

Thuật toán sắp xếp trộn (Merge Sort)

[Mô phỏng quá trình sắp xếp trộn](#)

- Giả sử ta có hai dãy $A[i], \dots, A[k]$ và $A[k+1], \dots, A[j]$ và hai dãy này đã được sắp.
- Thực hiện trộn hai dãy trên để được dãy $A[i], \dots, A[j]$ cũng được sắp
- Do hai dãy $A[i], \dots, A[k]$ và dãy $A[k+1], \dots, A[j]$ đã được sắp nên việc trộn hai dãy thành một dãy được sắp là rất đơn giản.
- **Vậy trộn như thế nào?**

Ví dụ: Trộn hai dãy sau



A

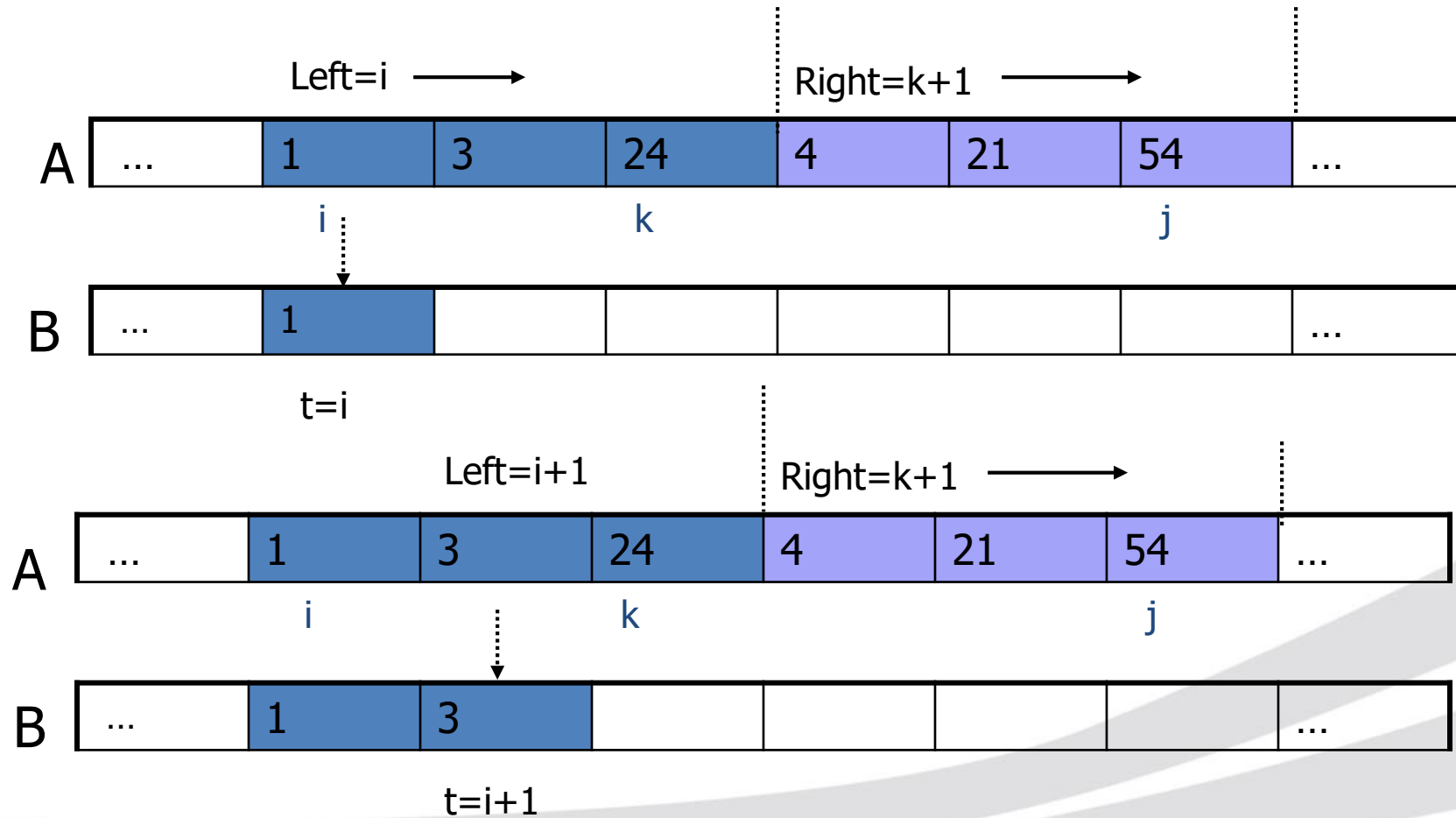
...	1	3	24	4	21	54	...
	i		k	k+1		j	

Thuật toán trộn

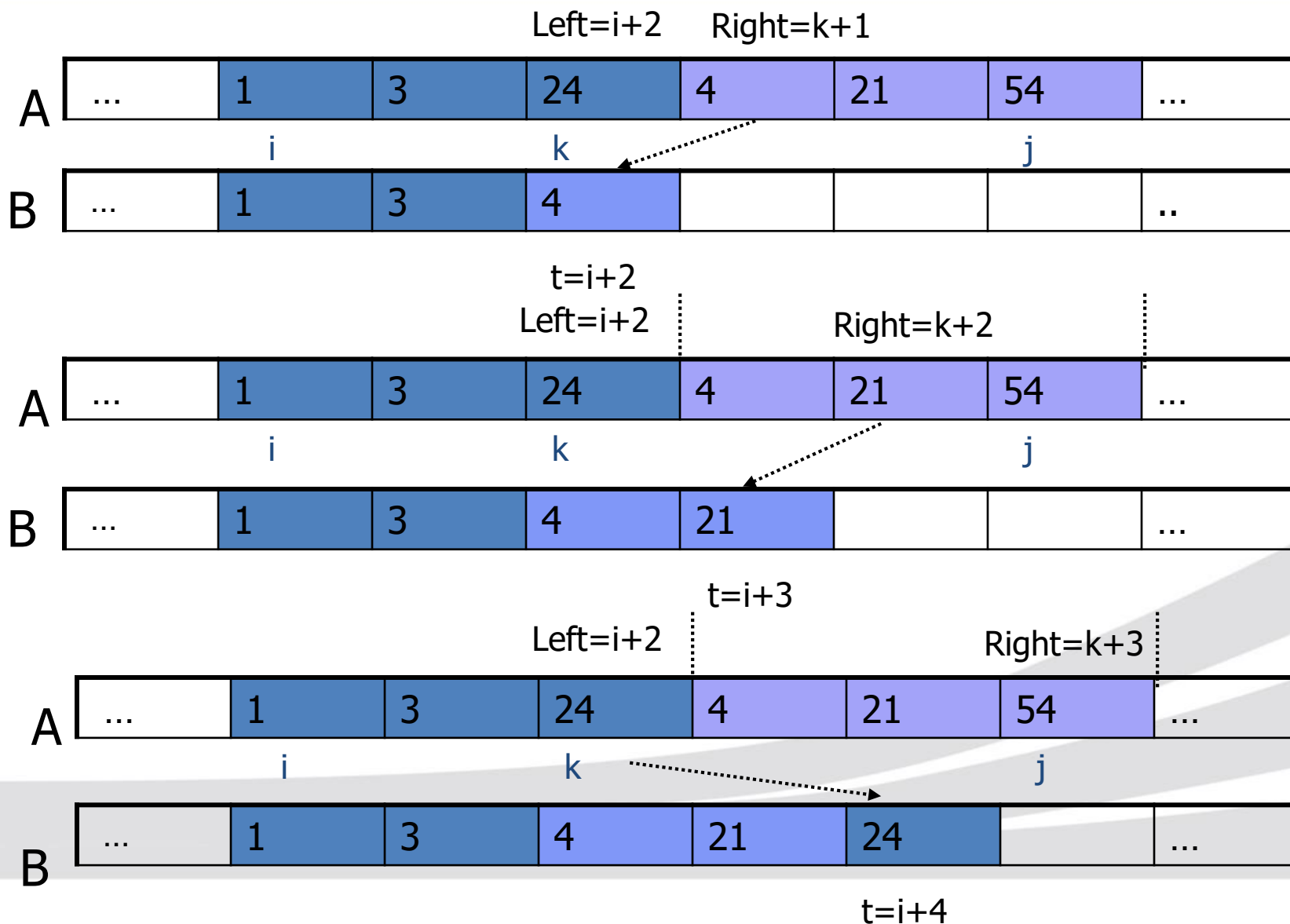


- Sử dụng hai biến $left$, $right$, t và sử dụng mảng phụ $B[i], \dots, B[j]$. $left$ xuất phát từ i , $right$ xuất phát từ $k+1$, t xuất phát từ i trên mảng phụ B .
- Nếu $A[left].key < A[right].key$ thì $B[t] \leftarrow A[left]$, $t \leftarrow t+1$ và $left \leftarrow left+1$
- Nếu $A[left].key \geq A[right].key$ thì $B[t] \leftarrow A[right]$, $t \leftarrow t+1$ và $right \leftarrow right+1$
- Quá trình trên được thực hiện cho đến khi $left > k$ hoặc $right > j$ thì dừng lại.
- Nếu $left > k$ thì $B[t] \leftarrow A[right], \dots, B[j] \leftarrow A[j]$.
- Nếu $right > j$ thì $B[t] \leftarrow A[left]$, $B[t+1] \leftarrow A[left+1], \dots, B[t+k-left] \leftarrow A[k]$.
- Gán $A[i] \leftarrow B[i], \dots, A[j] \leftarrow B[j]$

