Hive Code Structure

Hive yapısının SQL bağlantılı olarak çalıştığından bahsetmiştim.O yüzden hive’i kavramak benim için daha kolay oldu.Bugünkü çalışmamda Hive üzerinde çeşitli uygulamalar yaptım.Onlardan biri de Veritabanları arasında tablo taşımadır.

Veritabanları Arası Tablo Taşıma :

Öncelikle tablomuzu oluşturduk.

<strong>hive> use DB;</strong>

hive> create table rn1 ( id int);

hive> insert into rn1 values ( 100);

hive> insert into rn1 values ( 200);

hive> select \* from rn1;

hive> alter table rn1 rename to rn2;

Oluşan Tablo :

HOST]~$hadoop fs -ls  /home/NewDB/

Found 7 items

...

drwxrwx---+  - hive   supergroup          0 2016-01-29 16:20 /NewDB/DB/rn4

...

[HOST]~$hadoop fs -ls  /nEWdb/db/rn4

Found 3 items

-rwxrwx---+  3 hive supergroup          4 2016-01-29 16:22 /NewDB/DB/rn4/000000\_0

-rwxrwx---+  3 hive supergroup          2 2016-01-29 16:22 /NewDB/DB/rn4/000000\_0\_copy\_1

<strong>-rwxrwx---+  3 hive supergroup          3 2016-01-29 16:22 /NewDB/DB/rn4/000000\_0\_copy\_2

</strong>

Tabloyu dışa aktarma :

hive> export table rn4 to '/tmp/rn4';

Son olarak sıra tabloyu taşımaya geldi.

hive> import table rn42 from '/tmp/rn4';

Taşındığını gösteren komutlar :

[HOST]~$hadoop fs -ls  /NweDB/DB/

Found 7 items

...

drwxrwx---+  - hive   supergroup          0 2017-08-10 16:20 /NewDB/DB/rn4

drwxrwx---+  - hive   supergroup          0 2017-08-10 16:22 /New\_DB/DB/rn42

PİG CODE STRUCTURE

Aslında Apache Pig ve Apache Hive yazmış olduğumuz kodları arka planda Java MapReduce kodlarına çevirir.Pig ve Hive kullanmamızın en önemli sebebi Java MapReduce ile kod geliştirmek zahmetlidir , bu yüzden Pig ve Hive ile daha hızlı MapReduce uygulamaları geliştirebiliriz.

Ben pig ile veritabanına veri ekleme işlemi yapmak istedim.Sitelerde araştırmalarımı yaptıktaın sonra gerekli kodları yazdım.

Employee verisi Department verisi

1,Sales



2,Marketing

3,Engineering

Pig Join kodu

CUSTOMER = LOAD '/user/data/customerlog' USING PigStorage(',') AS

(

  name: chararray,

  age: int,

  dept\_id:int

);

DEPARTMENT = LOAD '/user/data/departmentlog' USING PigStorage(',') AS

(

  dept\_id:            int,

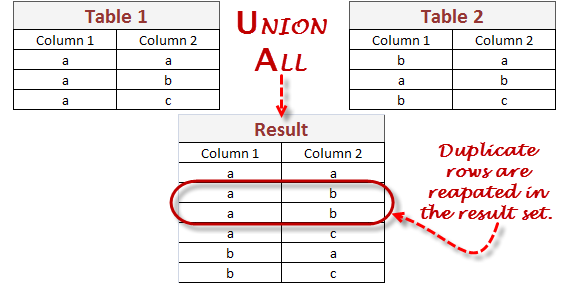
  dept\_name:    chararray

);

JOINED\_DATA = JOIN CUSTOMER BY dept\_id , DEPARTMENT BY dept\_id;

DUMP JOINED\_DATA;

#### **UNION**



customer log

http://www.buyukveri.co/apache-pig-join/