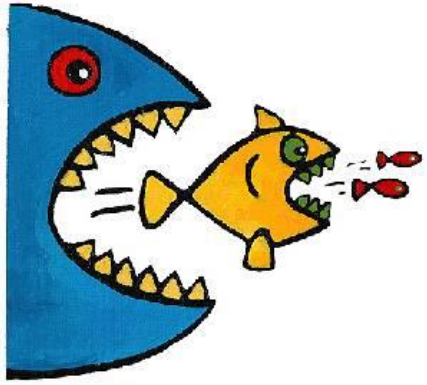


Architectures des systèmes de bases de données

Disk Merge Join

Disk Sort Merge

Tri Fusion

Conception De Base De Données

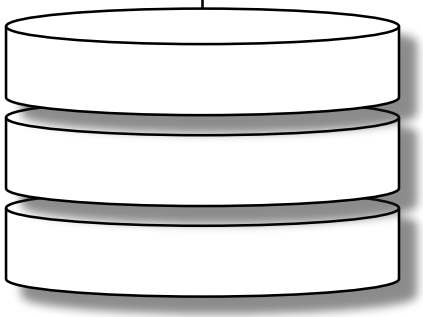
Conception Avancée De Base De Données

Java, C++, ...

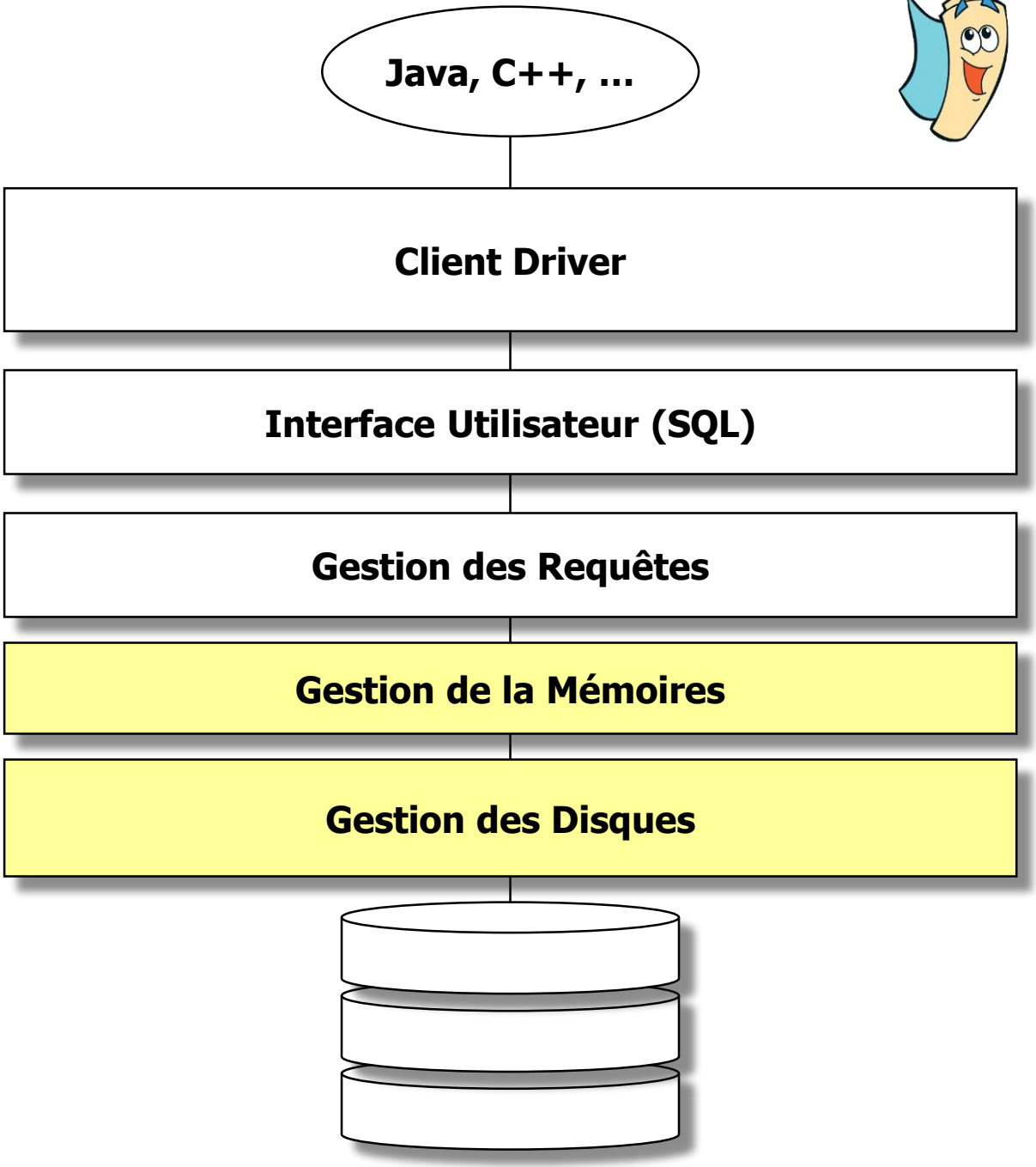
Client Driver

Interface Utilisateur (SQL)

Structure Interne

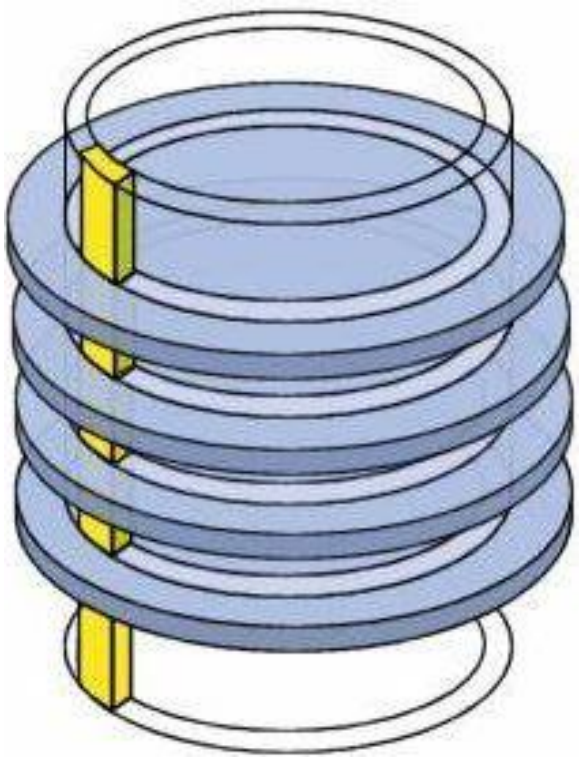


Emmanuel Fuchs

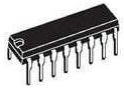


Problème.

- Que se passe-t-il si la taille d'une relation est plus grande que la taille de la mémoire ?



>



Problème.

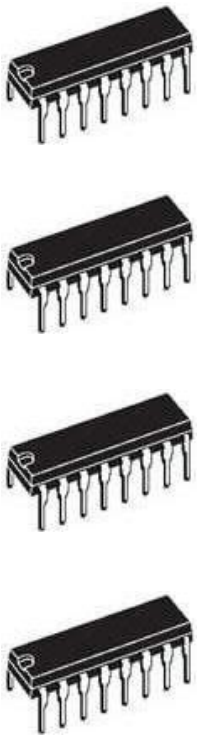
- Que se passe-t-il si la taille d'une relation est plus grande que la taille de la mémoire ?

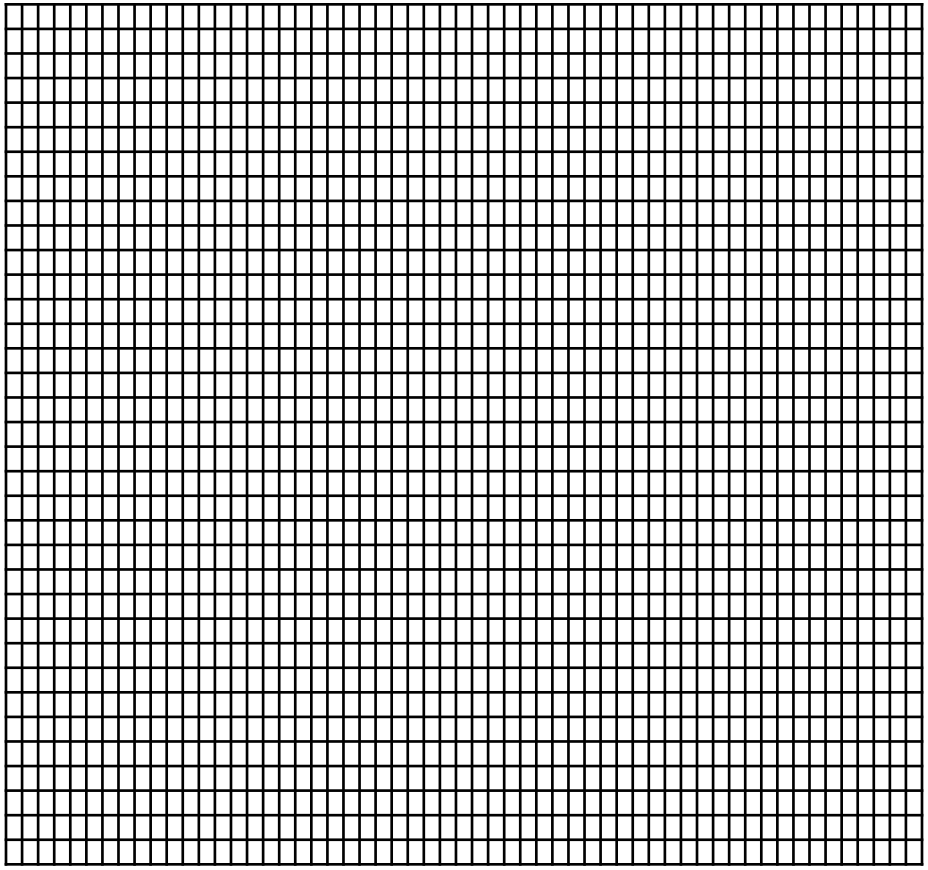
- ***Il faut découper la relation pour réaliser les opérations par « segments »***

- Un SGBD gère lui-même l'espace disque comme, et à la place, du système d'exploitation.
  - Les relations sont stockées sur disque sous forme de segments non contigus.