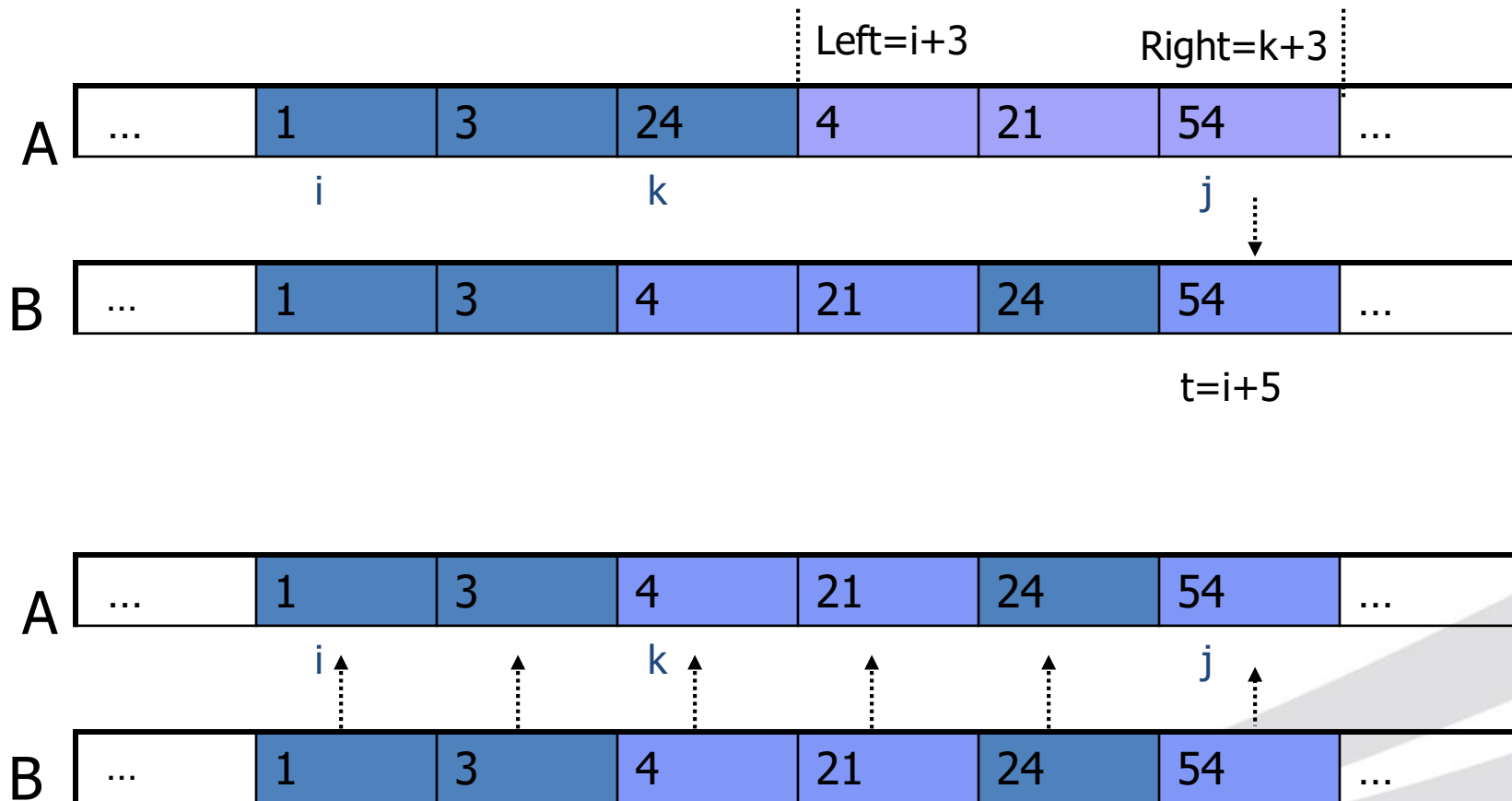
Ví dụ mô tả Quá trình trộn dãy



Ví dụ mô tả Quá trình trộn dãy



Ví dụ mô tả Quá trình trộn dãy



Thuật toán giả mã



Algorithm *Merge(array A, int i, int k, int j)*

Input: Hai dãy $A[i], \dots, A[k]$ và $A[k+1], \dots, A[j]$ đã được sắp và các số nguyên i, j

Output: Dãy $A[i], \dots, A[j]$ cũng được sắp

$\text{left} \leftarrow i; \text{right} \leftarrow k+1; t \leftarrow i;$

While $(\text{left} \leq k)$ and $(\text{right} \leq j)$ **do**

if $A[\text{left}].\text{key} < A[\text{right}].\text{key}$ **then**

$B[t] \leftarrow A[\text{left}];$

$\text{left} \leftarrow \text{left} + 1;$

$t \leftarrow t + 1;$

else

$B[t] \leftarrow A[\text{right}];$

$\text{right} \leftarrow \text{right} + 1;$

$t \leftarrow t + 1; \quad // \text{kết thúc while}$

If $\text{left} > k$ **then**

for $r \leftarrow \text{right}$ **to** j **do**

$B[t] \leftarrow A[r];$

$t++;$

else

for $r \leftarrow \text{left}$ **to** k **do**

$B[t] \leftarrow A[r];$

$t++;$

for $r \leftarrow i$ **to** j **do**

$A[r] \leftarrow B[r];$

Thuật toán sắp xếp trộn



- Để sắp xếp dãy $A[1], \dots, A[n]$ ta thực hiện như sau:
- Chia dãy trên thành hai dãy: $A[1], \dots, A[k]$ và dãy $A[k+1], \dots, A[n]$, trong đó $k = (n+1)/2$
- Thực hiện sắp xếp 2 dãy $A[1], \dots, A[k]$ và $A[k+1], \dots, A[n]$ độc lập cũng theo thuật toán Mergesort.
- Thực hiện trộn hai dãy: $A[1], \dots, A[k]$ và dãy $A[k+1], \dots, A[n]$ để được dãy $A[1], \dots, A[n]$ cũng được sắp

Thuật toán giả mã



Algorithm *Mergesort*(array *A*, int *i*, int *j*)

Input: **Dãy các phần tử $A[i], \dots, A[j]$**

Output: **Dãy $A[i], \dots, A[j]$ được sắp.**

if $i < j$ then

$k \leftarrow (i+j)/2$;

Mergesort(*A*, *i*, *k*);

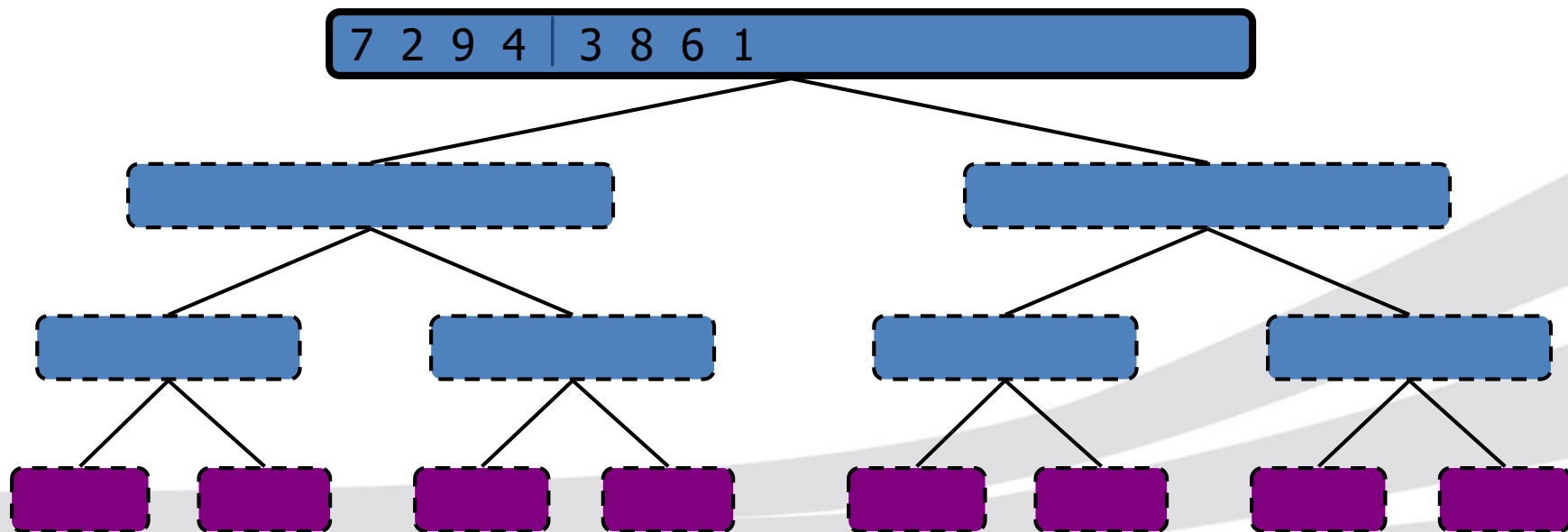
Mergesort(*A*, *k*+1, *j*);

Merge(*A*, *i*, *k*, *j*);

