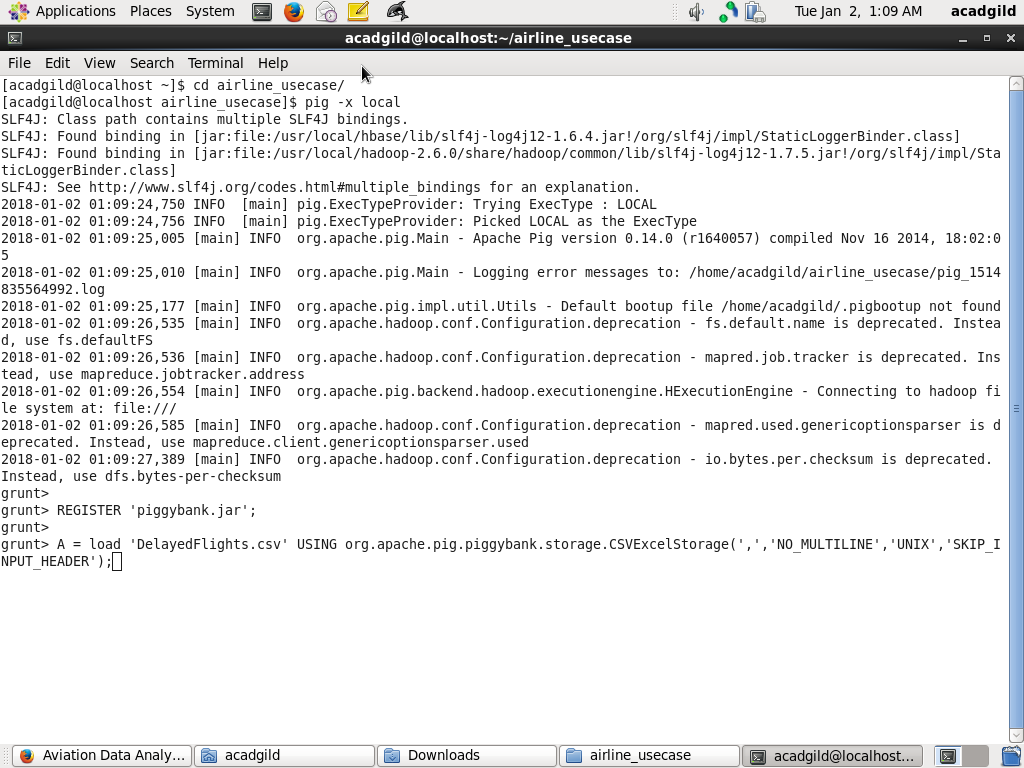
**Problem Statement 4**

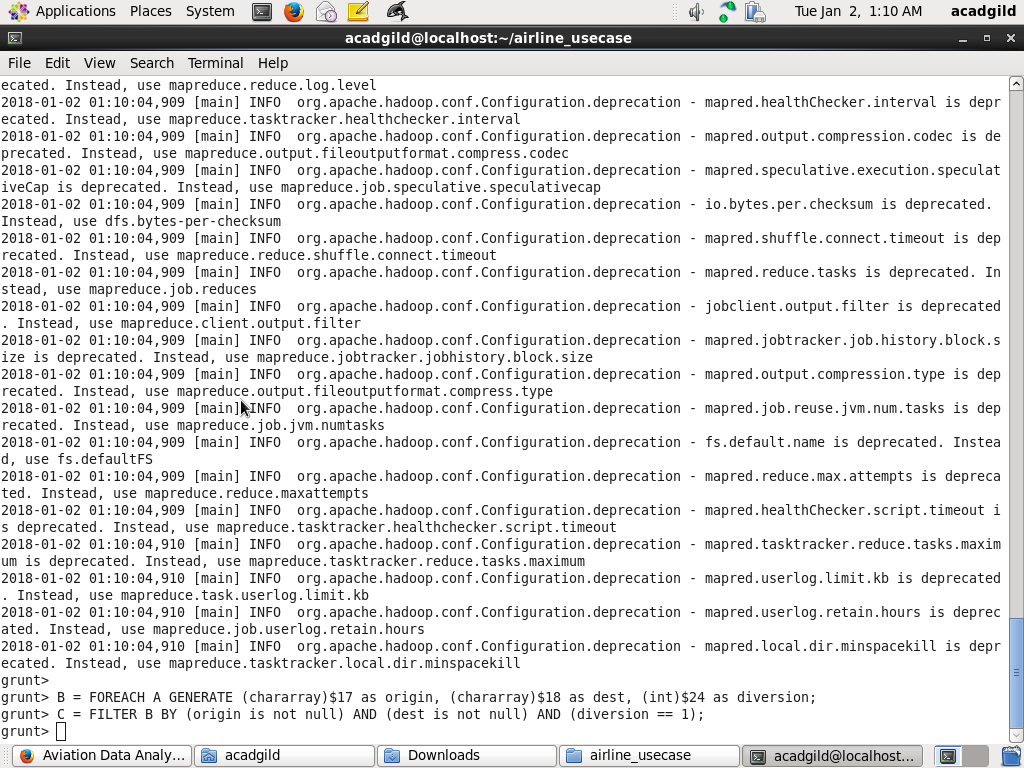
Which route (origin & destination) has seen the maximum diversion?