Big Relation => Big Data

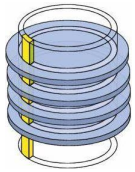
Relation



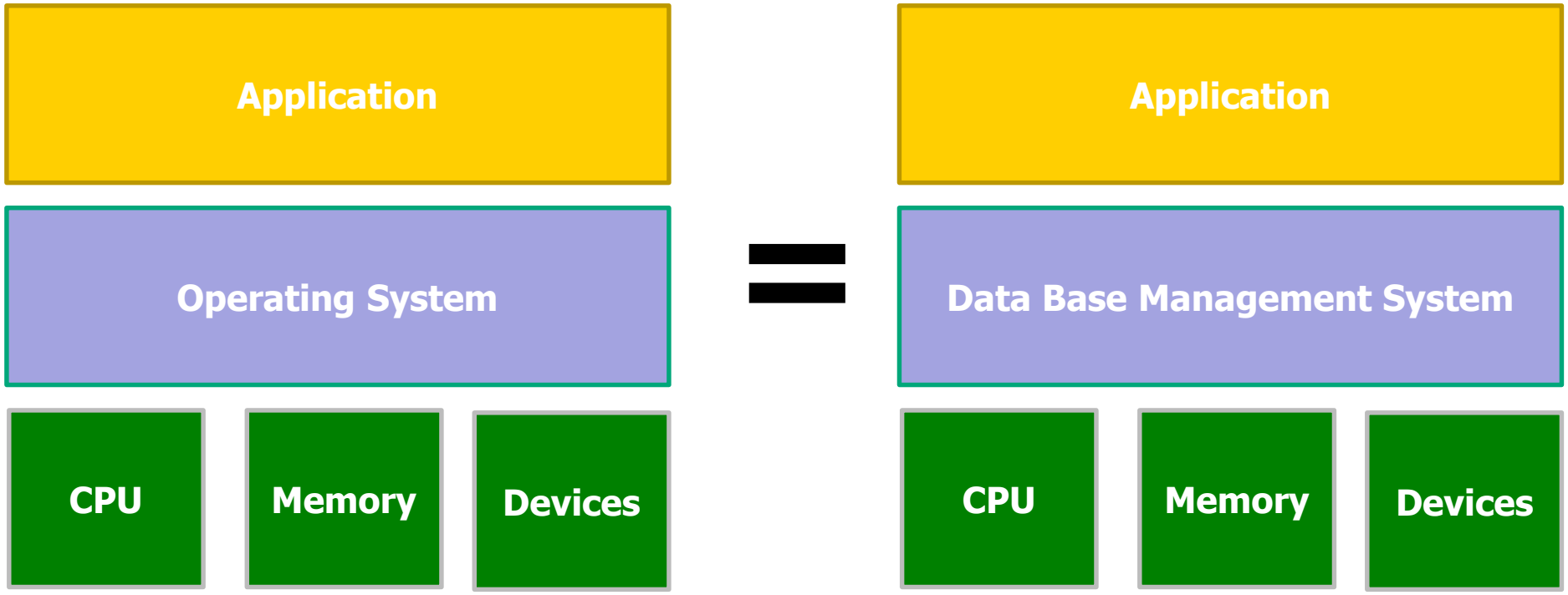


Memory

Disc



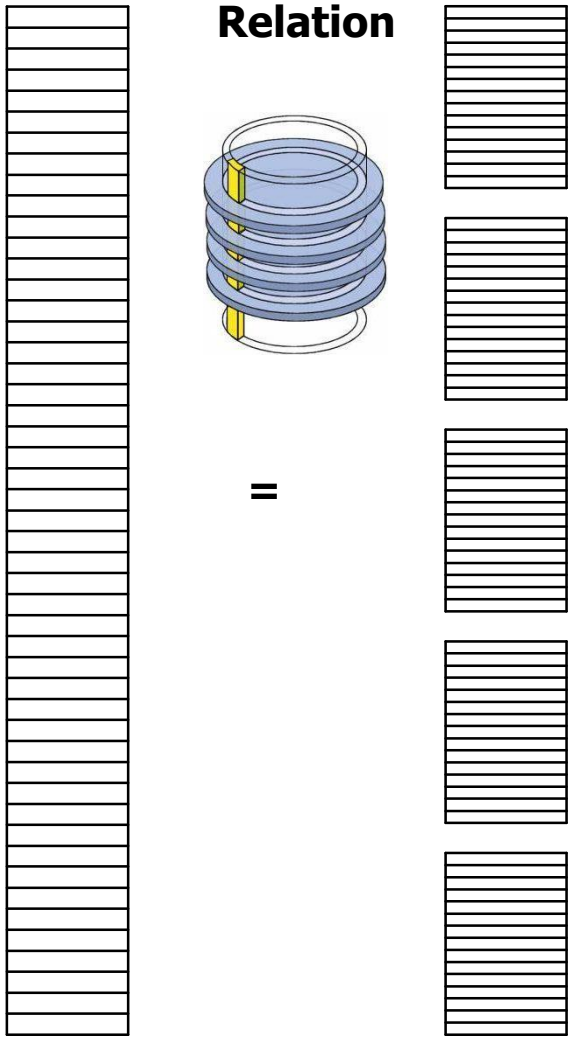
Data Base Management System

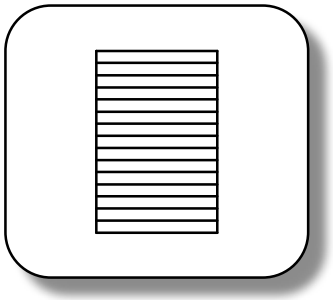


Google Big Table

http://actu.abondance.com/2008/04/google-pourrait-proposer-big-table-son.html

Transfert Disque Mémoire



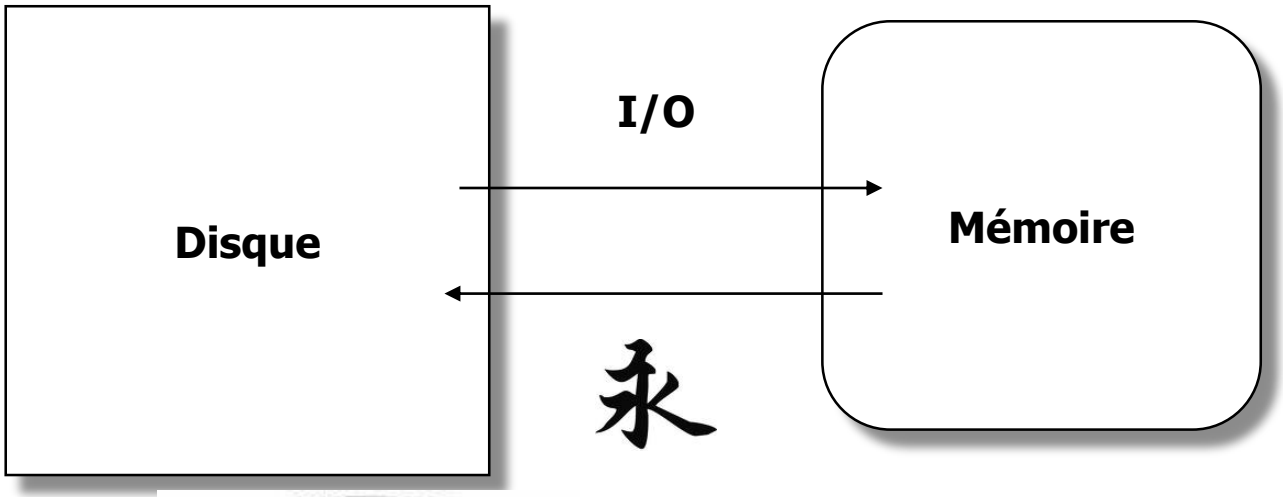


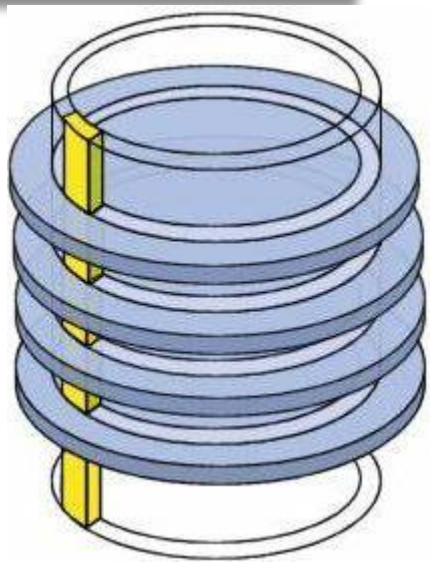
Lecture

Écriture

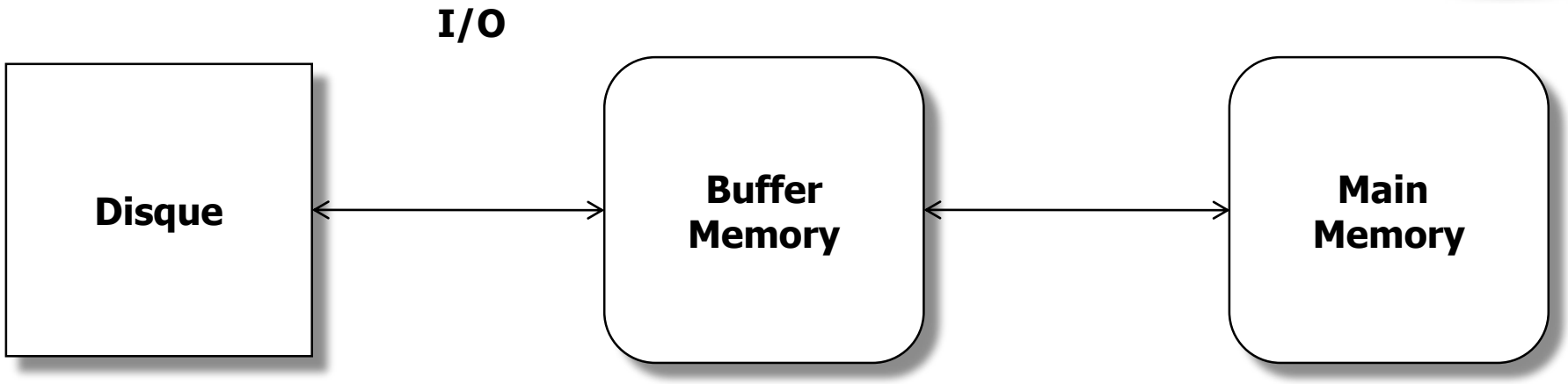
I/O

Entrées/Sorties, Lectures écritures disques, IO

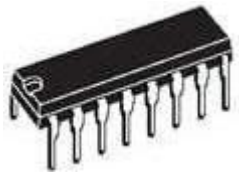


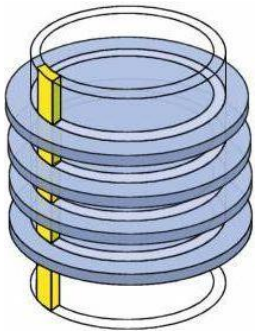


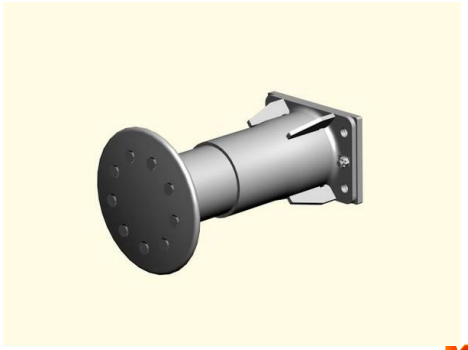
Entrées/Sorties, Lectures écritures disques, IO



