkey,endkey)表示，不同的region会被Master分配给相应的RegionServer进行管理，如下图所示：



## 第三部分：HBase的部署架构

### 1.Client：

- 包含访问hbase的接口，Client维护着一些cache来加快对hbase的访问，比如region的位置信息

### 2.ZooKeeper：

- 保证任何时候，集群中只有一个master。
- 存储所有Region的寻址入口。
- 实时监控Region Server的状态，将Region server的上线和下线信息实时通知给Master。
- 存储Hbase的schema,包括有哪些table，每个table有哪些column family。

### 第三部分：HBase的部署架构

#### 3.HMaster :

- 为Region Server分配Region
- 负责Region Server的负载均衡
- 发现失效的Region Server并重新分配其上的Region
- HDFS上的垃圾文件回收
- 处理Schema更新请求

---

---

---

---

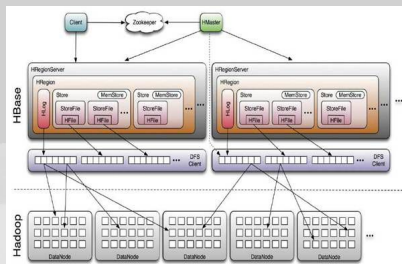
---

---

---

---

### 第三部分：HBase的部署架构




---

---

---

---

---

---

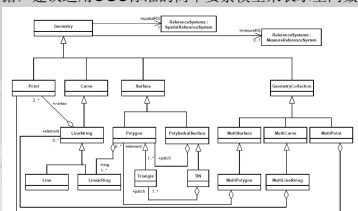
---

---

### 第四部分：HBase储存空间数据

#### 1.空间数据模型：

- 为了方便HBase储存管理空间数据，应选用合适的空间数据模型表达和组织矢量数据，建议选用OGC标准的简单要素模型来表示空间数据。




---

---

---

---

---

---

---

---

#### 第四部分：HBase储存空间数据

##### 2.空间数据物理储存：

- HBase不同于RDBMS具有丰富的数据类型和存储方式，它只有简单的字符串类型，所有的类型须用户自行处理。因此建议使用WKB的方式，对空间信息进行储存。
- WKB通过序列化的字节对象描述几何对象，具有更高的读写和存储效率，所以本文采用WKB格式对空间数据进行存储。

#### 第四部分：HBase储存空间数据

##### 3.矢量数据的存储模型：

- 我们可以把空间数据的信息分为几何信息，属性信息，拓扑信息（起点、终点、左多边形、右多边形）三部分，根据矢量数据的特点设计的基于Hbase的存储图层表结构如图3所示，表中的列族依次是几何信息列族、属性信息和拓扑信息列族，每种数据格式都是字符串类型，使用时解析成相应的数据类型。

表 4 矢量数据图层的数据字典表结构  
Table 4 Table structure of data dictionary of vector data layer

Row Key	Time Stamp	Column Family <sup>†</sup> attribute		Column Family <sup>†</sup> topo	
		info	value	info	value
Layer_ID1	t5	"fcode"	"String"		
	t4	"fname"	"String"		
	t3	"flevel"	"int"		
	t2			toponame1	"int"
t1				toponame2	"string"

图2

#### 第四部分：HBase储存空间数据

##### 4.空间索引方式：

- 在HBase中对空间数据进行索引，可根据不同的比例尺应划分不同尺度的网格。网格索引ID的值可按照Hilbert空间填充曲线来确定。下图显示的是图层的索引表结构，索引ID 作为RowKey，列族IDS表示一个网格所包含的所有空间要素的ID，不同的空间要素ID使用标准的逗号分隔取值格式。

表 5 空间数据索引表结构  
Table 5 Table structure of spatial data index

Row Key	Time Stamp	Column Family <sup>†</sup> IDS	
		列	值
Index_ID	t1	IDS*1	"...."
	t2	IDS*2	"...."

#### 第四部分：HBase储存空间数据

##### 5.HBase在空间索引生成过程中的优势：

- 传统建立空间索引通常是以空间换取时间的串行方式，效率较低。利用MapReduce编程模型并行构建矢量数据空间索引，可有效缩短索引建立的时间。
- MapReduce算法分为Map阶段和Reduce阶段。Map函数读取一条空间要素记录，计算要素所在的所有网格索引ID，并以索引ID为Key、以空间要素的ID为Value写入本地HDFS；Reduce阶段把从不同Mapper计算的索引Key-Value数据整合并进行排序，然后调用Reduce函数对输入的<索引ID, list (Fea\_ID)> 对进行合并处理，把得到的<索引ID, Fea\_IDS>写入相应比例尺、相应图层的HBase索引表中。

---

---

---

---

---

---

---

---

#### 第四部分：HBase储存空间数据

##### 6.HBase在海量空间数据存储和查询过程中的优势：

- 随着空间数据存储量增加，HBase自动将存储的数据拆分为不同的Region块，由一个RegionServer管理
- 数据录入时，利用MapReduce框架实现数据高速录入，高效率创建空间索引。
- 实现海量空间数据的空间查询操作。

---

---

---

---

---

---

---

---

以上是本周学习报告

请批评指正

2015年11月25日

北京大学城市与环境学院

尹懿行

交流邮箱：501948754@qq.com

---

---

---

---

---

---

---

---