Mô tả quá trình thực hiện sắp xếp



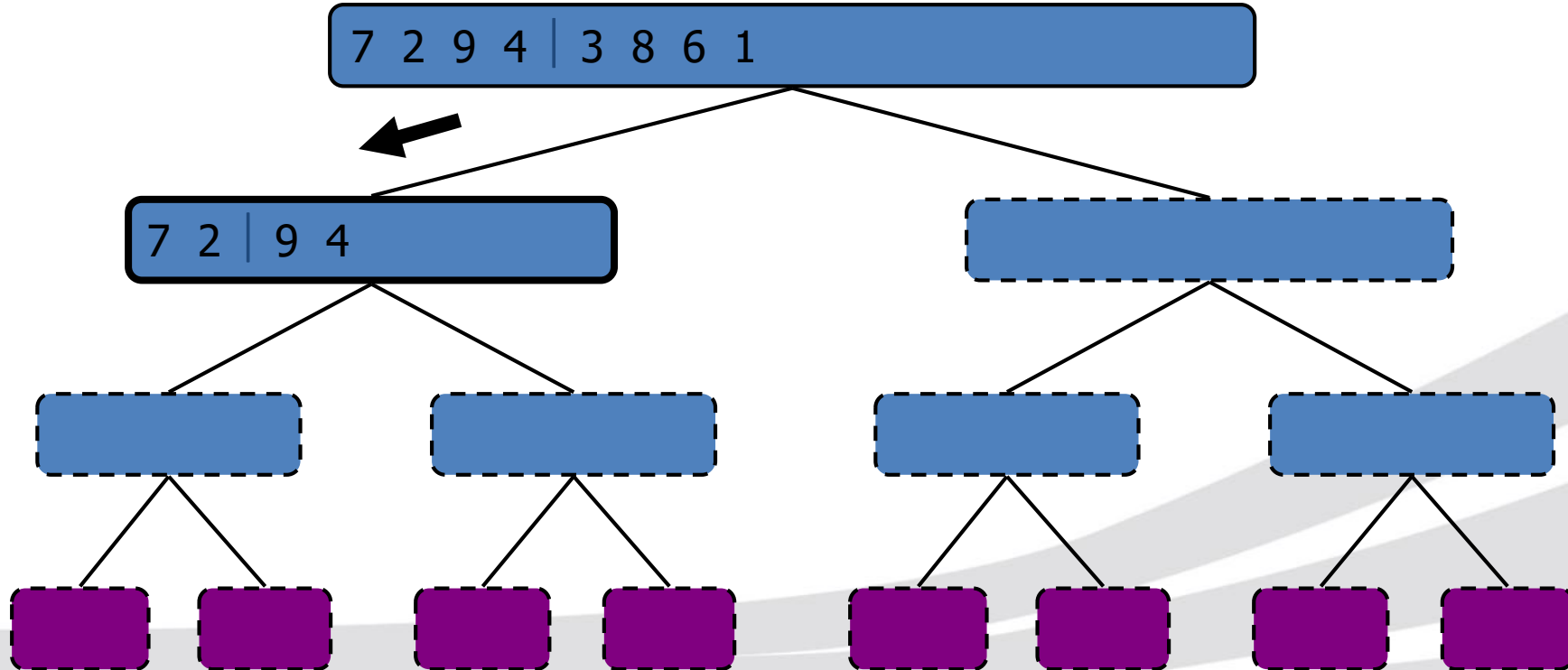
- ❖ Ví dụ sắp xếp dãy: $A = 7\ 2\ 9\ 4\ 3\ 8\ 6\ 1$
- Gọi thủ tục MergeSort($A, 1, 8$), chia đôi dãy



Mô tả quá trình thực hiện sắp xếp



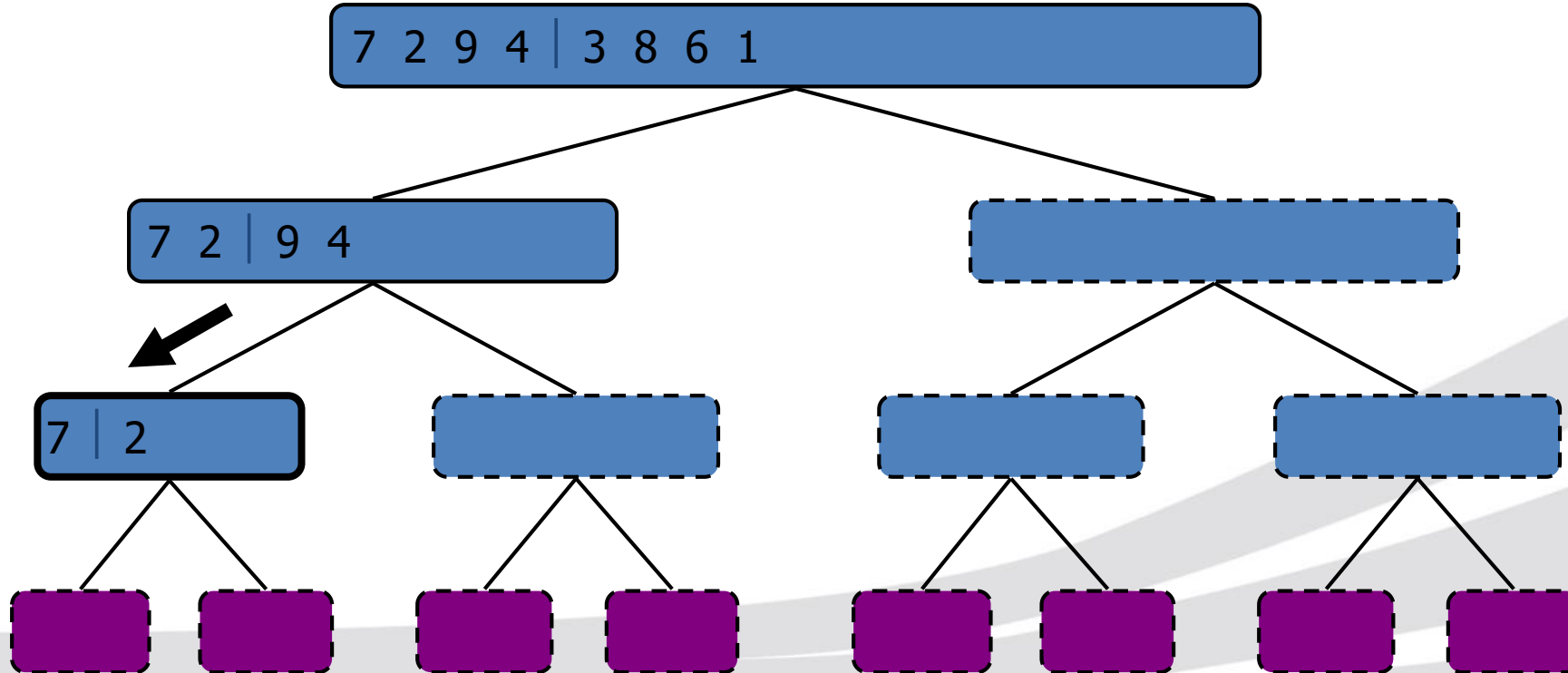
❖ Gọi đệ qui và phân chia Mergesort(A,1,4)