永

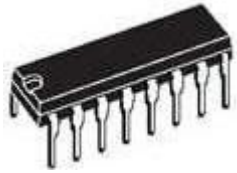




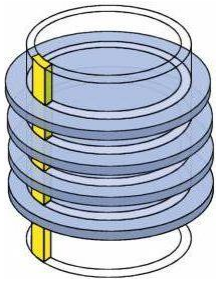


Jointures Physiques

- Jointure en mémoire
  - Nested Loop
  - Merge join
  - Hash join

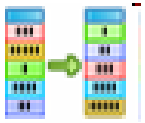


- Jointure sur disque
  - Nested Loop
  - Sort Merge Join
  - Hash join



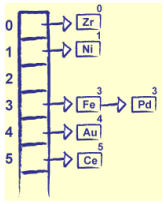
Memory join

- Nested loop

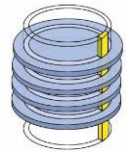


- Merge join

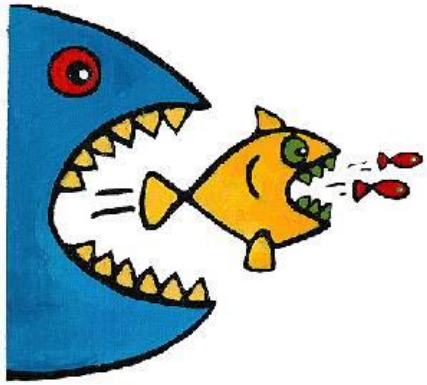
- Hash join



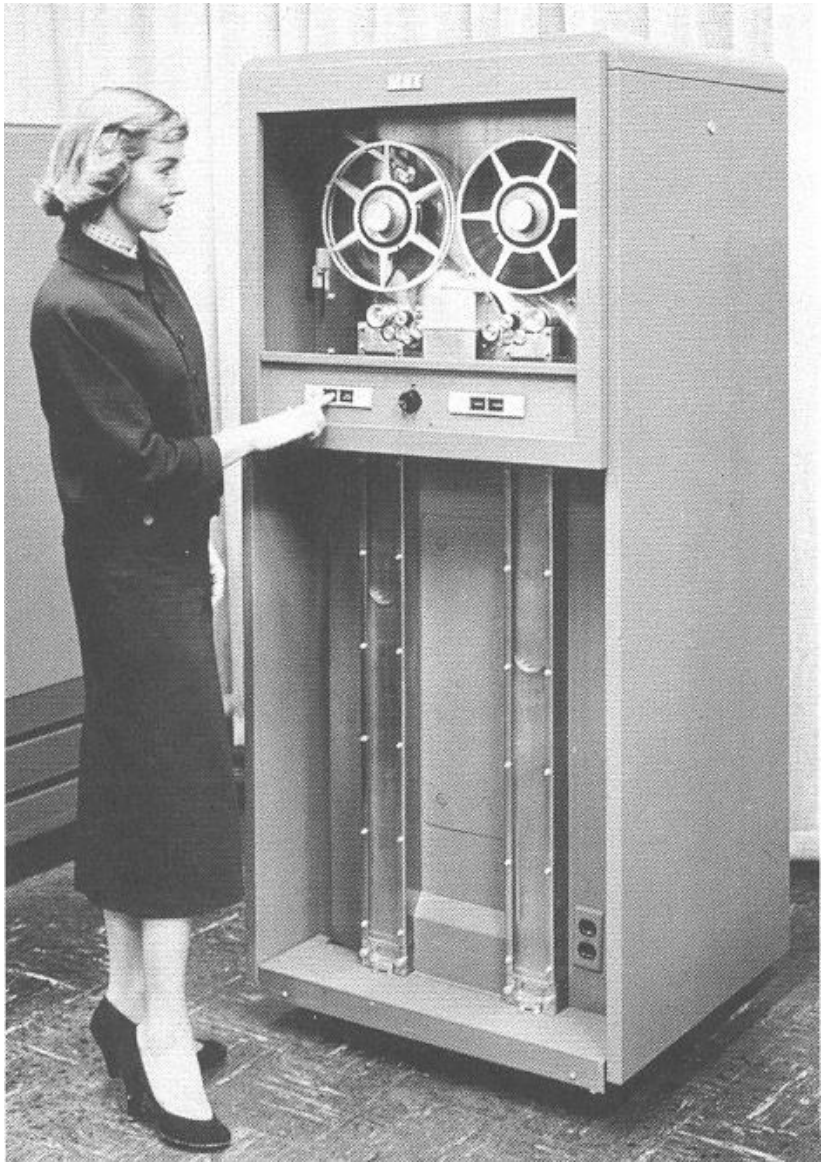
Disk Join algorithms

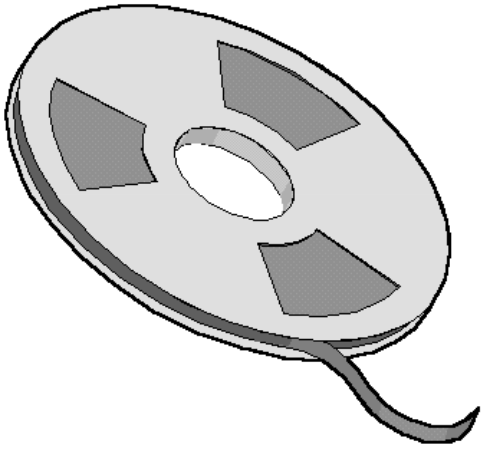


- Disk Merge join
  - Disk Sort Merge



History

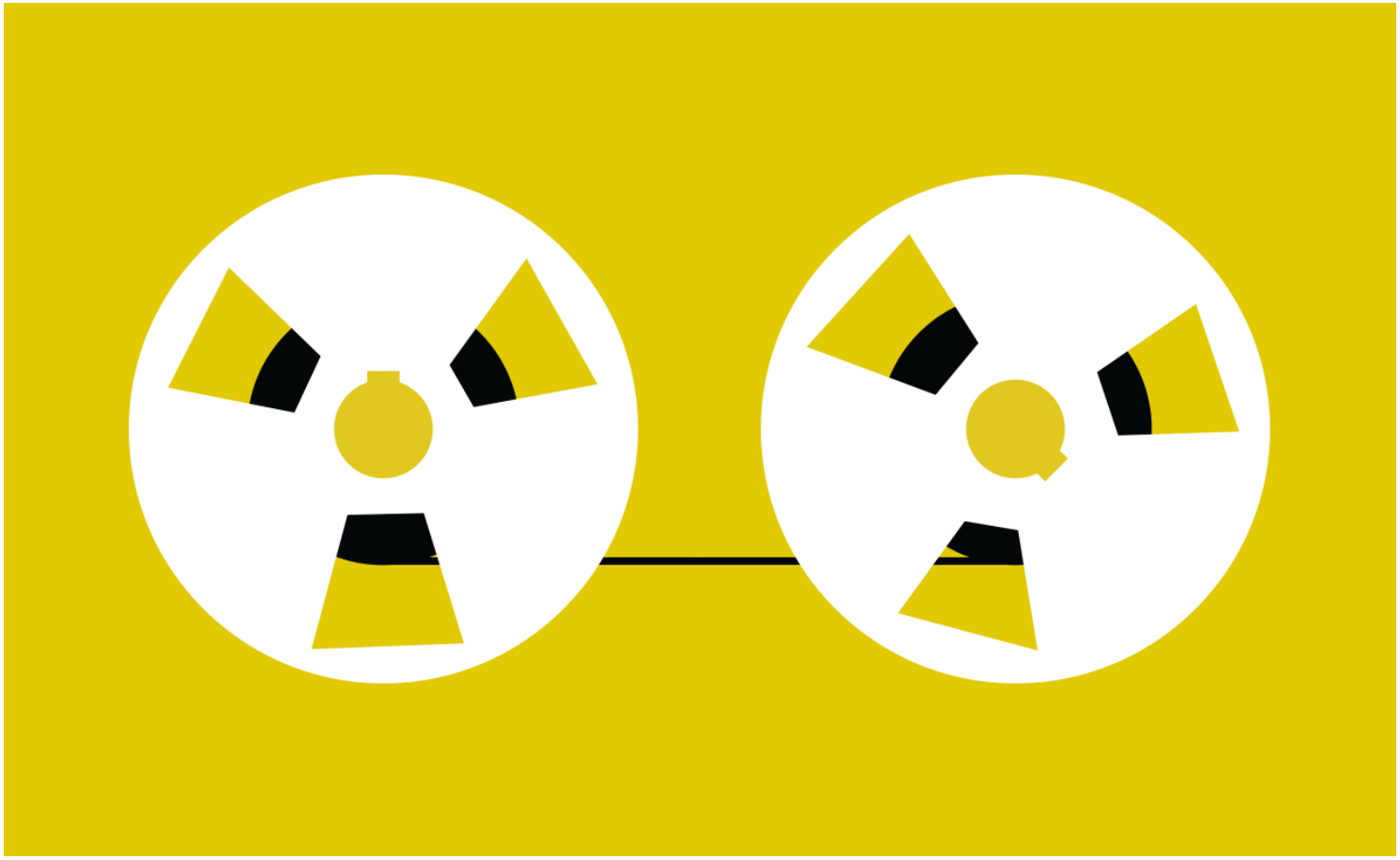




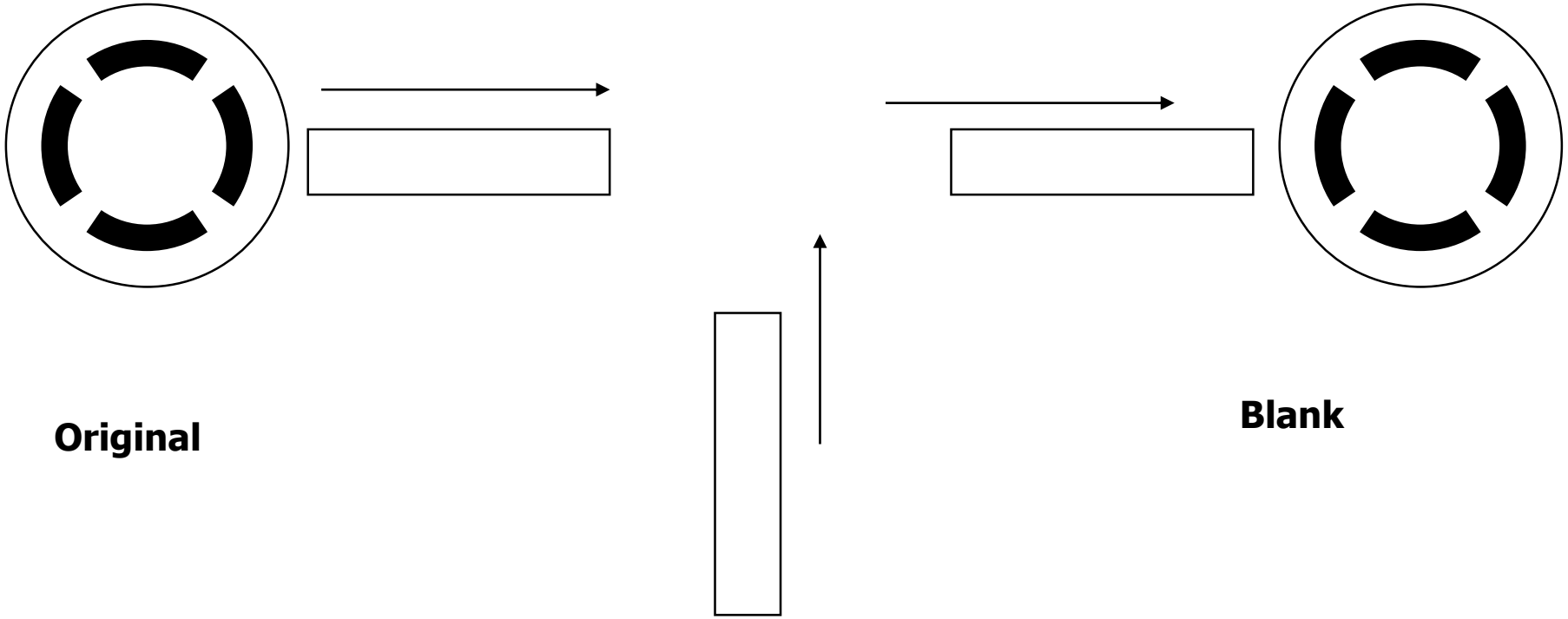
History

http://en.wikipedia.org/wiki/IBM_System/360

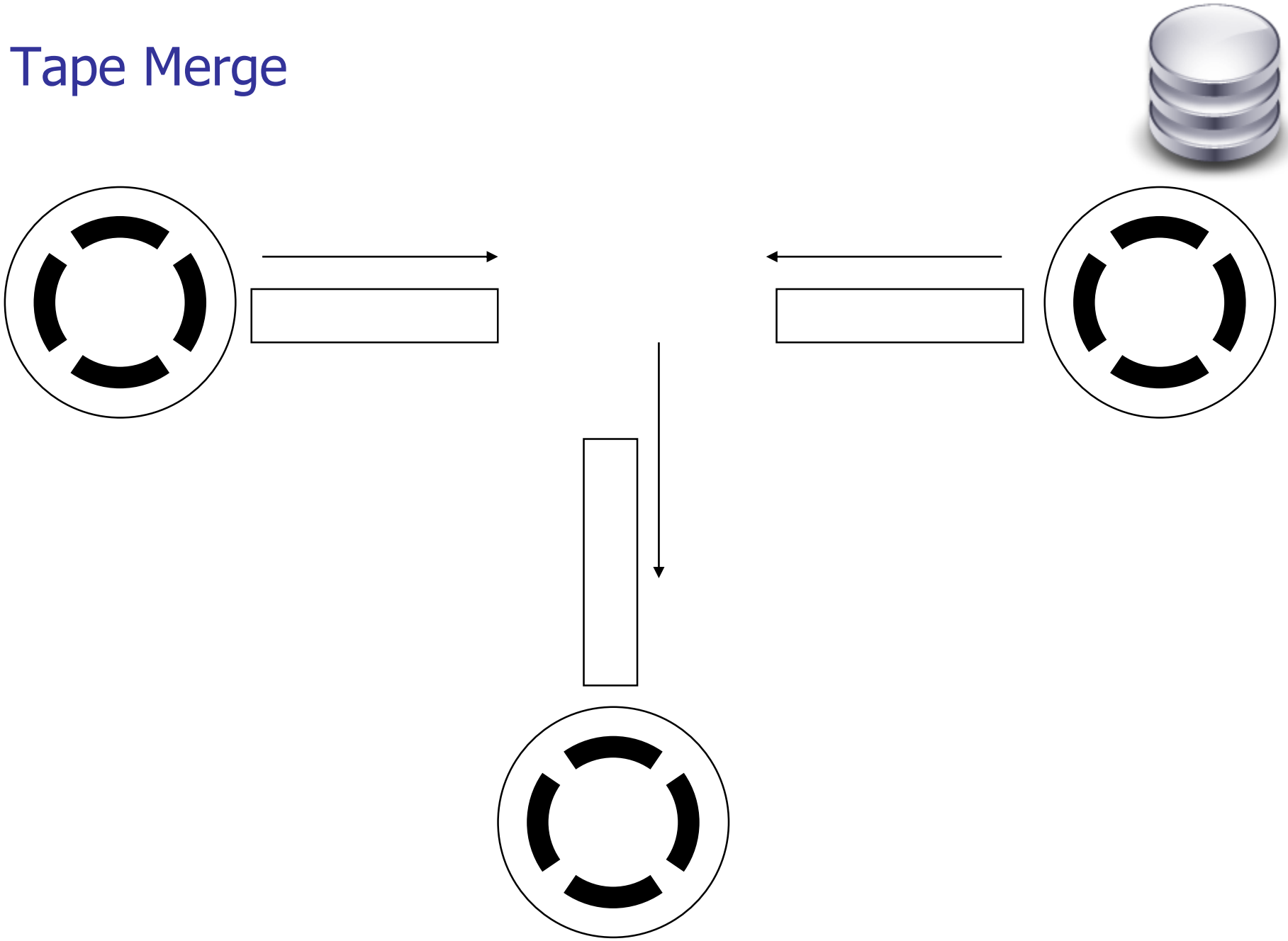
Tape Driver



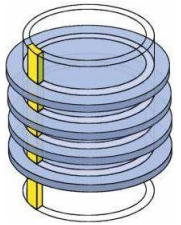
Tape Record Insert



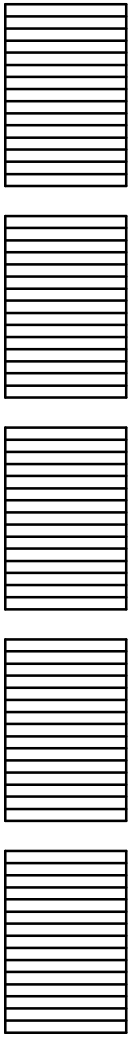
Tape Merge



Relations sur disque disques

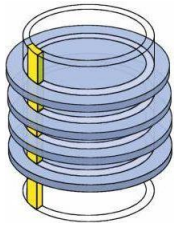


Relation



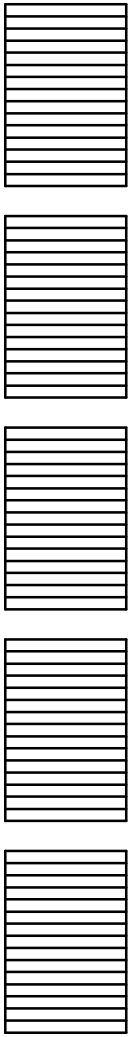
Segments
De relation

Relations sur disque disques

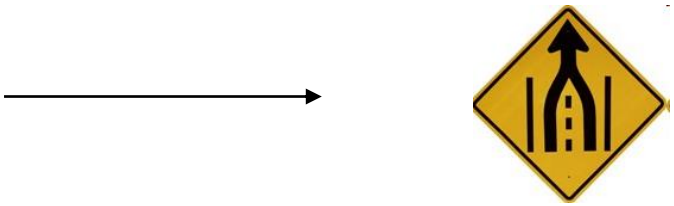


Relation

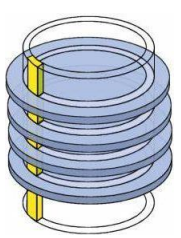
Segments
De relation



Les opérations sont réalisées sur les segments

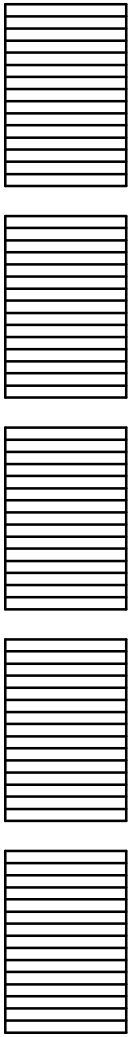


Relations sur disque

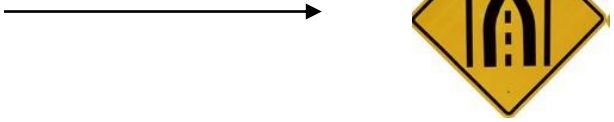


Relation

Segments
De relation



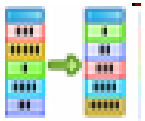
Les opérations sont
réalisées sur les segments



Chaque segment
Est chargé en mémoire
Au cours de l'opération

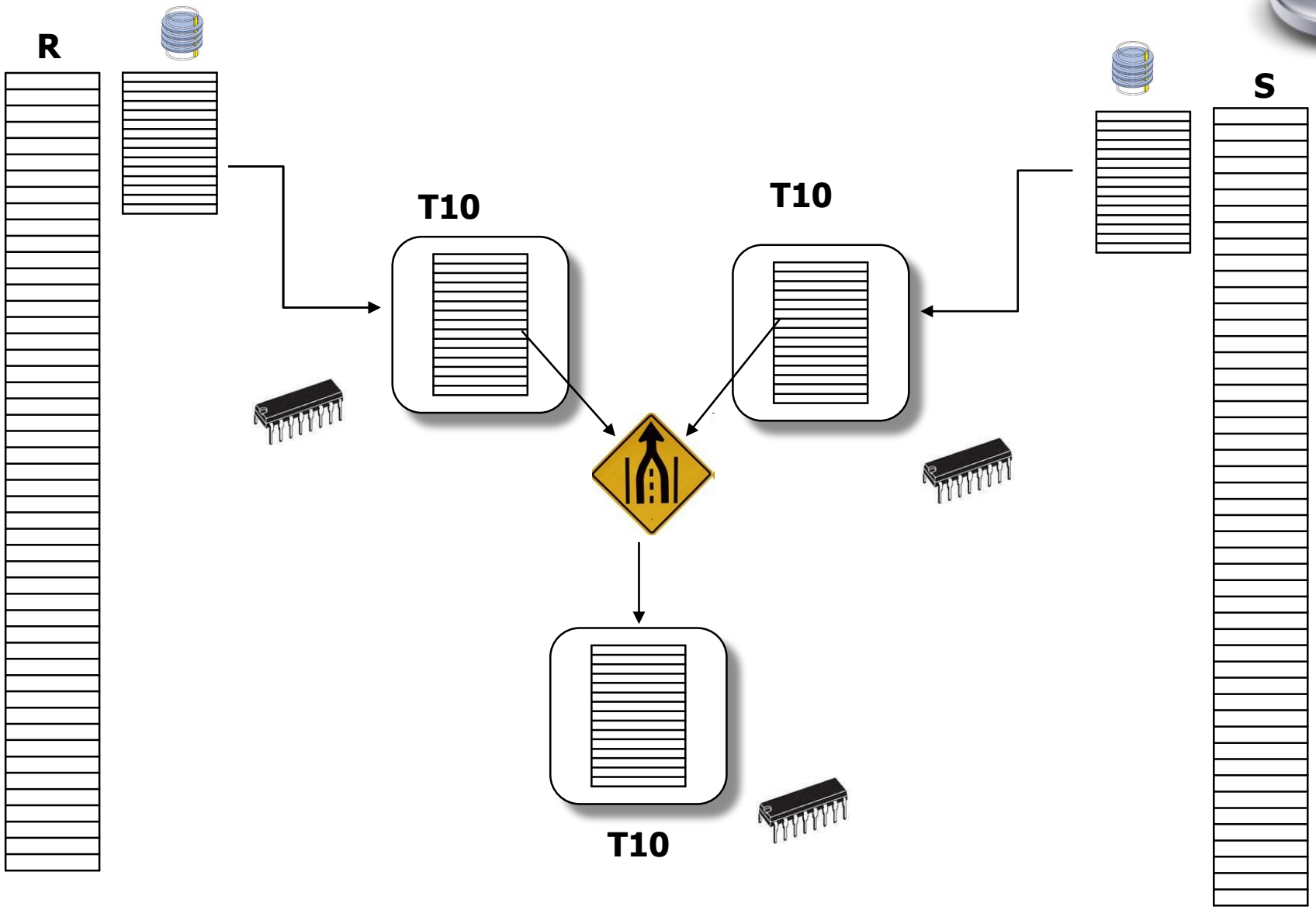
Two phases

- Disk Sort Merge

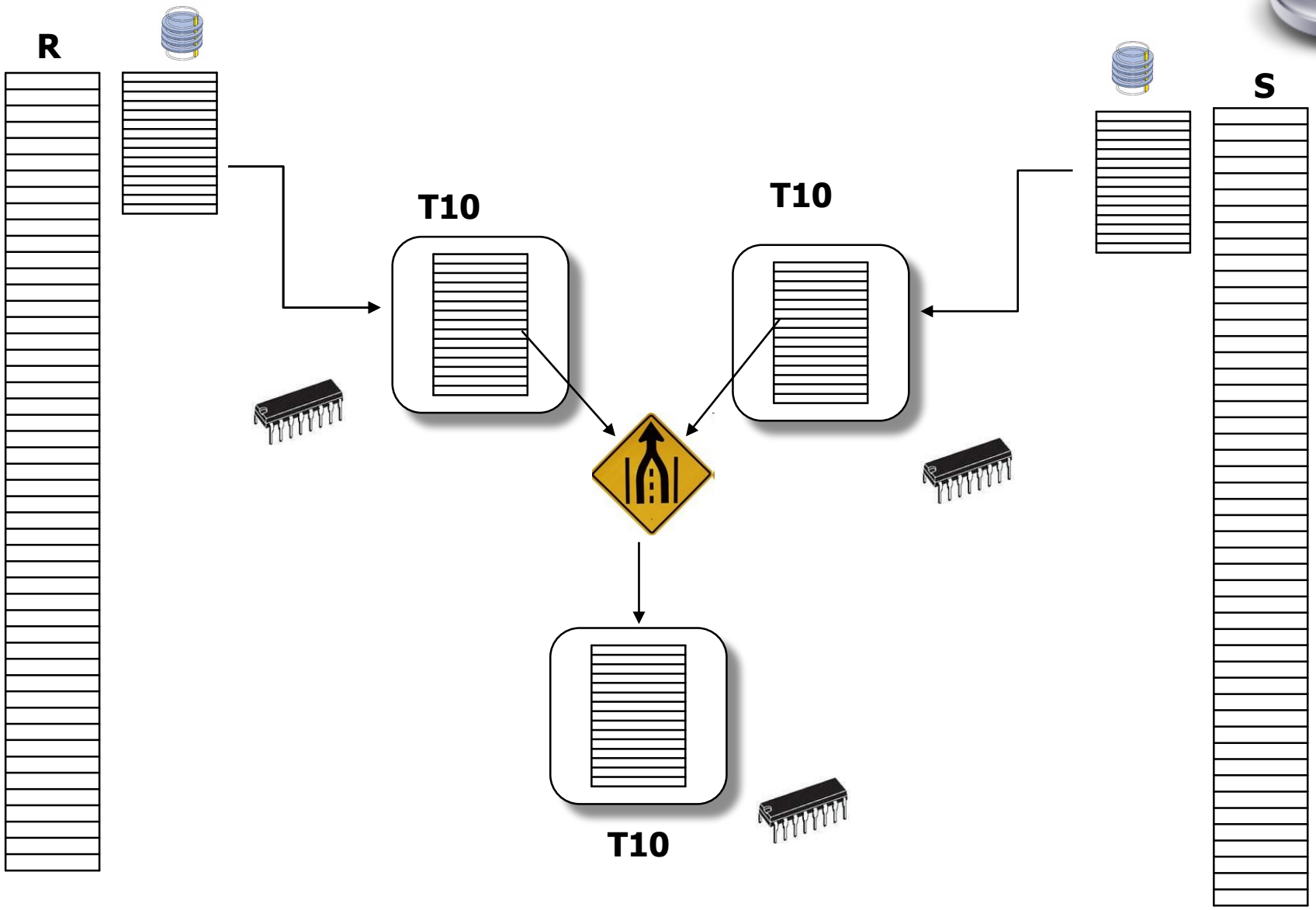


- Disk Merge Join

Merge join



Merge join



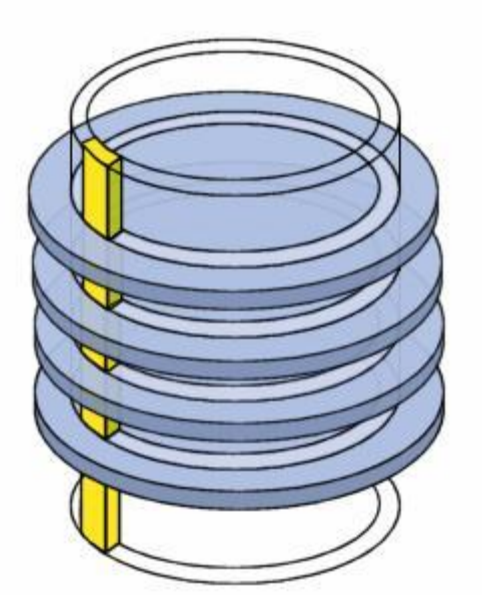
Two Phases

- (1) the sorting phase.
  - Sorting block in memory pages
    - Internal sorting

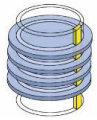
- (2) the merging phase.
  - Merging blocks in memory pages to disk block
    - Several passes

