

### Exercice 1

Soit les relations A et B aux caractéristiques suivantes :

- A : schéma A(A1, A2), 30 tuples / page
- B : schéma B(B1, B2), 50 tuples / page.

Supposons une instance de A avec 3000 tuples, dont 480 ont la valeur 1 sur A1, et le reste ont la valeur 8 sur A1.

Supposons une instance de B avec 150 tuples, dont 100 ont la valeur 1 sur B1, et le reste ont la valeur 5 sur B1.

Supposons une jointure de A et B sur  $A1=B1$ .

Quel est le coût de la jointure pour ces deux instances avec un algorithme de jointure à base de hachage (hash join)

utilisant pour le partitionnement la fonction  $h(x) = x \bmod 3$

avec A comme relation « externe » (= ses partitions sont chargées « en premier »)

si on dispose d'un total de 4 buffers ?

Rappel : nous calculons le coût d'un algorithme de jointure en I/Os et *en ignorant le coût de l'écriture du résultat final de la jointure.*