Mô tả quá trình thực hiện sắp xếp



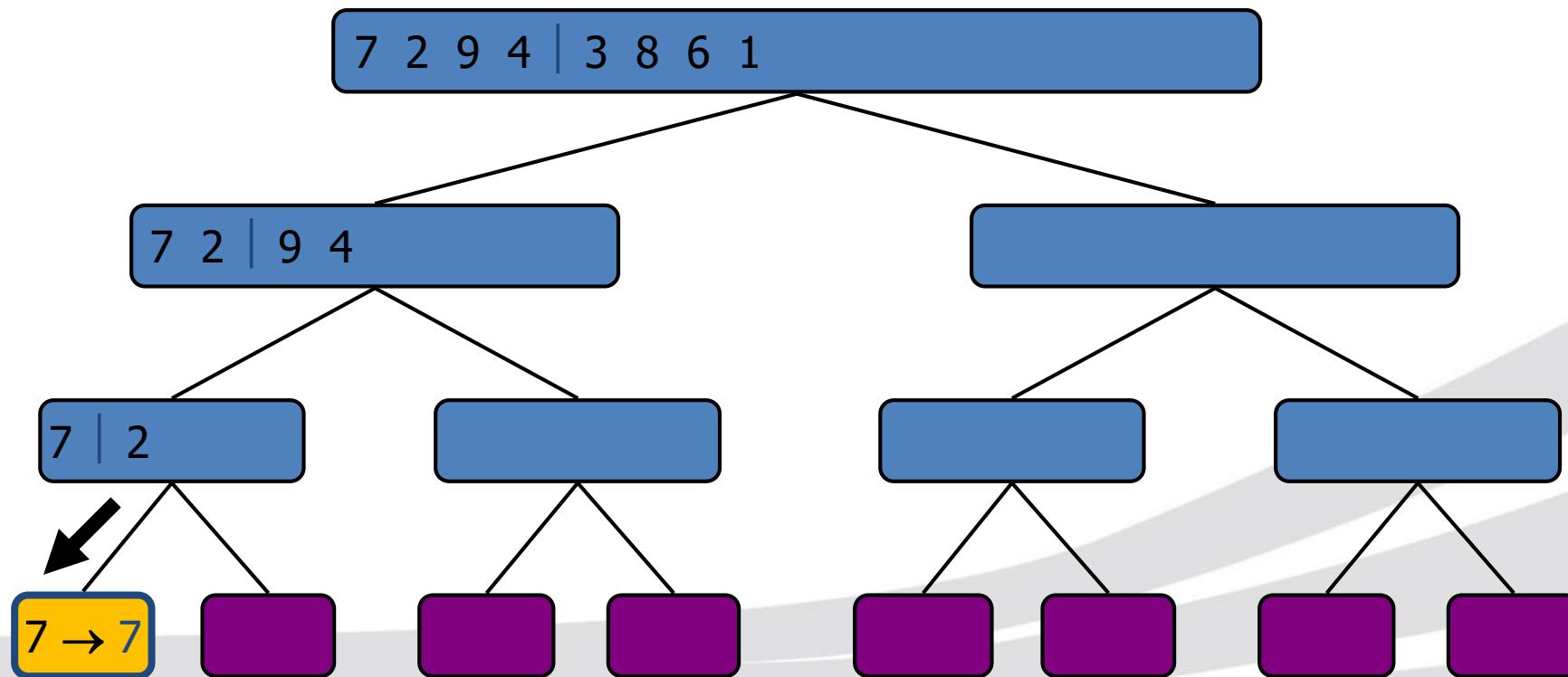
❖ Gọi đệ qui và phân chia Mergesort(A,1,2)



Mô tả quá trình thực hiện sắp xếp



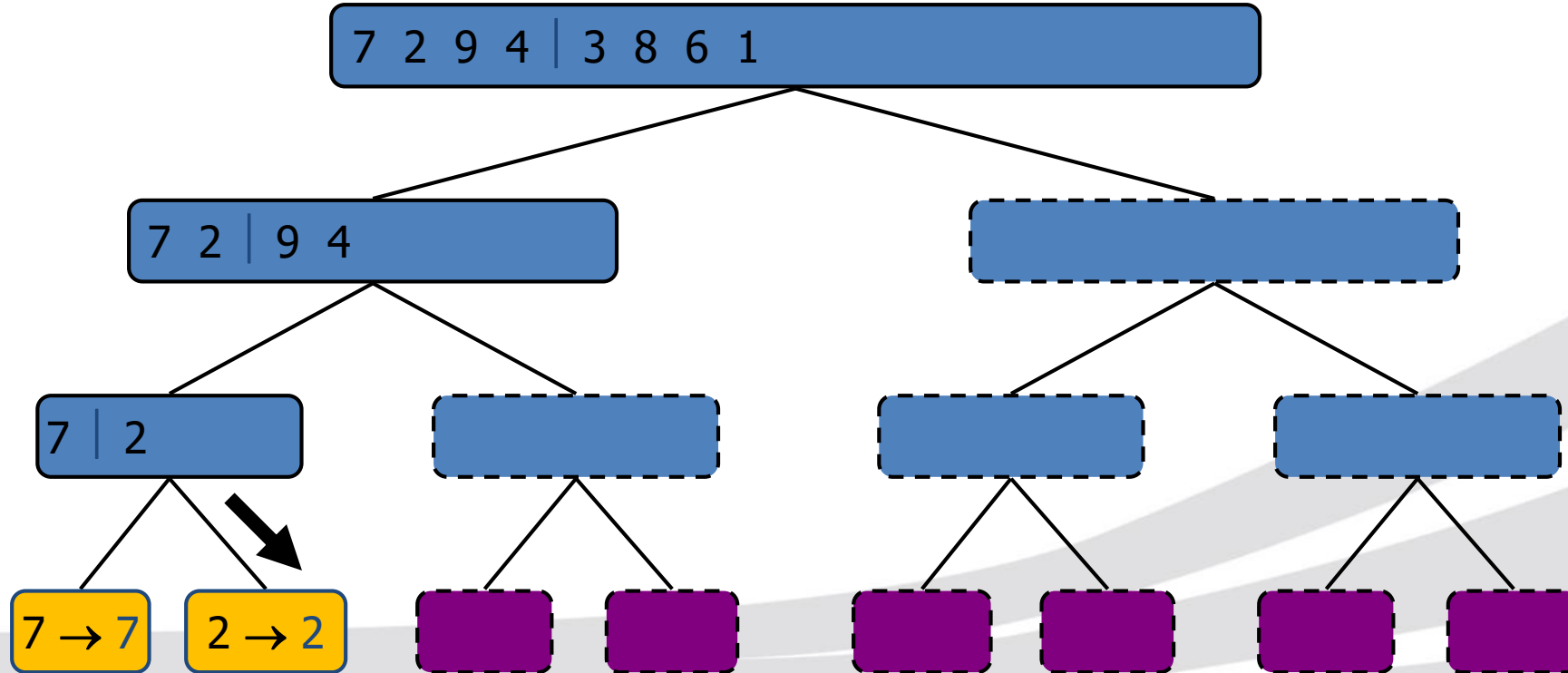
❖ Gọi đệ qui Mergesort(A,1,1), đây là trường hợp cơ sở