Disque

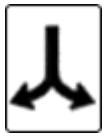
Mémoire



Non Trié	Non Trié	Non Trié	Non Trié	Non Trié	Non Trié	



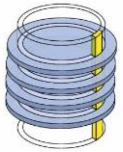
Passe 0

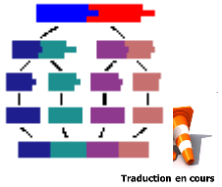


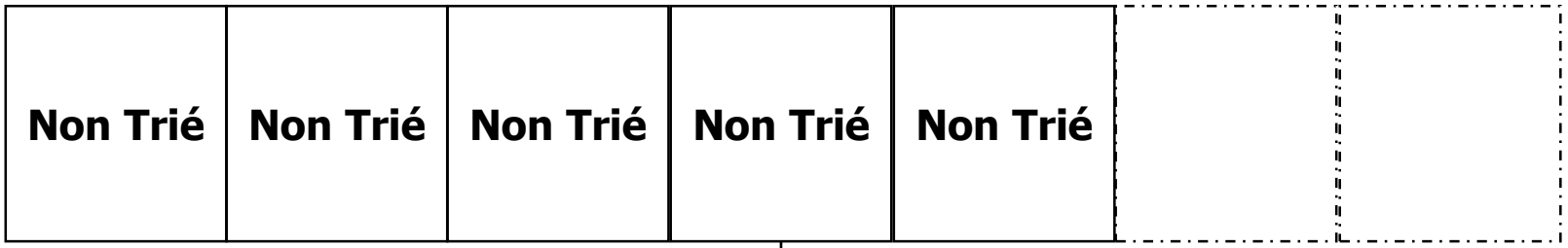
Non Trié

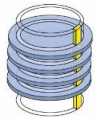
Tri d'un block à la fois
Quick Sort

Trié

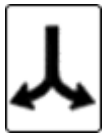




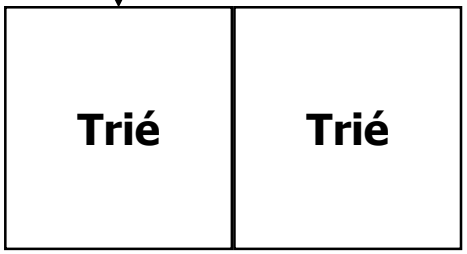


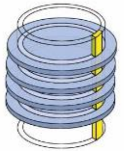


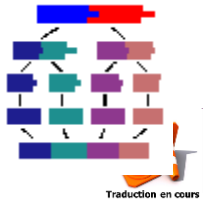
Passe 0

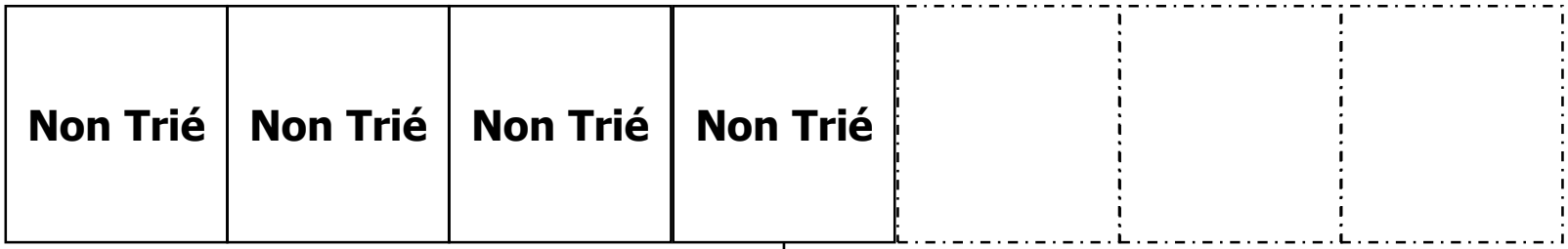


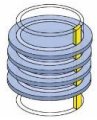
Tri d'un block à la fois
Quick Sort



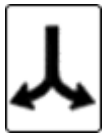




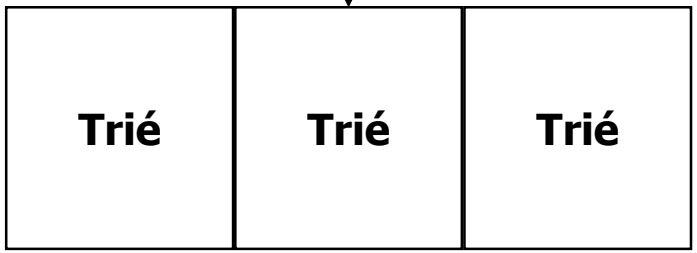


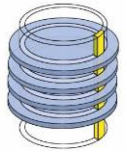


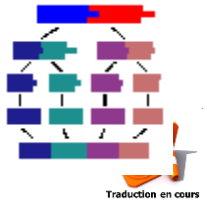
Passe 0

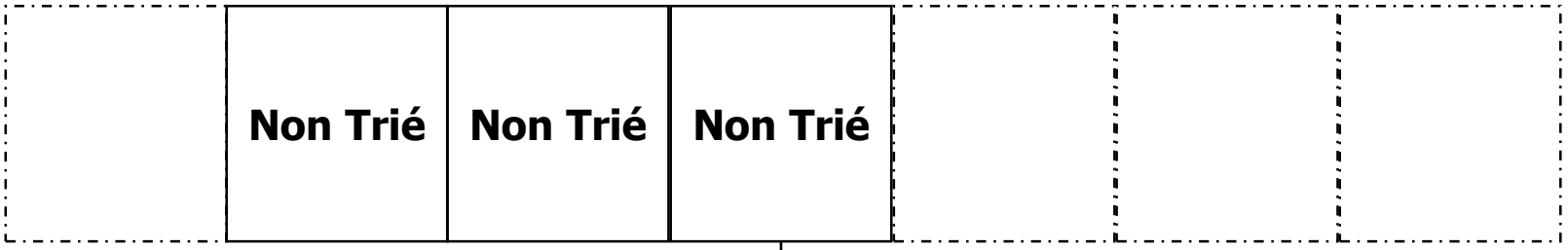


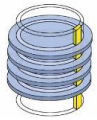
Tri d'un block à la fois
Quick Sort



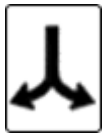






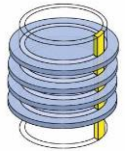


