

Ans:→

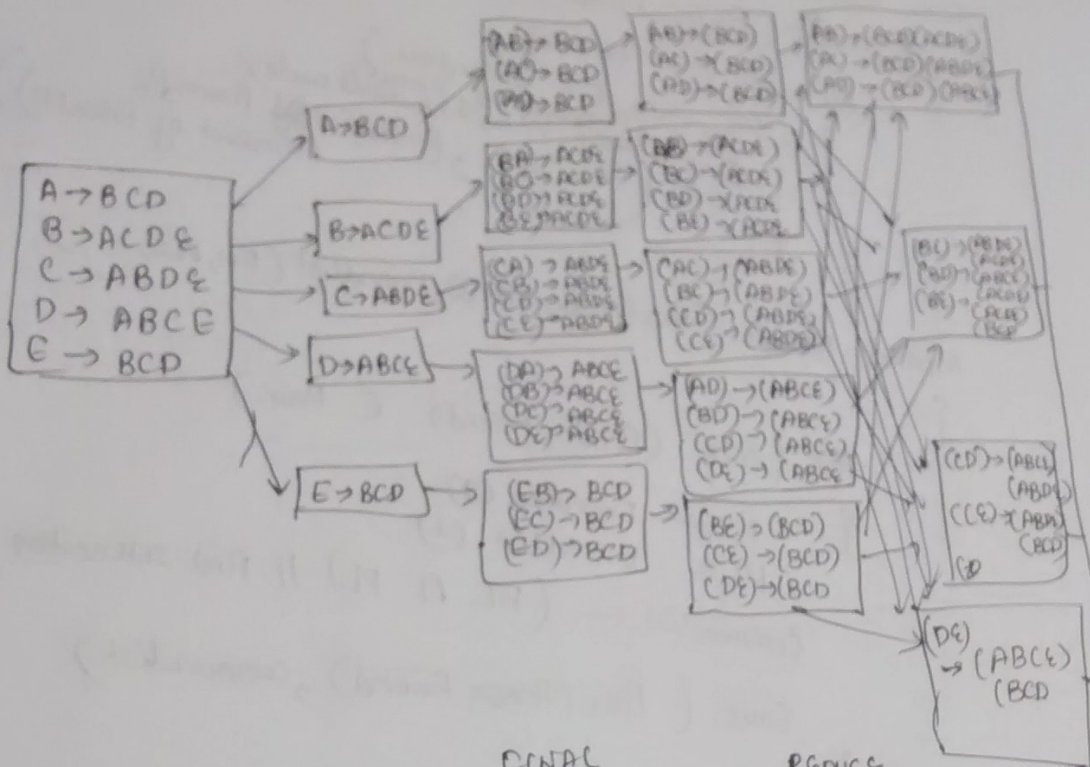
Input

SPLIT

MAPPING

SORTING

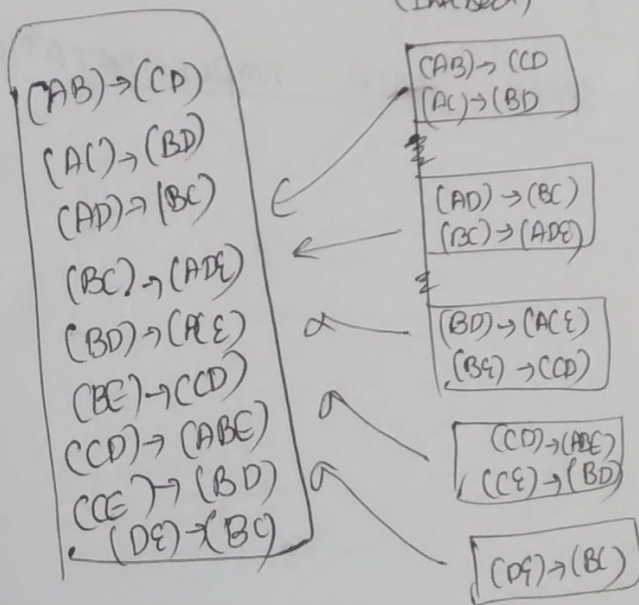
GROUP



FINAL RESULT

REDUCE

(Intersect)



② MapReduce Algorithm

class Mapper

{ method Map (Person_{list}, FriendList)

{ (for each P ∈ Person_{list})

{ (for each F ∈ FriendList)

FL ← (FriendList of Friend)

Sort P and F in alphabetical order

emit (Pair (P, F), (FriendList of Person P, FriendList of friend F))

}

Method Reduce (Pair (Person, Friend), FriendList (FL, PL, ...))

{

For (each (Person, Friend) ∈ Pair)

{

FL ← FriendList (2)

PL ← FriendList (1)

Common_List ← (FL ∩ PL) // find Intersection

Emit (Pair (Person, Friend), common_list)

}

}

③ SPARK SCALA IMPLEMENTATION

PairFlatMap function $\langle T, K, V \rangle$

$T = \text{Iterable} \langle \text{Tuple2} \langle K, V \rangle \rangle$

Reduced():

// key = Tuple (user1, user2)

// values = List ((list (user))

Reduce (key, values)

{

commonFriends = intersection (values);
emit (key, friends);

}

JavaPair Rdd $\langle \text{Tuple2} \langle \text{K}, \text{V} \rangle, \text{Iterable} \langle \text{Iterable} \rangle \rangle$

grouped = Pairs.groupByKey();