Mô tả quá trình thực hiện sắp xếp



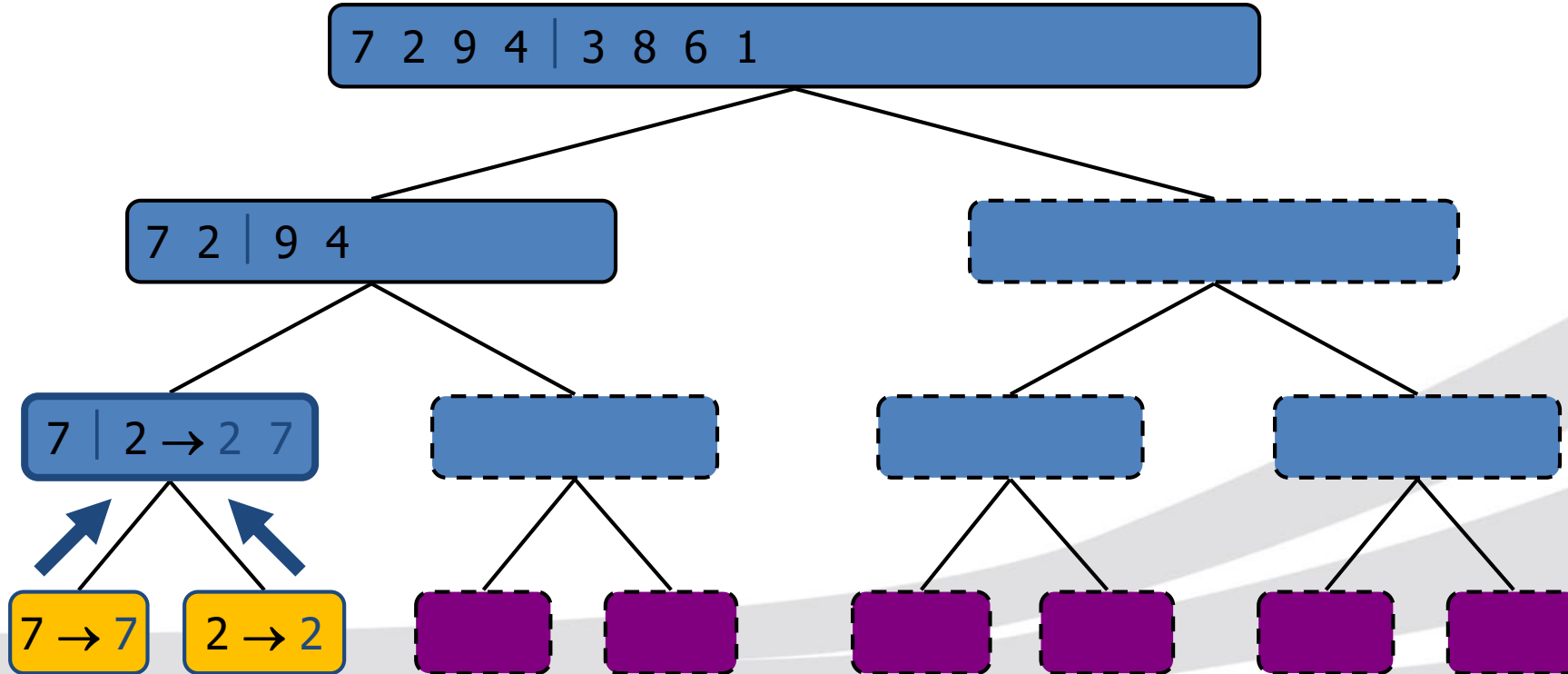
❖ Gọi đệ qui Mergesort(A,2,2), đây là trường hợp cơ sở