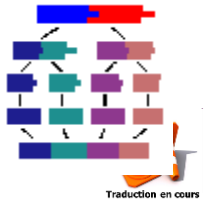
Passe 0

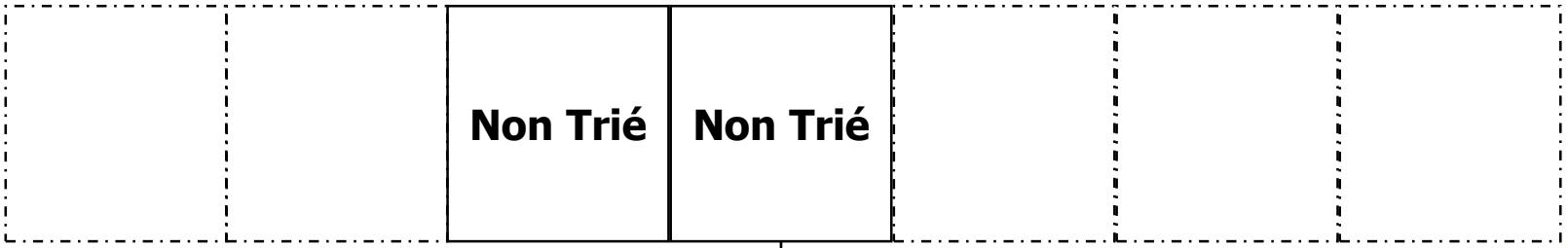


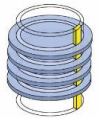
Non Trié

Tri d'un block à la fois
Quick Sort

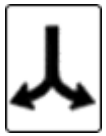




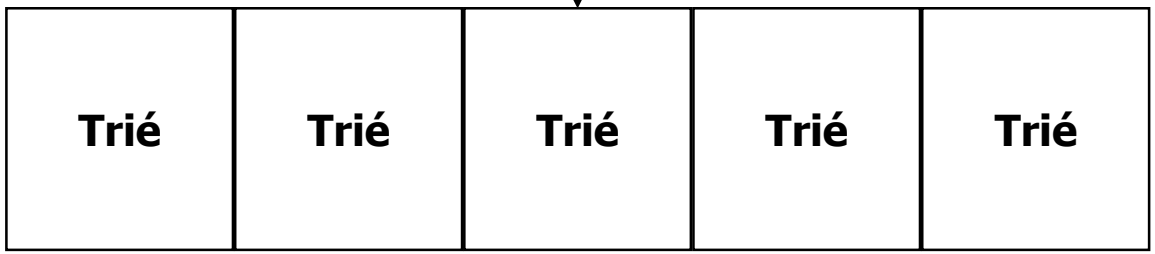


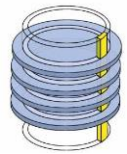


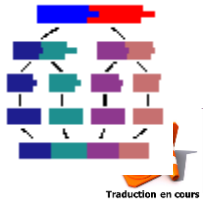
Passe 0

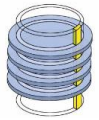


Tri d'un block à la fois
Quick Sort

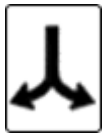




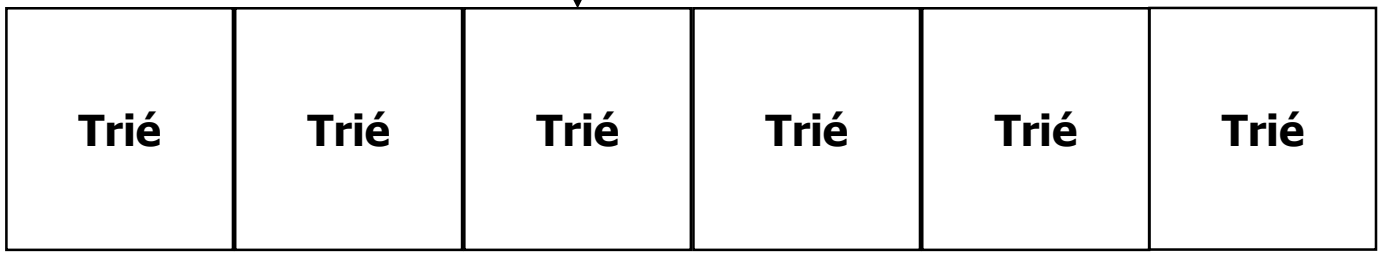


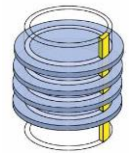


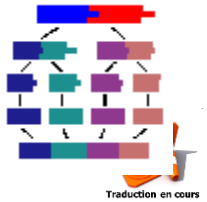
Passe 0

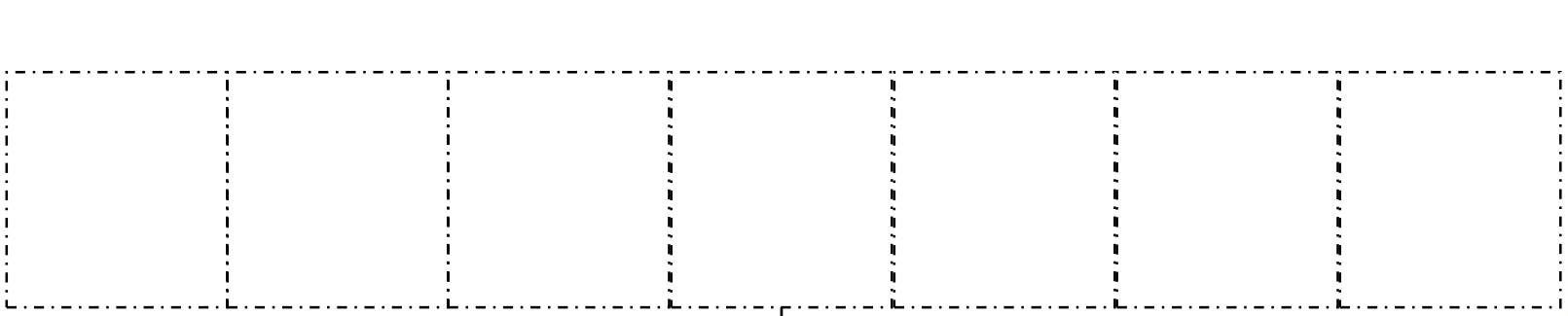


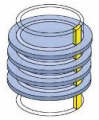
Tri d'un block à la fois
Quick Sort



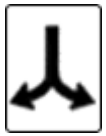








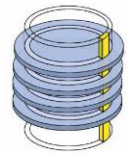
Passe 0

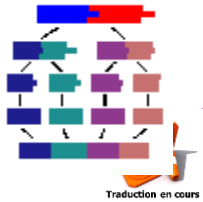


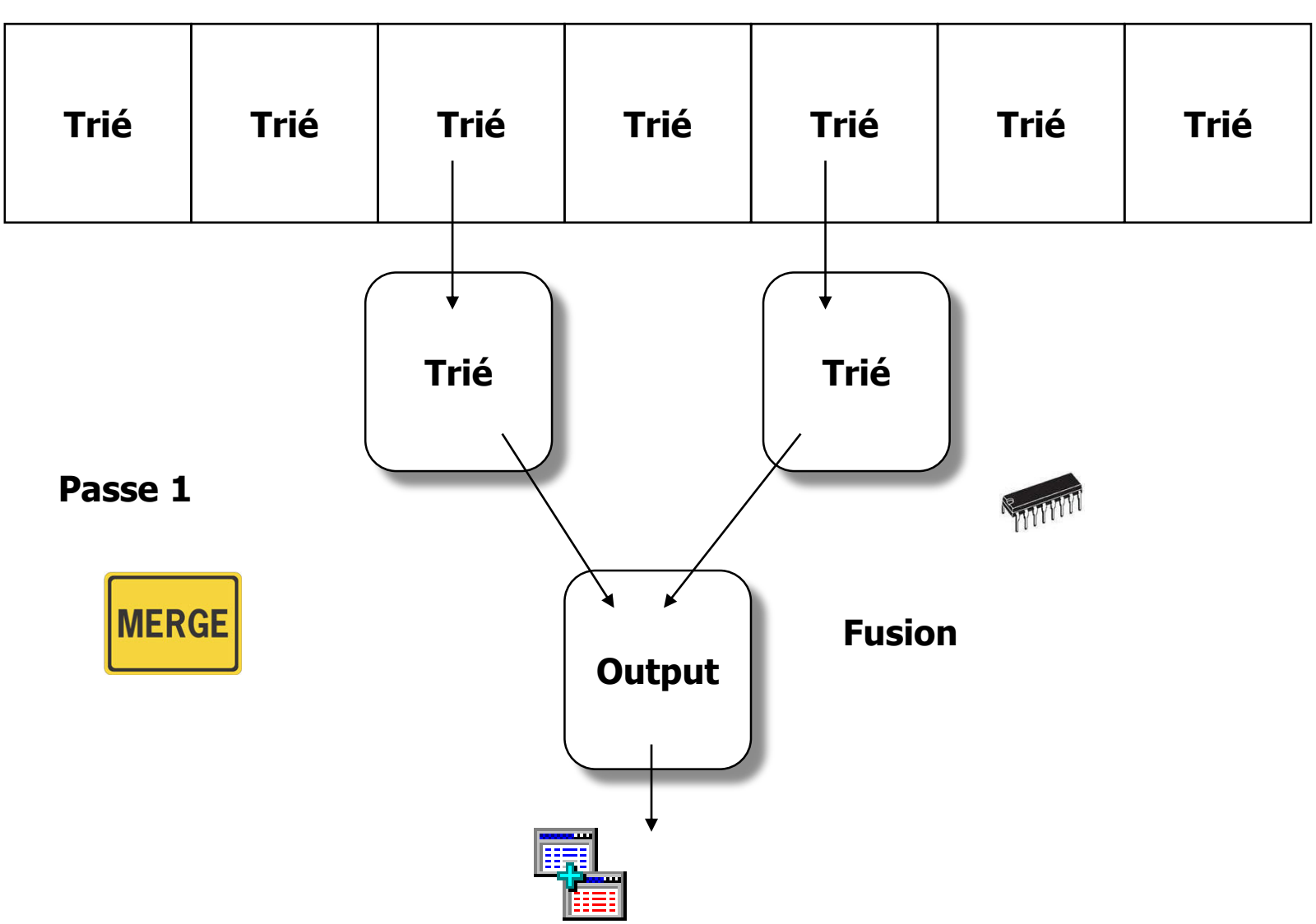
Non Trié

Tri d'un block à la fois
Quick Sort

Trié	Trié	Trié	Trié	Trié	Trié	Trié



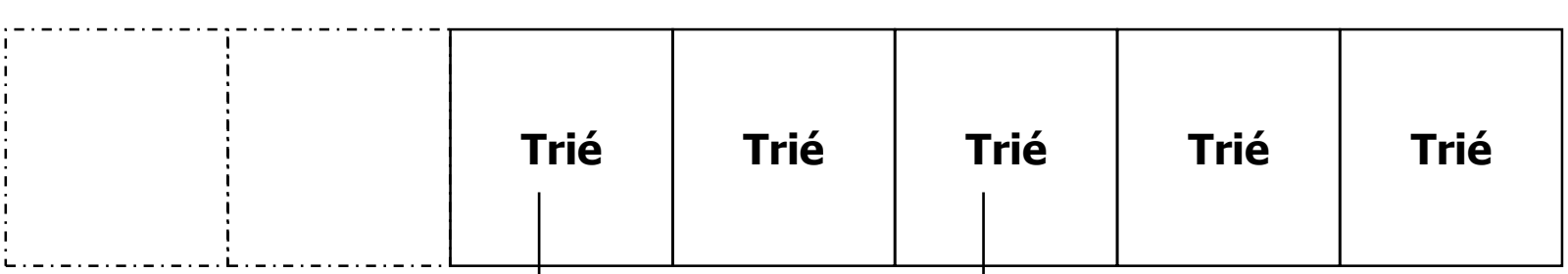


