# cartesian( ) Transformation

Return the cartesian product of this RDD and another one, that is, the RDD of all pairs of elements (a,b) where a is in the first RDD and b is in the second RDD.

In [1]:
```
from pyspark.sql import SparkSession

spark = SparkSession \
        .builder \
        .master("local[2]") \
        .appName("cartesian Transformation") \
        .enableHiveSupport() \
        .getOrCreate()

#First RDD
a = [('a', 2.5), ('b', 4.0), ('c', 5.5) ]

rdd1 = spark.sparkContext.parallelize(a)
print(rdd1.collect())

#Second RDD
b = [('u', 3.5), ('v', 4.0), ('x', 2.5), ('y', 3.5), ('z', 10.0) ]

rdd2 = spark.sparkContext.parallelize(b)
print(rdd2.collect())
```

```
22/10/14 01:32:39 WARN Utils: Your hostname, Vaishalis-MacBook-Pro.local resolves to a loopback address: 127.0.
0.1; using 192.168.0.105 instead (on interface en0)
22/10/14 01:32:39 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
22/10/14 01:32:40 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builti
n-java classes where applicable
22/10/14 01:32:41 WARN Utils: Service 'SparkUI' could not bind on port 4040. Attempting port 4041.
22/10/14 01:32:41 WARN Utils: Service 'SparkUI' could not bind on port 4041. Attempting port 4042.
[Stage 0:>                                                          (0 + 2) / 2]
[('a', 2.5), ('b', 4.0), ('c', 5.5)]
[('u', 3.5), ('v', 4.0), ('x', 2.5), ('y', 3.5), ('z', 10.0)]
```

In [2]:
```
cartesianValue = rdd1.cartesian(rdd2)
print("count = ", cartesianValue.count())
print("resultant RDD =", cartesianValue.take(5))
```

```
count =  15
resultant RDD = [(('a', 2.5), ('u', 3.5)), (('a', 2.5), ('v', 4.0)), (('a', 2.5), ('x', 2.5)), (('a', 2.5),
('y', 3.5)), (('a', 2.5), ('z', 10.0))]
```