Mô tả quá trình thực hiện sắp xếp



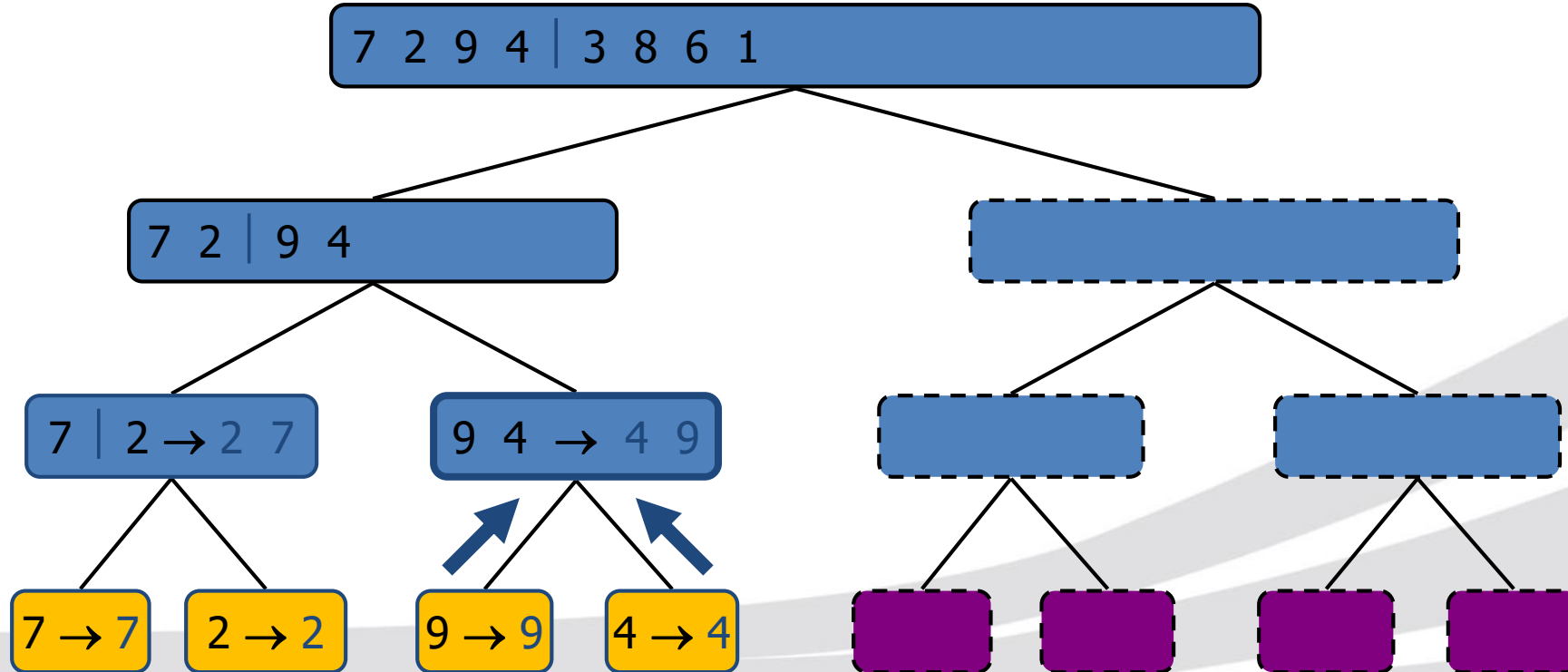
❖ Trộn merge(A,1,1,2)



Mô tả quá trình thực hiện sắp xếp



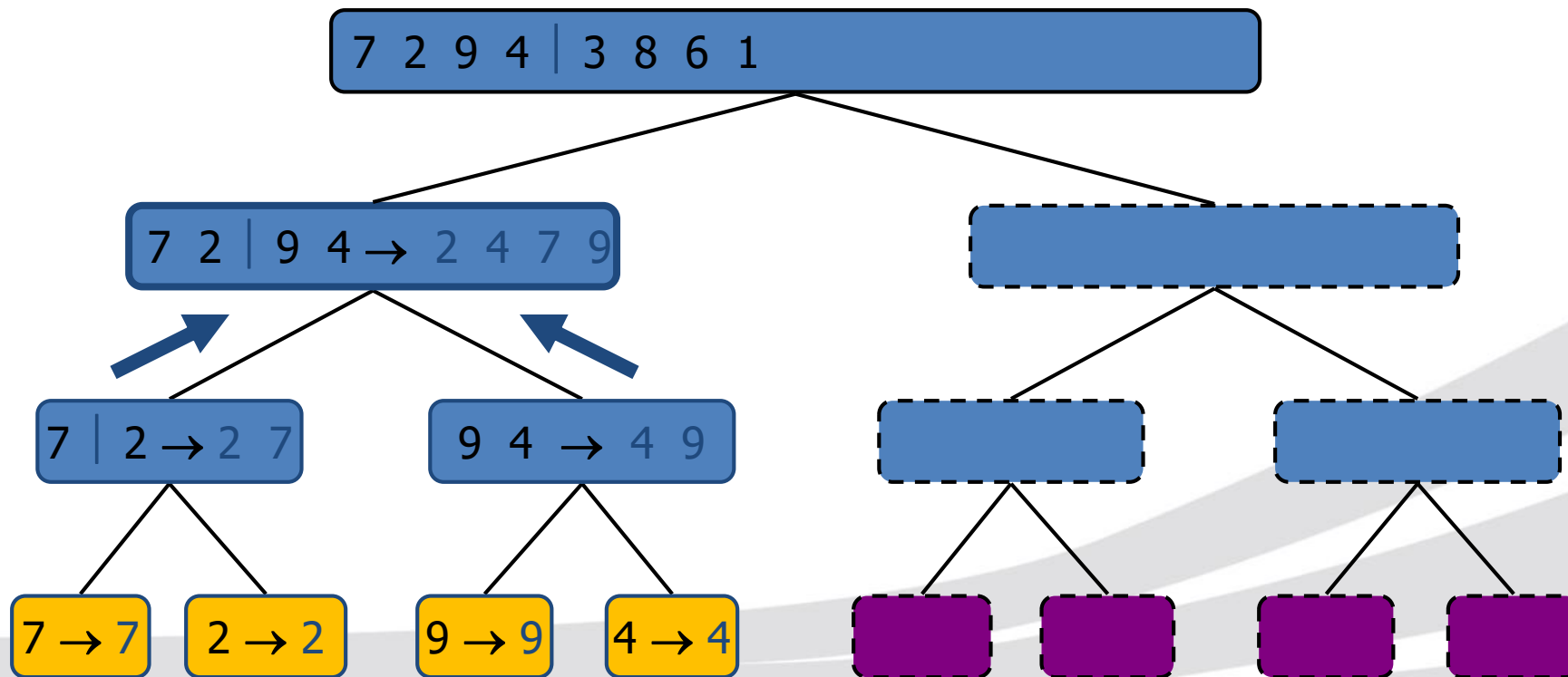
❖ Gọi đệ quy Mergesort(A,3,3), Mergesort(A,4,4) và trộn merge(A,3,3,4)