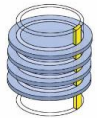


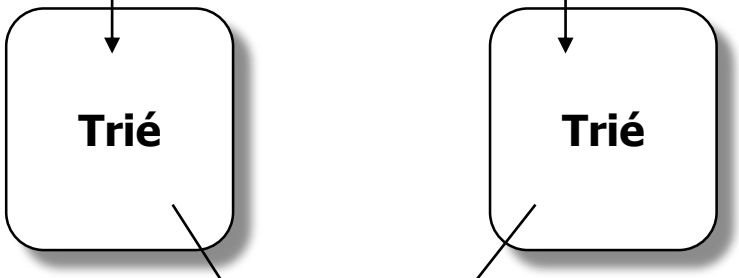
3 pages

Emmanuel Fuchs

Traduction en cours

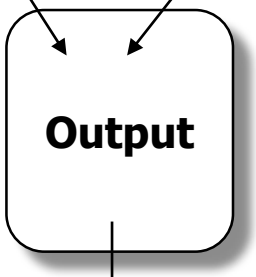


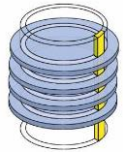


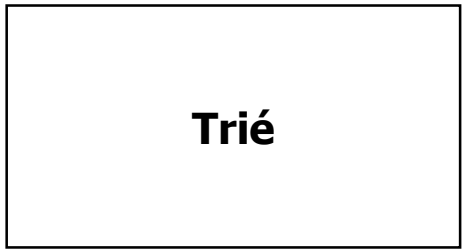


Passe 1

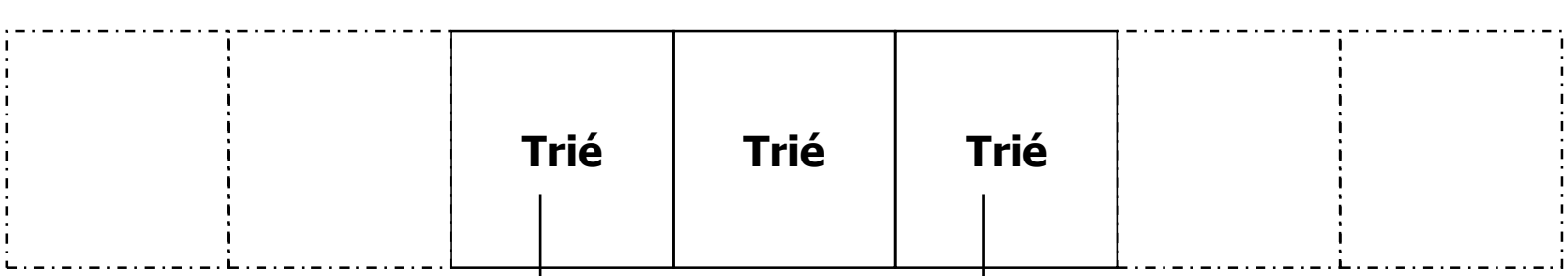
Fusion

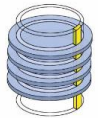


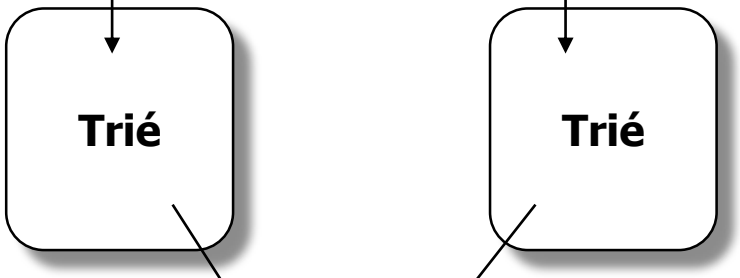




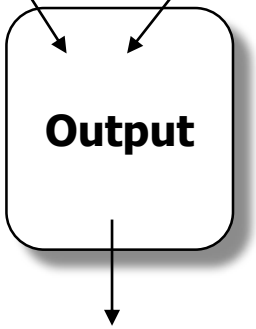
3 buffers



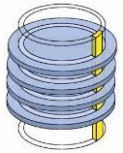


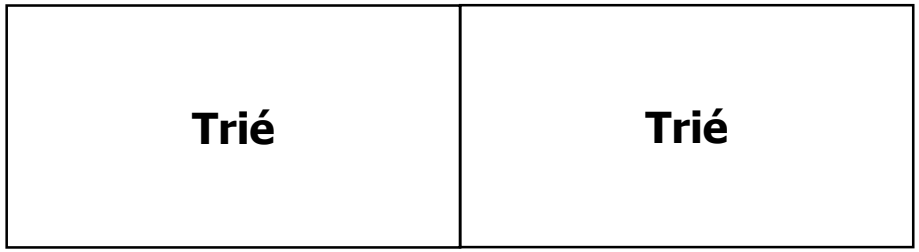


Passe 1

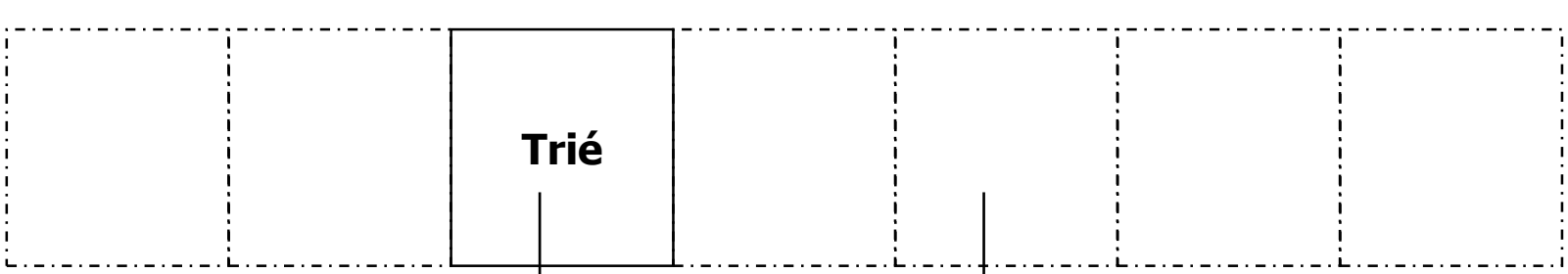


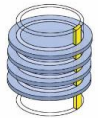
Fusion

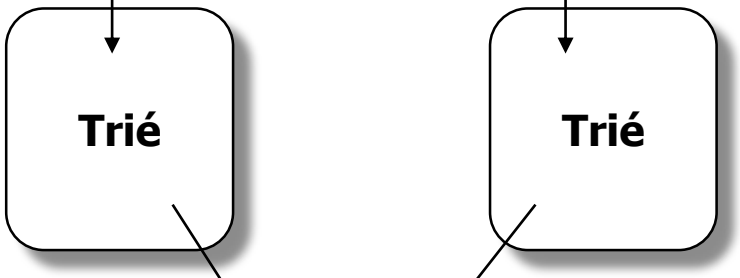




3 buffers

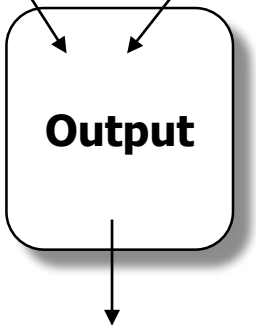


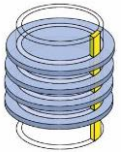


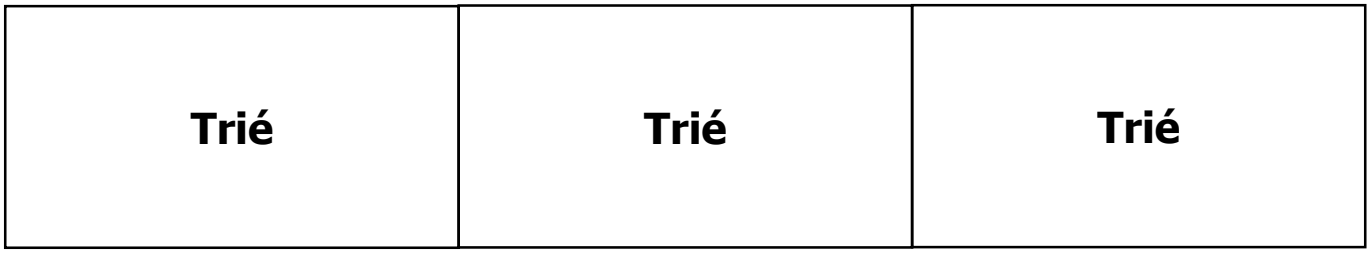


Passe 1

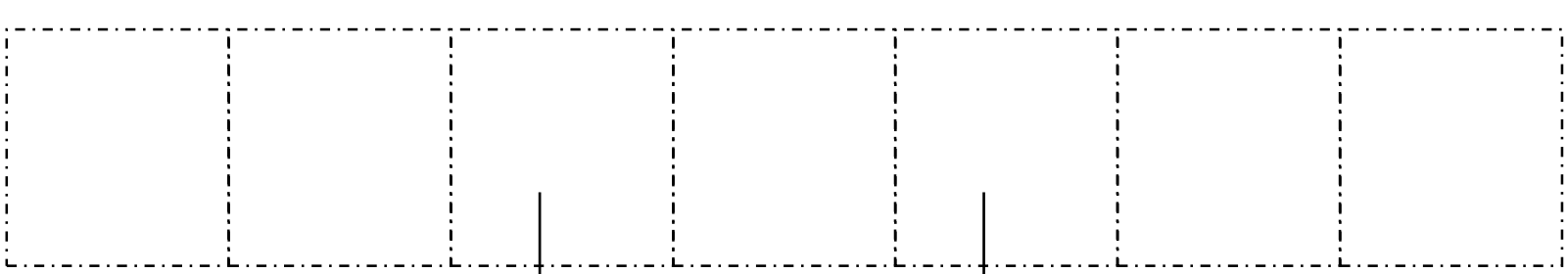
Fusion

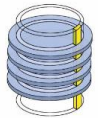






3 buffers





Trié

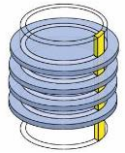
Trié

Passe 1

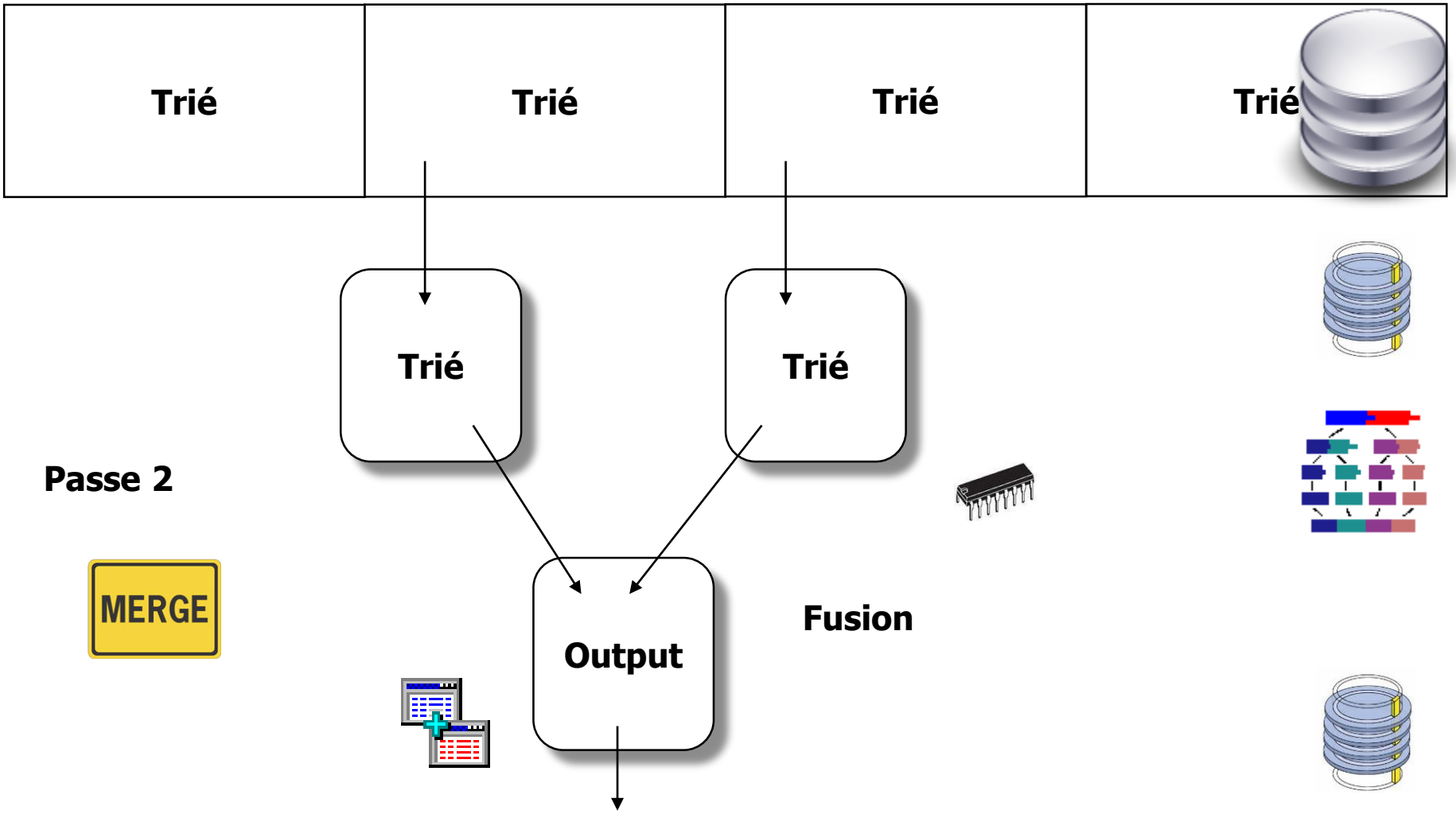
