
Subject: Sorting and comparing 2 files
Posted by [sanam](#) on Tue, 28 Feb 2006 23:03:19 GMT
[View Forum Message](#) <> [Reply to Message](#)

I have a scripting problem and hope you could help me.

Perl, Python or shell scripting is fine.

I have 2 files fileA is a list of assets owned by us and fileB is a list of assets held by the whole industry which includes our assets too.

I tried fgrep but it is painfully slow

```
/usr/xpg4/bin/fgrep -f sorted_car_mark_history.csv sorted_duedates.csv  
> /tmp/output.csv
```

FileA is 2456320 bytes (153520 lines)
FileB is 100028430 bytes (3334281 lines)

FileA has 2 columns separated by comma where as fileB has 4 columns separated by commas.

I need to match the 2 columns in fileA with the 2 columns in fileB (BTW the column position is 1 and 2) and for every matched record in FileB spool it out to an output file.

Sample data

FileA

```
AAMX,0000021004  
AAMX,0000021027  
AAMX,0000814223  
AAMX,0000816800  
AAMX,0000816801  
AAMX,0000816802  
AAMX,0000816803  
AAMX,0000816804  
AAMX,0000816805  
AAMX,0000816806
```

FileB

```
AA ,0000000537,DU13,20020101  
AA ,0000000537,DU58,20060130  
AAIX,0000000001,DU13,20060801
```

AAIX,0000000001,DU58,20100701
AAIX,0000000002,DU13,20060801
AAIX,0000000002,DU58,20100701
AAIX,0000000003,DU13,20060801
AAIX,0000000003,DU58,20100701
AAIX,0000000004,DU13,20060801
AAIX,0000000004,DU58,20100701

I have sorted the 2 files.

Help appreciated.
