



UNIVERSITAT DE
BARCELONA

Sistemes operatius II

Informe pràctica 5

Arnau Gris
Eric Duque

Resposta a les questions

Q1) Quines variables haurà de passar (per argument) el fil principal al fils secundaris?

El fil principal haurà de passar als fils secundaris l'arxiu a llegir, el nombre de línies N i els arxius passats per arguments com num_flights i airports.

Q2) Els fils secundaris accedeixen a recursos (variables) compartits entre ells. Quines seran les seccions crítiques? Quines parts del codi són les que s'han de protegir? Comenteu la vostra resposta.

El funcionament consisteix a llegir N línies del fitxer compartit i escriure-les a una matriu compartida. Es produeix un decrement de temps d'execució perquè mentre un fil llegeix l'altre escriu a la matriu. Les seccions crítiques que cal protegir doncs, són la lectura del fitxer i l'escriptura de la matriu de manera que els fils no interfereixin entre si.

Execució amb fils vs sense fils:

Comparem el temps d'execució del processament del fitxer amb fils i sense fils, per aquest experiment utilitzarem 2 fils i un valor N de 100.

Amb el fitxer de dades 2007.csv obtenim:

Execució amb fils:

```
Tiempo para procesar el fichero: 4.613793 segundos
```

Execució sense fils:

```
Tiempo para procesar el fichero: 8.613220 segundos
```

Com podem veure utilitzant concurrència, concretament 2 fils, obtenim un millor temps d'execució, aproximadament un 50% menys de temps.

Provem-ho ara amb el fitxer 2008.csv:

2008.csv amb fils:

```
Tiempo para procesar el fichero: 4.280476 segundos
```

2008.csv sense fils:

```
Tiempo para procesar el fichero: 8.393218 segundos
```

Amb el fitxer 2008.csv obtenim també un increment del 50% en la velocitat d'execució. Podem dir ara, que la concurrència és la millor opció en aquest cas.

Proves amb la variable N (nombre de fitxers que llegeix a la vegada):

Fitxer de dades:	N = 1	N = 10	N = 100	N = 10000
2007	5.027s	4.739s	4.589s	4.660s
2008	4.814s	4.606s	4.536s	4.643s

Com podem veure amb pocs o molts valors de N el temps d'execució augmenta, però amb un valor intermedi com 100 o 1000 obtenim el millor temps d'execució, això és degut al fet que si es tarda molt a llegir les dades, l'altre fil es queda esperant per poder llegir ell. O si el valor de N és baix és menys eficient perquè farà moltes vegades el lock.

Proves amb diferent nombre de fils secundaris:

Per realitzar aquest experiment hem modificat el nombre de CPU's de la màquina virtual a 8. Definim N=100 i fitxer de proves 2007.csv, i amb diferent nombre de fils obtenim els següents temps d'execució:

Nombre de fils	F = 2	F = 4	f = 8
Temps d'execució	4.652s	2.491s	1.544s

Com podem veure a la taula a més fils menys temps d'execució.