MERGE

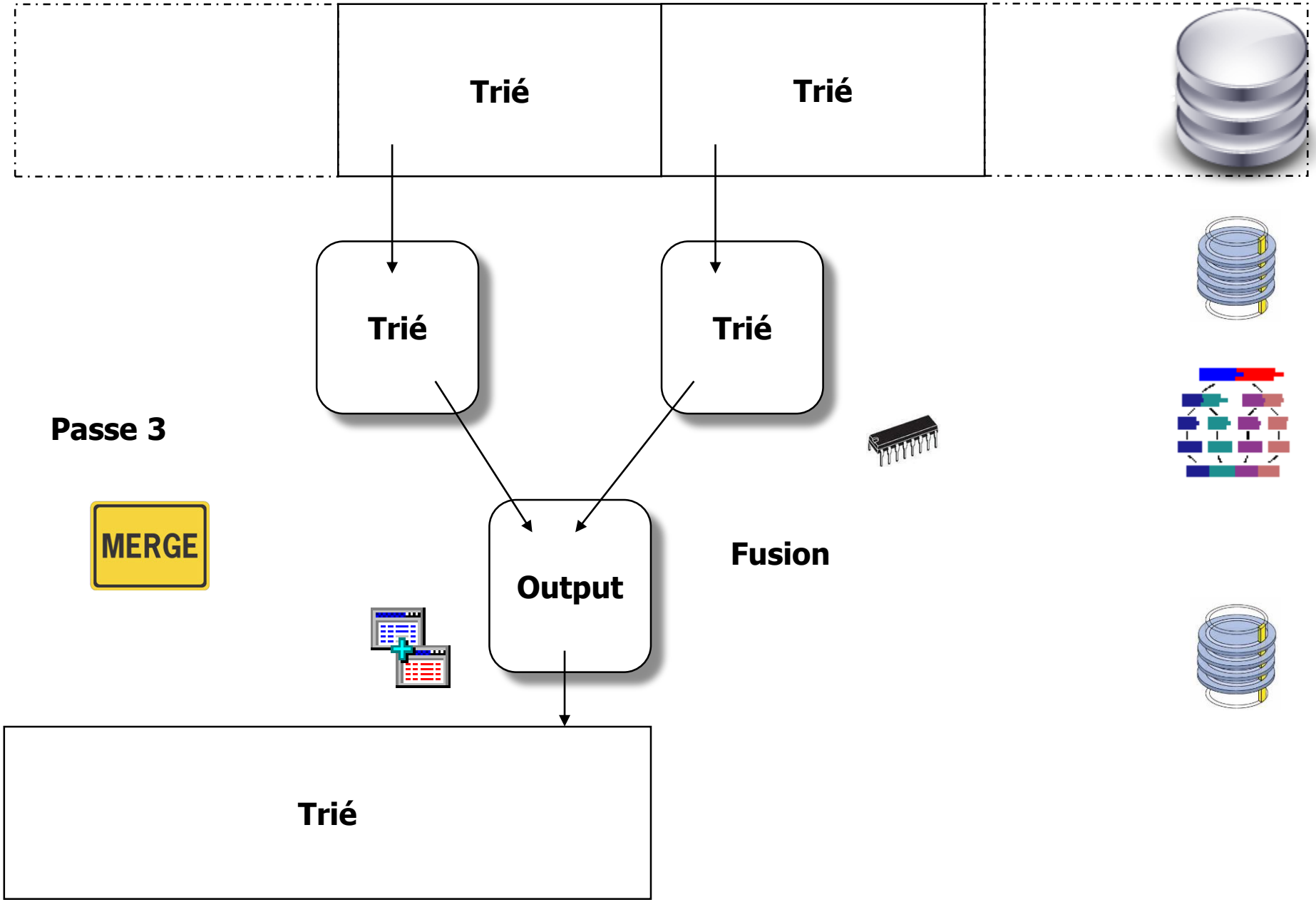
Output

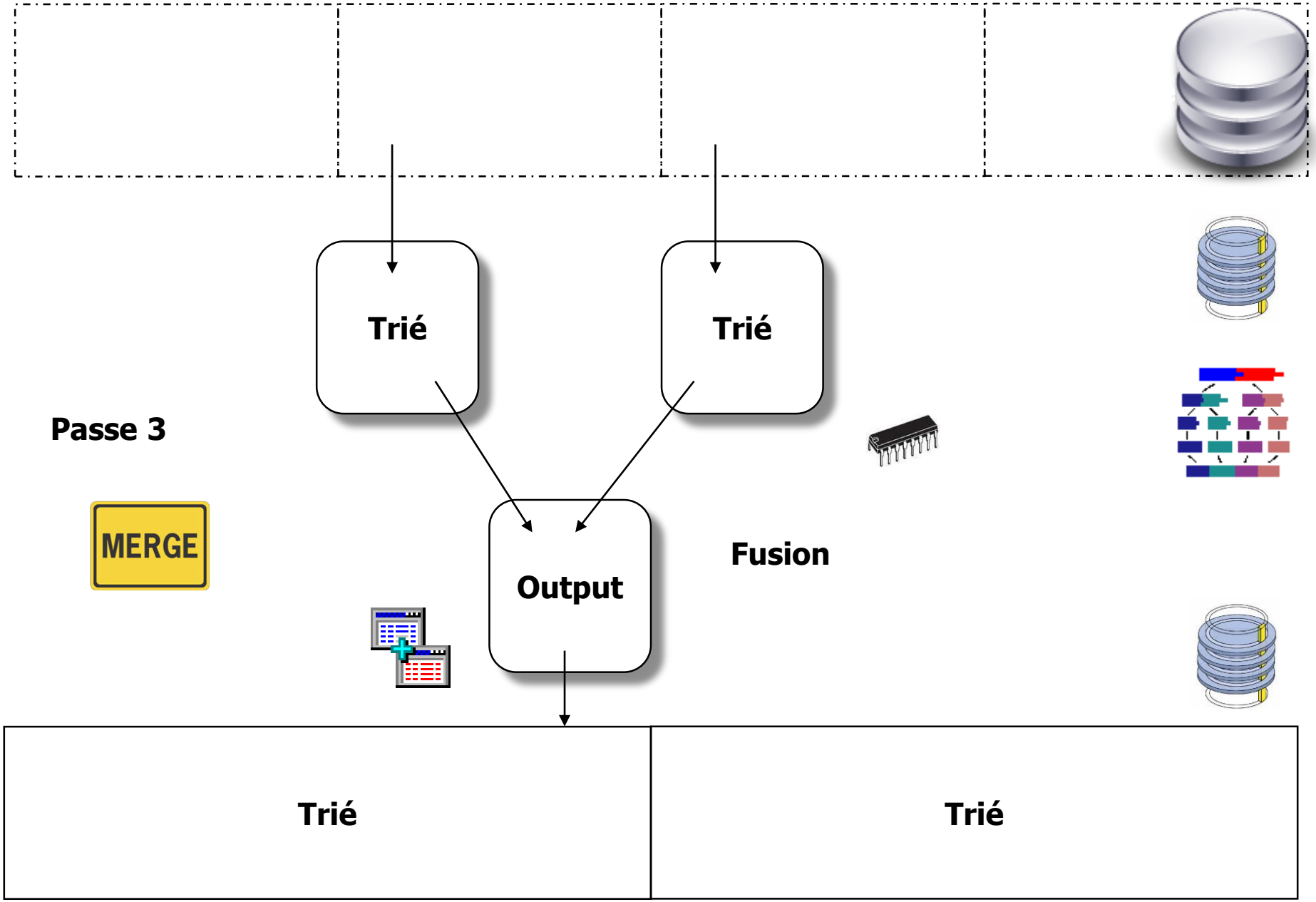
Fusion

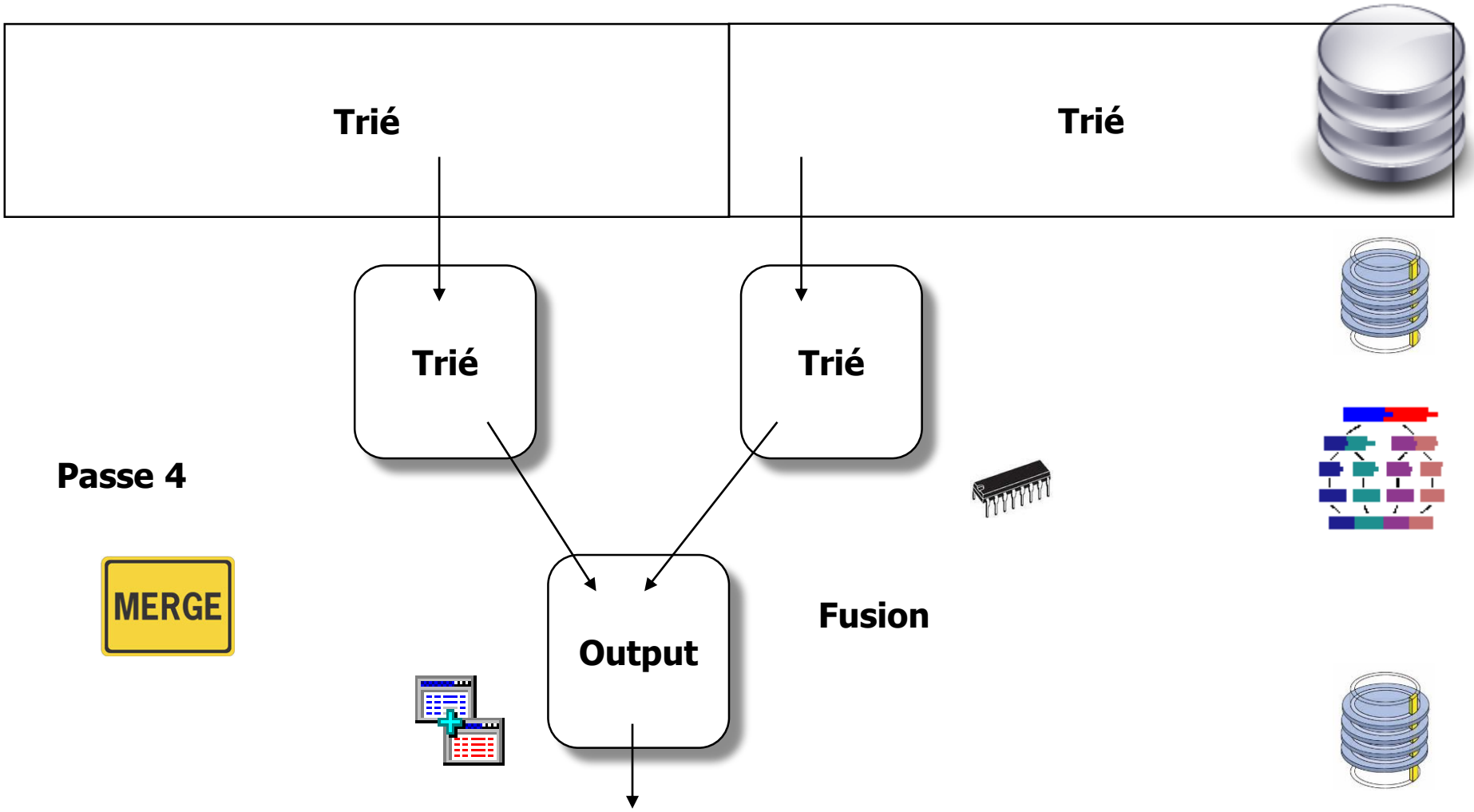


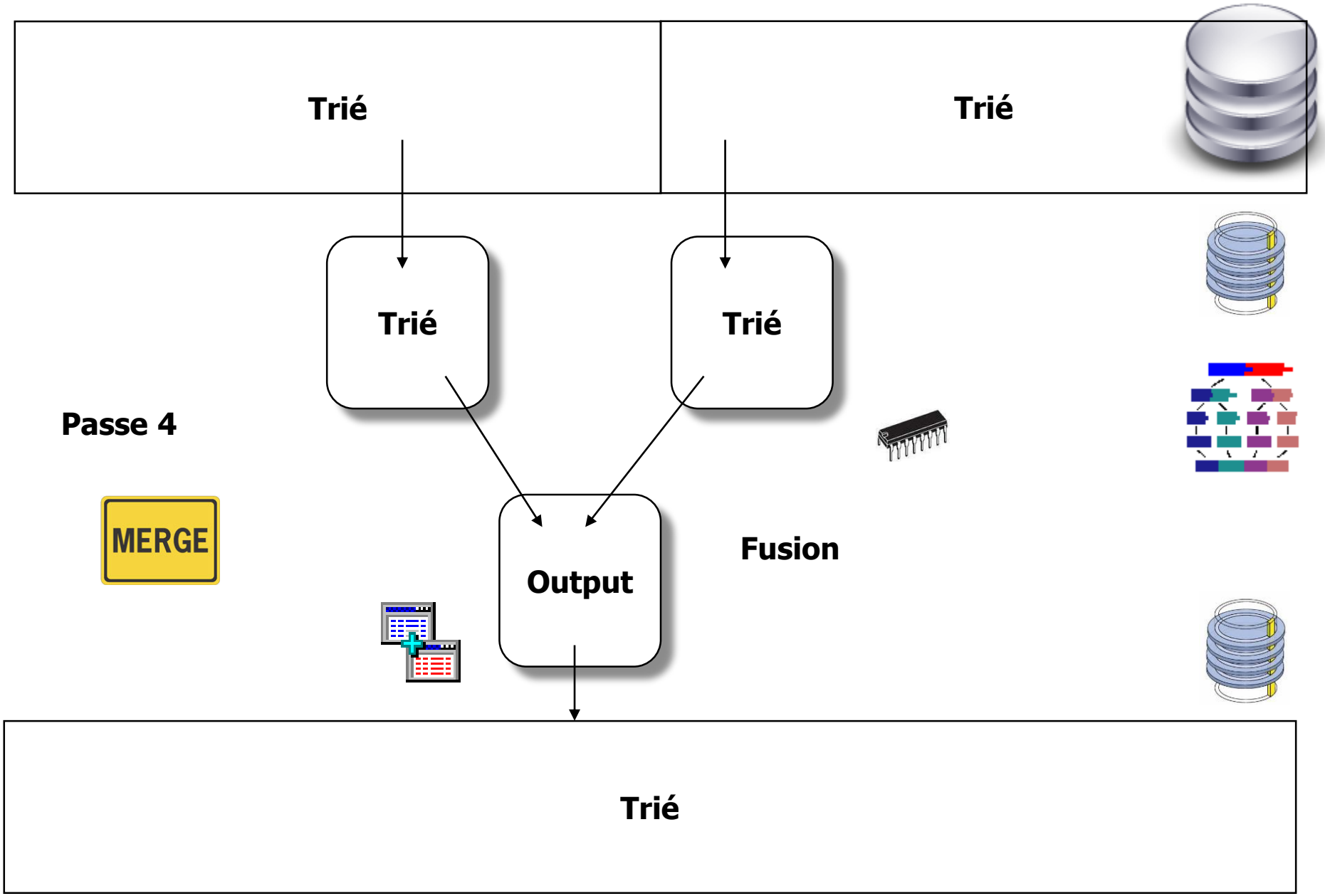
Trié	Trié	Trié	Trié

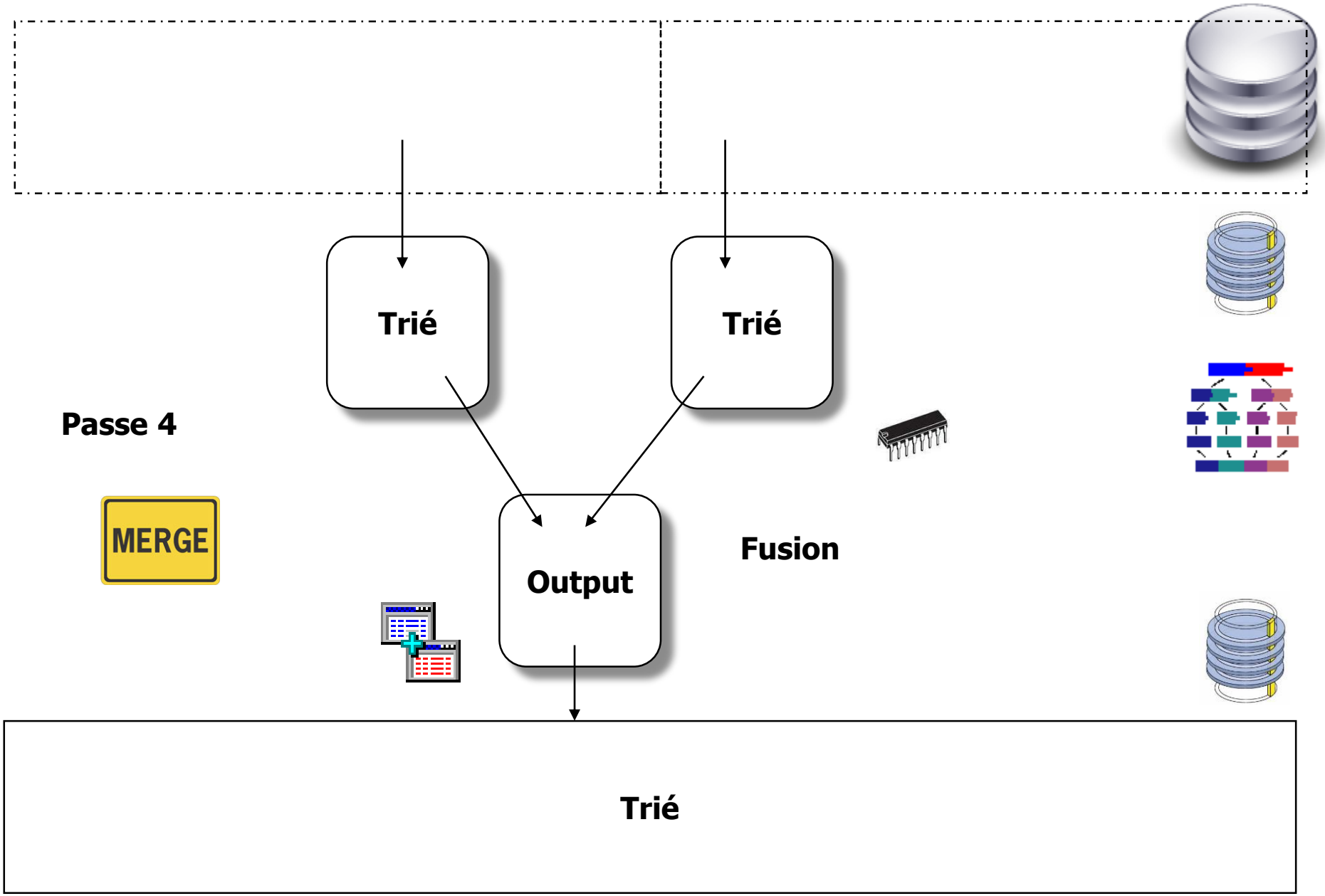

Trié
Trié
Trié
Trié
Trié
Trié
Passe 2
MERGE
Output
Fusion
3 buffers


Trié
Trié
Trié
Trié
Passe 3
MERGE
Output
Fusion
Trié
3 buffers

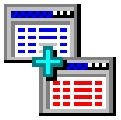


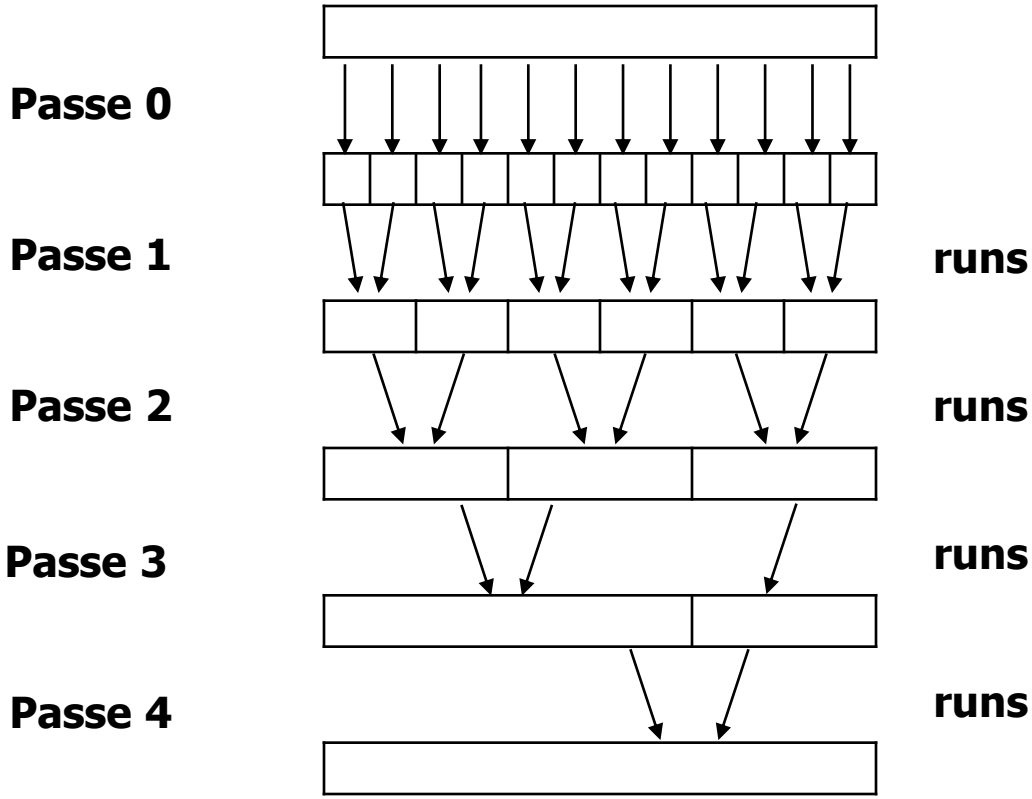

Trié
Trié
Trié
Trié
Passe 4
MERGE
Output
Fusion
3 buffers



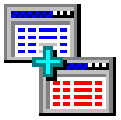


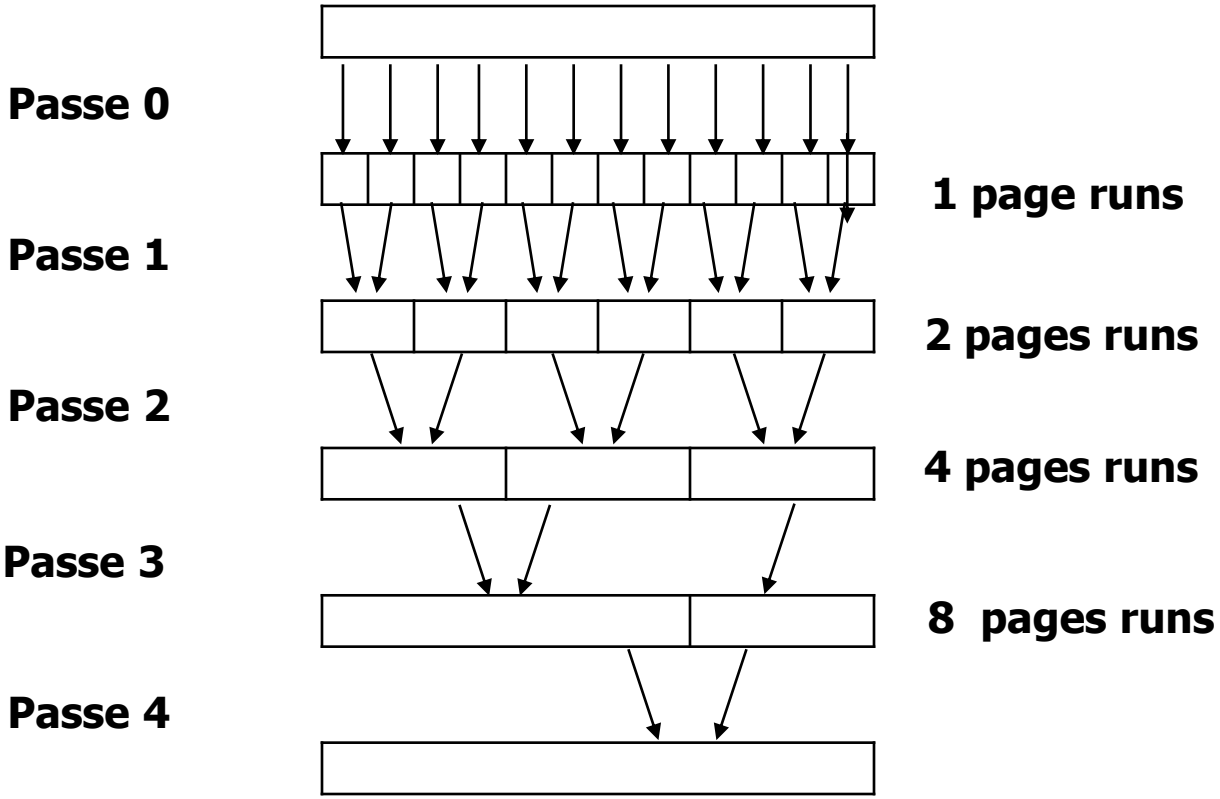
2 way Sort Merge



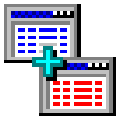


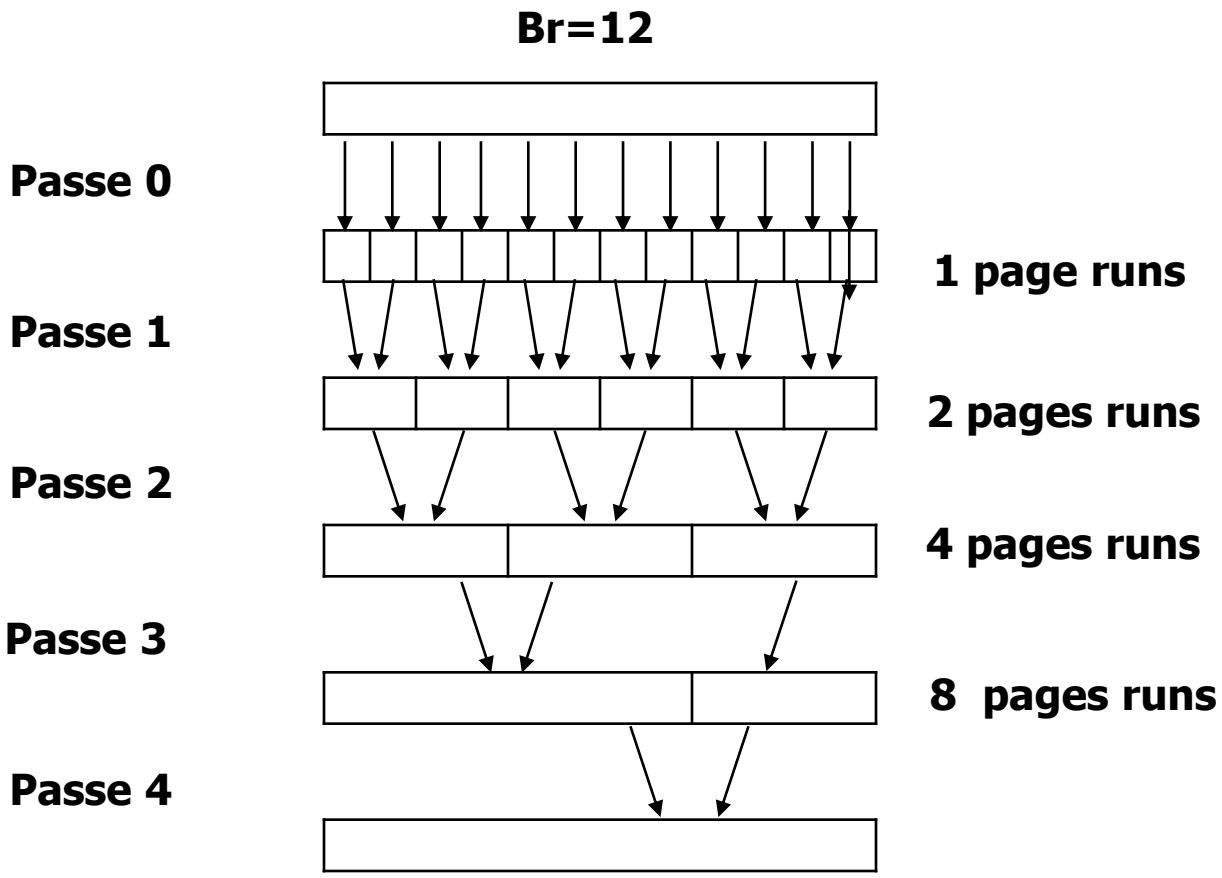
2 way Sort Merge



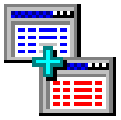


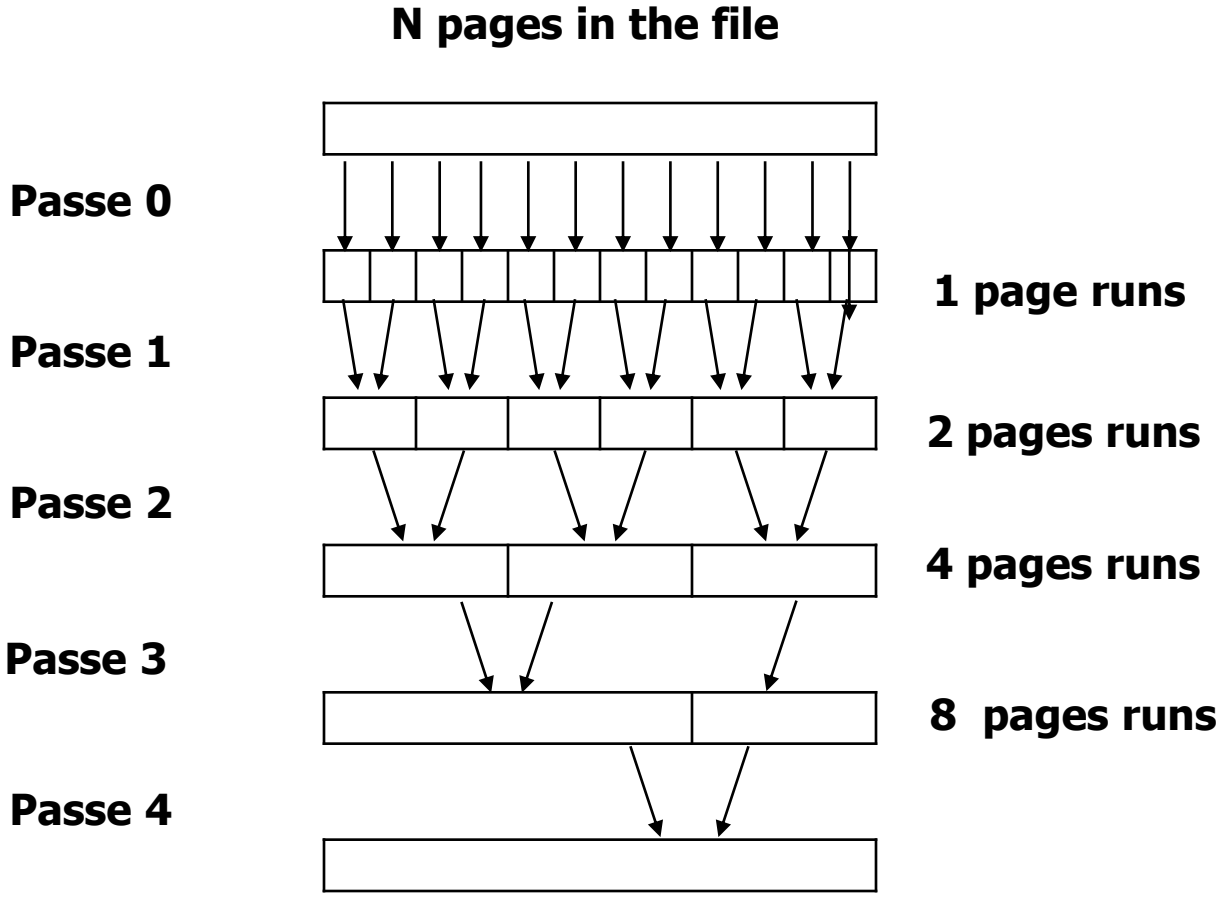
2 way Sort Merge





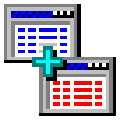
2 way Sort Merge

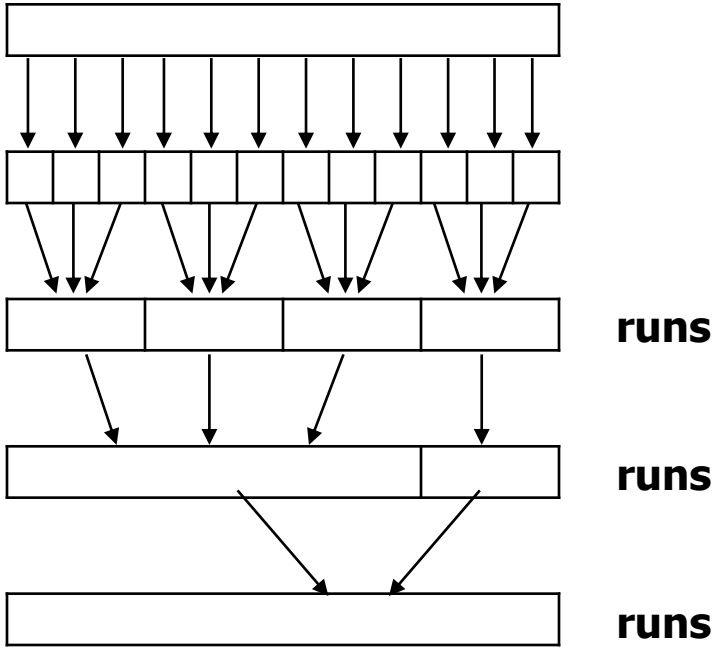




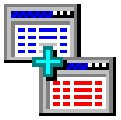
Number of passes : $2\ N\ (\lceil \log_2 N \rceil + 1)$