1. **REGISTER '/home/acadgild/airline\_usecase/piggybank.jar';**

1. **A = load '/home/acadgild/airline\_usecase/DelayedFlights.csv' USING org.apache.pig.piggybank.storage.CSVExcelStorage(',','NO\_MULTILINE','UNIX','SKIP\_INPUT\_HEADER')**

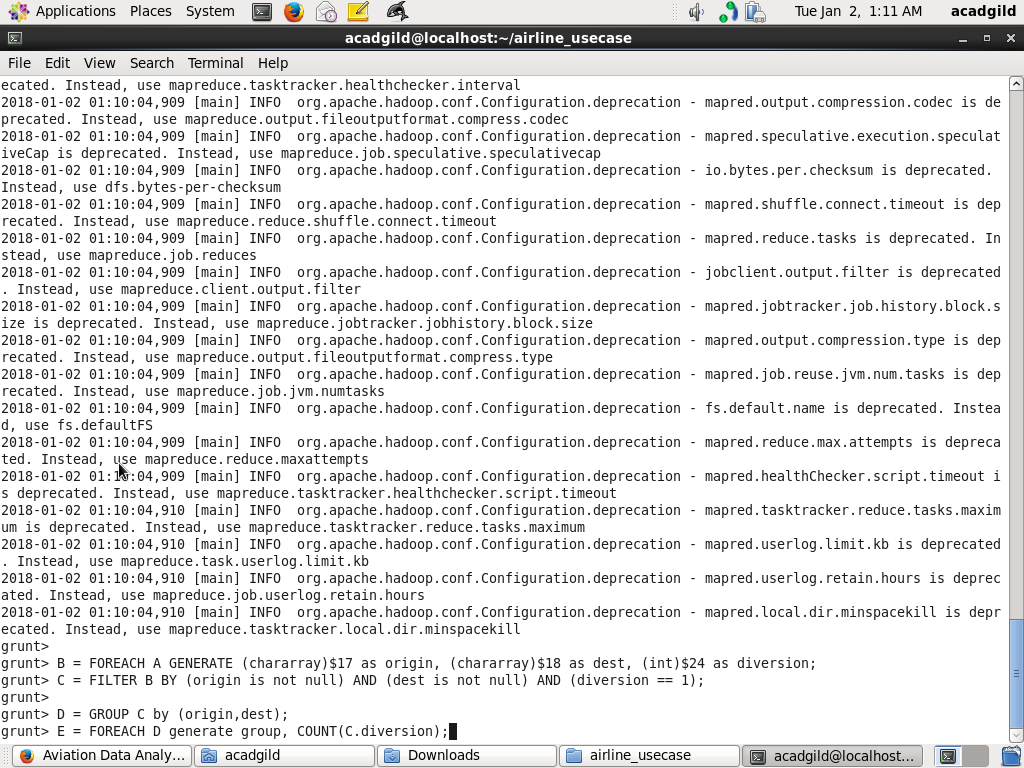


1. B = FOREACH A GENERATE (chararray)$17 as origin, (chararray)$18 as dest, (int)$24 as diversion;
2. C = FILTER B BY (origin is not null) AND (dest is not null) AND (diversion == 1);



1. D = GROUP C by (origin,dest);

1. E = FOREACH D generate group, COUNT(C.diversion);



1. **F = ORDER E BY $1 DESC;**

1. **Result = limit F 10;**

1. **dump Result;**

