stratebi
open business intelligence

Hbase – Up & Running

H·BASE

Índice

- Presentación Stratebi
- Introducción
  - ¿Qué es?
  - Características Principales
- Instalación & Configuración
- Comandos

Sobre nosotros - Stratebi

www.stratebi.com

stratebi
open business intelligence

Confían en Business Intelligence Open Source

Sector Privado

BBVA
EROSKI
Blanco
Redsýs ServiRed
YELL
sage
AMERSAM
amper
SIGAUS
MRW
pepecar.com
opinat
20:20mobile
MútuaTerrassa
EQUIFAX
BOSCH
accenture High performance. Delivered.
PKW
ASMIUS
Italfarmaco
HAMBURG SÜD
ELDON
Unilever
educaweb

Sector Público

ICEX
Agencia Tributaria
SERVICIO CANARIO DE SALUD
SaludMadrid
UOC www.uoc.edu
Ayuntamiento de Arrecife
Instituto Cervantes
Diputació de Tarragona
Extremadura
CSIC
Port de Barcelona
EOI Escuela de organización industrial
Mancomunidad de Montejurra Jurramendiko Mankomunitatea
REUS TRANSPORT
Dipsalut
Universitat de Lleida
Gobierno de La Rioja
Instituto Tecnológico de Canarias ITC
Txingudiko Zerbitzuak Servicios de Txingudi
ICANE Instituto Cántabro de Estadística
Govern de les Illes Balears
Caixa d'Enginyers
infoCaja

stratebi
open business intelligence

Hbase – Up & Running

¿Qué es?

- Es una base distribuída, no relacional, y open source modelada después de Google BigTable y escrita en Java.
- Está desarrollada bajo el parágüas del proyecto Apache Hadoop de la fundación Apache Software y se ejecuta bajo HDFS (Hadoop Distributed Filesystem), ofreciendo posiblidades como las de BigTable pero para hadoop.
- Esto, nos proporciona una escasa tolerancia a fallos cuando estamos almacenando grandes volúmenes de datos.

Características Principales

- Las características principales de Hbase son la compresión, operaciones en memoria, y Bloom filters sobre una columna en concreto.
- Las tablas en Hbase pueden servir de input y output para los trabajos MapReduce que se ejecuten en Hadoop, y pueden ser accedidos mediate una API Java pero también mediante REST, Avro o Thrift.
- Hbase no es un reemplazo de la base de datos clásica SQL, aunque su rendimiento ha aumentado considerablemente y empresas como Facebook lo utilizan para su plataforma de mensajería.

Requisitos mínimos

- Ubuntu 32 bits
- Java 6 sun
- Hadoop, Zookeepers instalados y configurados

Variables de entorno y creación de directorios

- Paso1. Vamos al servidor que actuará como Hmaster con el usuario hduser
- Paso 2. Crear variables de entorno:

```
$ export HADOOP_HOME=/home/hduser/HP020/hadoop
$ export HBASE_HOME=/home/hduser/HP020/hbase
```

- Paso 3. Ir al directorio home de hduser y crear el directorio HP020:

```
$ cd ~/
$ mkdir HP020
$ cd HP020
```

Descargar Hbase

- Paso 4. Crear directorio ZIP, cambiar a el y descargar Hbase:

```
$ mkdir ZIP
$ wget http://archive.apache.org/dist/hbase/hbase-
0.90.0/hbase-0.90.0.tar.gz
```

Descomprimir y modificar hbase-env.sh

- Paso 5. Descomprimir, colocar en la ruta HP020 e ir a hbase:

```
$ tar xfz hbase-0.90.0.tar.gz
$ mv hbase-0.90.0 ../hbase
$ cd ../hbase
```

- Paso 6. Ir al directorio conf y modificar el archivo hbase-env.sh. Añadir las siguientes líneas:

```
export JAVA_HOME=/usr/lib/jvm/java-6-sun
export HBASE_LOG_DIR=/home/hduser/HP020/hbase/logs
export HBASE_MANAGES_ZK=false
```

Modificar hbase-site.xml (I)

- Paso 7. Modificar el archivo hbase-site.xml. Añadir:

```
<property>
<name>hbase.rootdir</name>
<value>hdfs://namenode:8020/hbase</value>
</property>
<property>
<name>hbase.cluster.distributed</name>
<value>true</value>
</property>
<property>
```

Modificar hbase-site.xml (II)

```
<name>hbase.tmp.dir</name>
 <value>/home/hduser/HP020/hbase/var</value>
</property>
<property>
 <name>hbase.zookeeper.quorum</name>
 <value>zookeeper1,zookeeper2</value>
</property>
<property>
 <name>zookeeper.session.timeout</name>
 <value>60000</value>
```

Modificar hbase-site.xml (III)

```
</property>
<property>
 <name>hbase.zookeeper.property.dataDir</name>
 <value>/home/hduser/HP020/zookeeper/var/data</value>
</property>
<property>
 <name>dfs.support.append</name>
 <value>true</value>
</property>
```

Copiar librerías de Hbase a Hadoop

- Paso 8. Copiar librería de Hbase al namenode, y los datanodes de hadoop:

```
rsync hbase-*.jar
   hduser@namenode:/home/hduser/HP020/hadoop/lib
rsync hbase-*.jar
   hduser@datanodeX:/home/hduser/HP020/hadoop/lib
```

Copiar librerías de Hadoop a Hbase y sincronizar

- Paso 9. Copiar librería de hadoop a la carpeta lib de hbase en el hmaster. Ir al namenode y:

```
$ cd $HADOOP_HOME
$ rsync hadoop*core.jar
  hduser@hmaster:/home/hduser/HP020/hbase/lib
```

- Paso 10. Sincronizamos hbase con las regiones:

```
$ rsync -avz /home/hduser/HP020/hbase
  hregionserver1:/home/hduser/HP020/
$ rsync -avz /home/hduser/HP020/hbase
  hregionserver2:/home/hduser/HP020/
```

Utilización de herramientas web

- Paso 11. Herramientas web:
- Hmaster: En un navegador web ir a la dirección ip del hmaster al puerto 60010
- HregionserverX: en un navegador web ir a la dirección ip del hregionserverX al puerto 60030

Creación de tablas, inserción y lectura

- Crear una tabla con una sola columna:

```
create 'test', 'cf’
```

- Insertar datos en tabla:

```
put 'test', 'row1', 'cf:a', 'valor 1’
put 'test', 'row2', 'cf:b', 'valor 2’
put 'test', 'row3', 'cf:c', 'valor 3’
```

- Listar valores de la tabla 'test:

```
scan 'test’
```

Borrado de tabla y Comandos adicionales

- Obtener una sola fila:

```
get 'test', 'row1’
```

- Deshabilitar 'test' y borrar la tabla:

```
disable 'test’
drop 'test’
```

- Iniciar Hbase:

```
./bin/start-hbase.sh
```

- Salir de Hbase:

```
exit
```

- Para Hbase:

```
./bin/stop-hbase.sh
```

Preguntas

Stratebi: Quiénes somos

Madrid: Pº de la Castellana, 164, 1º
Barcelona: C/ Valencia, 63
Brasil: Av. Paulista, 37 4 andar

Mas información

info@stratebi.com
www.stratebi.com
Tfno: 91.788.34.10

TodoBI
www.TodoBI.com

stratebi
open business intelligence