3 way Sort Merge





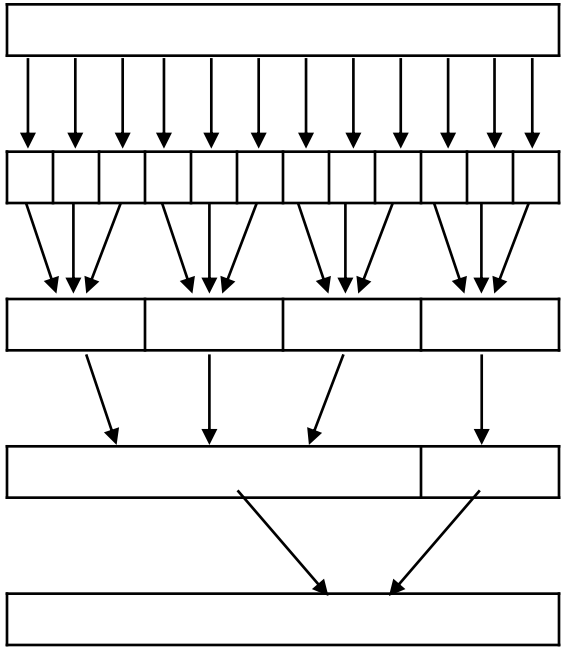
3 way Sort Merge



Passe 0

Passe 1

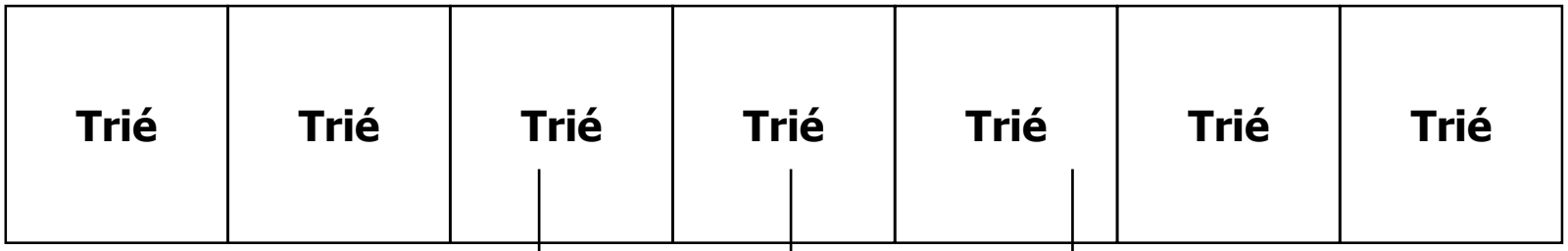
Passe 2

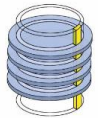


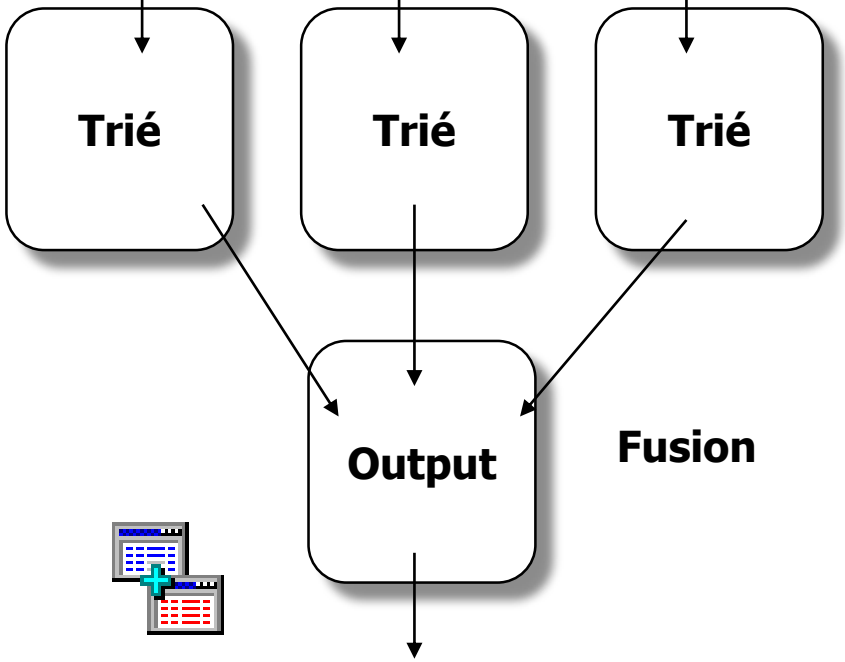
3 pages runs

6 pages runs

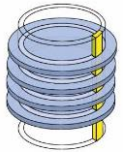
12 pages runs



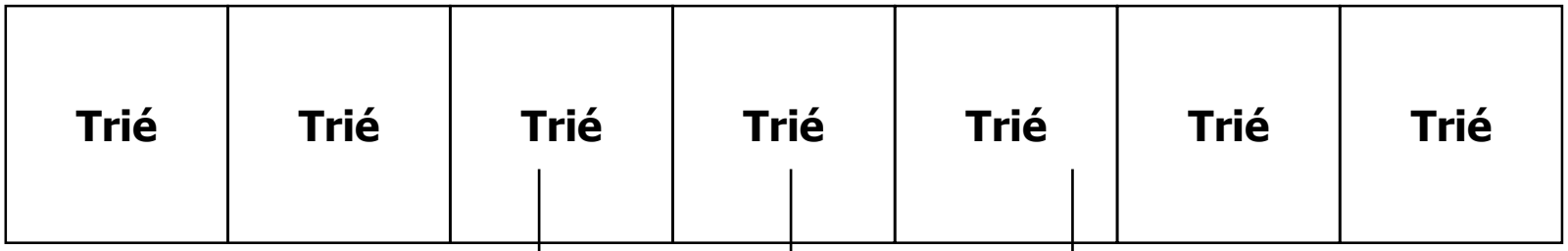


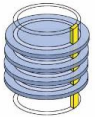


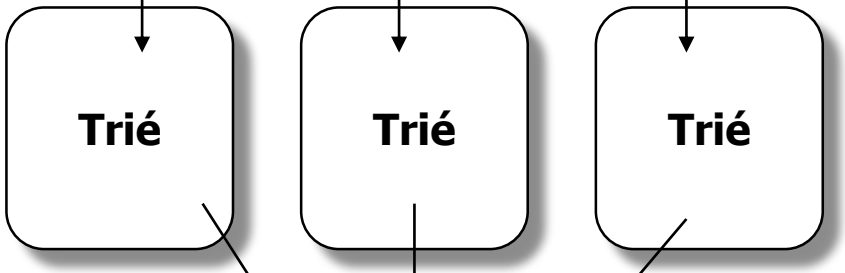
Fusion



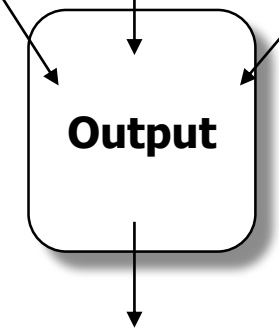
4 buffers



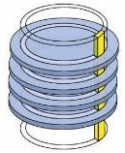


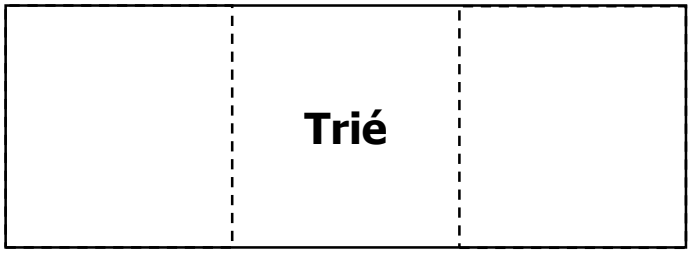


Passe 1

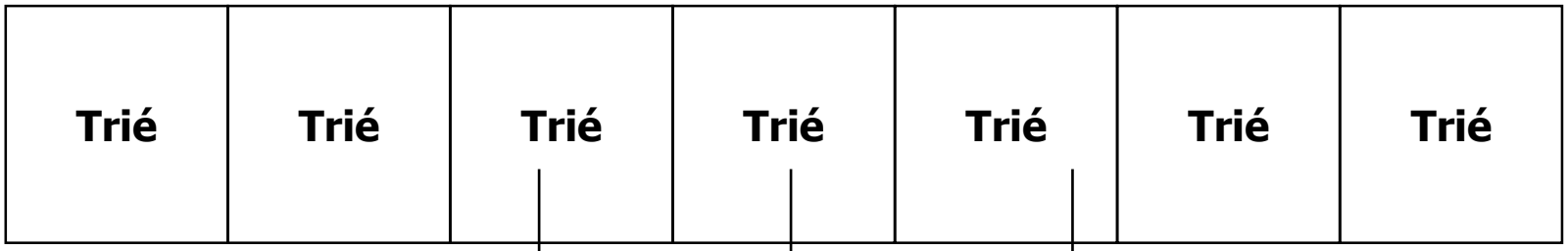


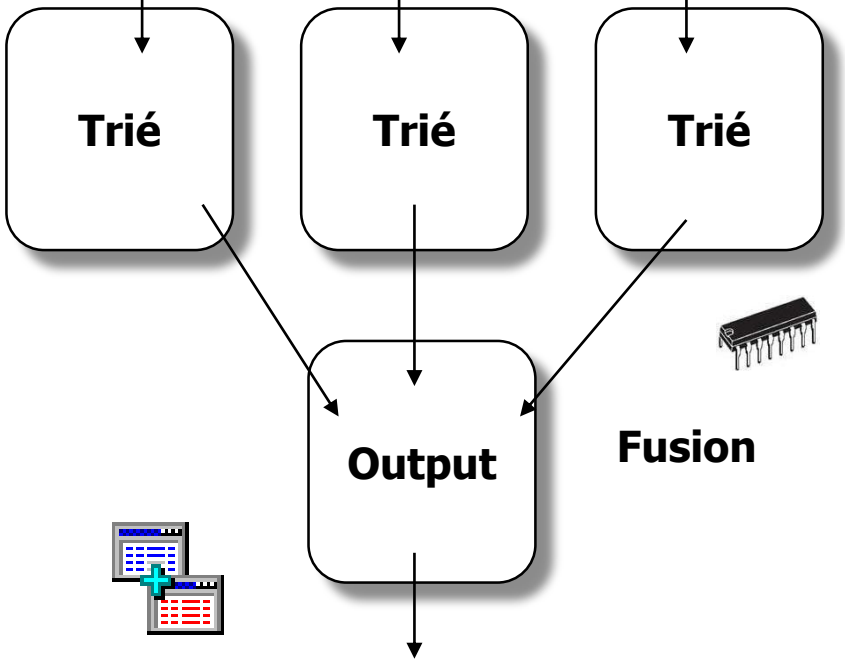
Fusion





3 WAY MERGE

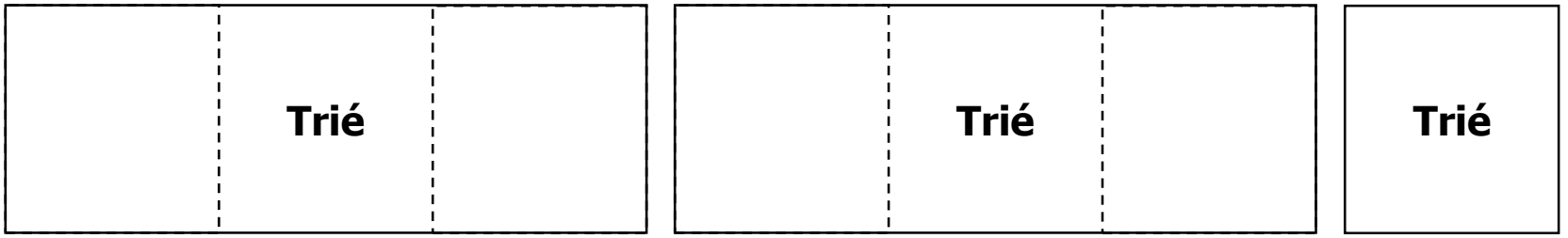


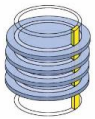


Passe 1

Fusion

3 WAY MERGE





Trié

Trié

Trié

Passe 1

MERGE

Output

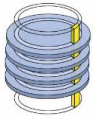
Fusion

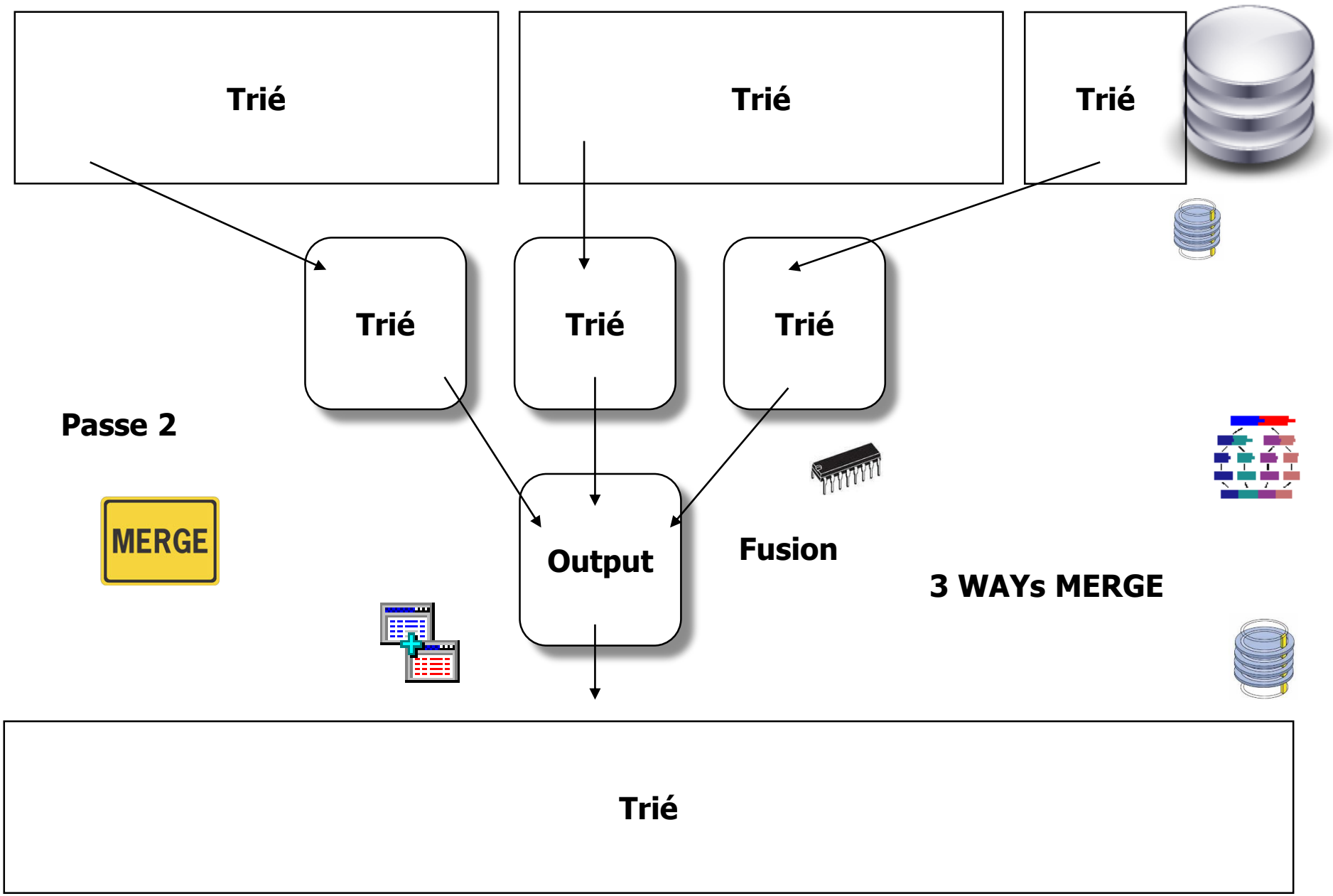
3 WAY MERGE

Trié

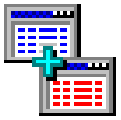
Trié

Trié





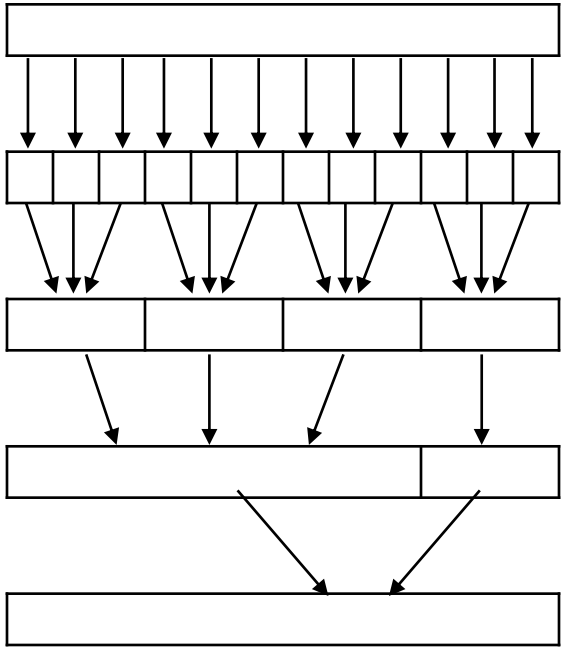
3 way Sort Merge



Passe 0

Passe 1

Passe 2

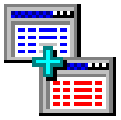


3 pages runs

6 pages runs

12 pages runs

3 way Sort Merge

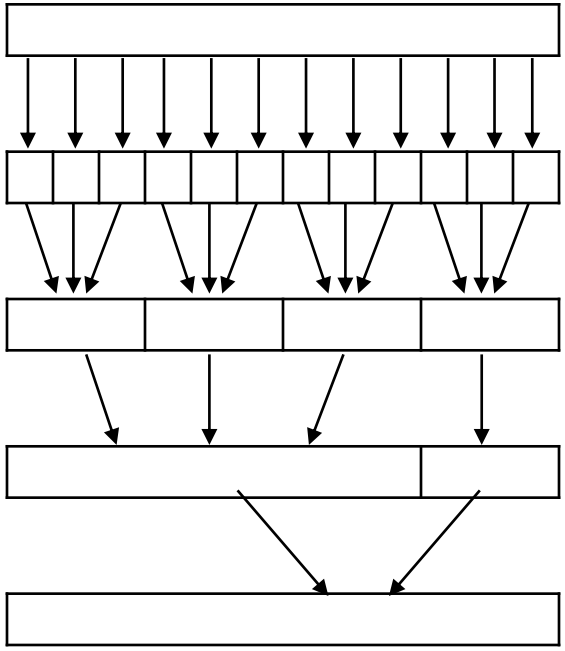


Br=12

Passe 0

Passe 1

Passe 2



3 pages runs : 3 * 4

6 pages runs : 6 * 1 + 4

12 pages runs : 12 * 1

4 buffers