Mô tả quá trình thực hiện sắp xếp



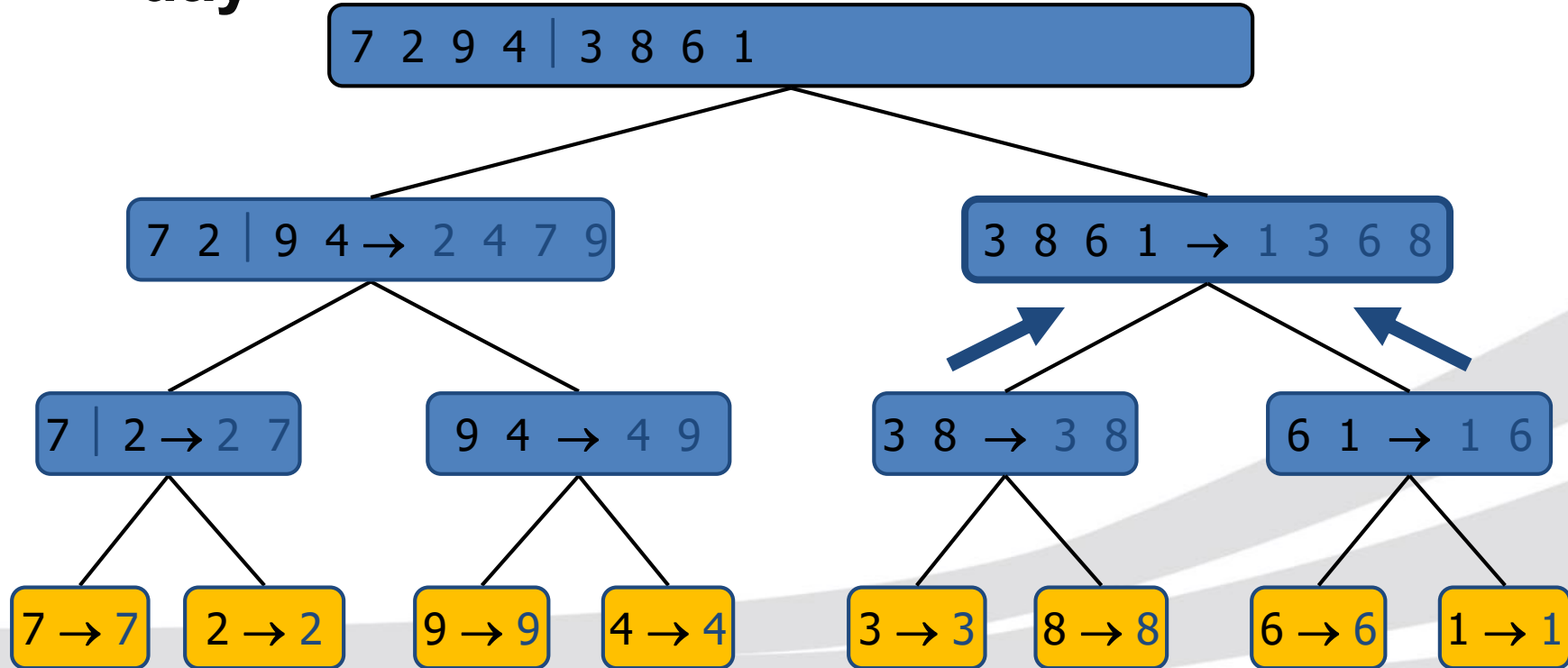
❖ Trộn merge(A,1,2,4)



Mô tả quá trình thực hiện sắp xếp



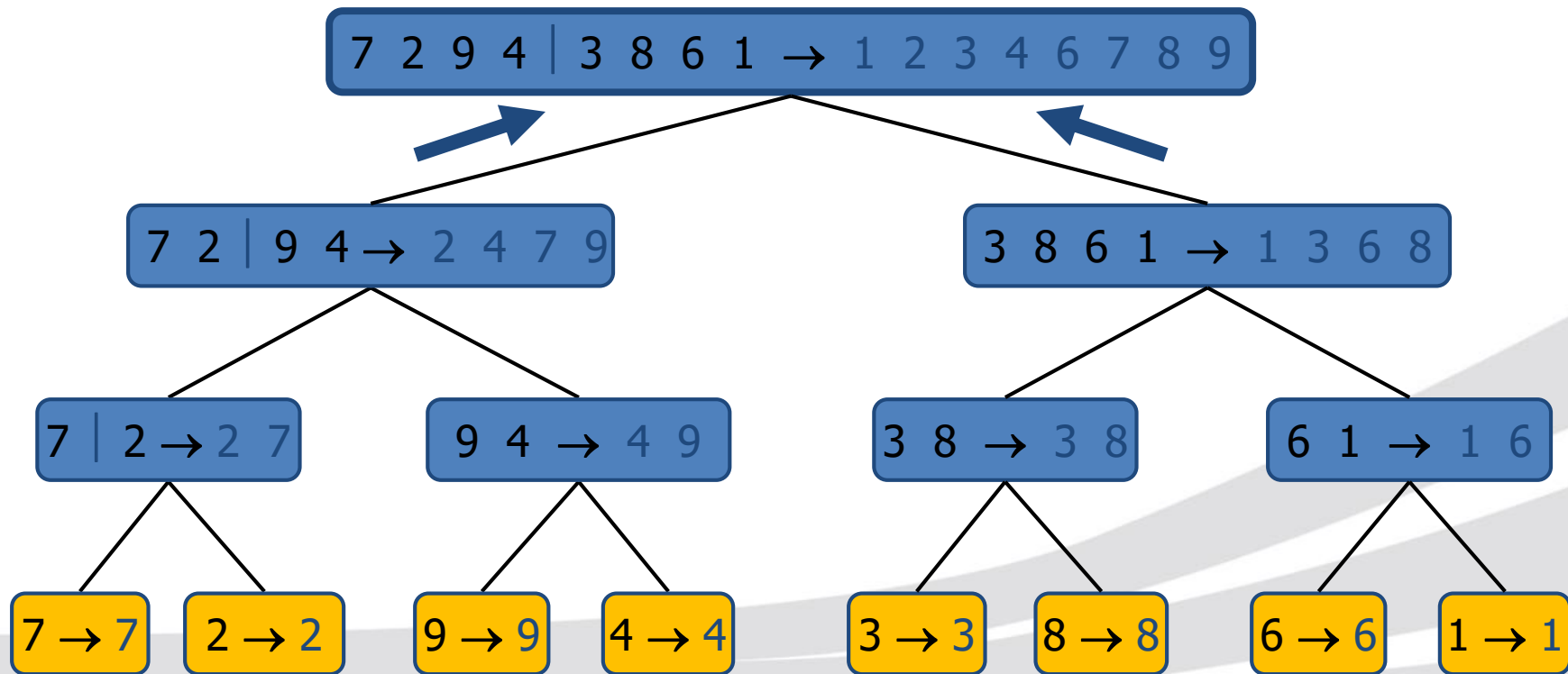
❖ Tương tự như trên với nửa bên phải của dãy



Mô tả quá trình thực hiện sắp xếp



- ❖ Trộn hai nửa dãy thành dãy được sắp merge(A, 1, 4, 8)

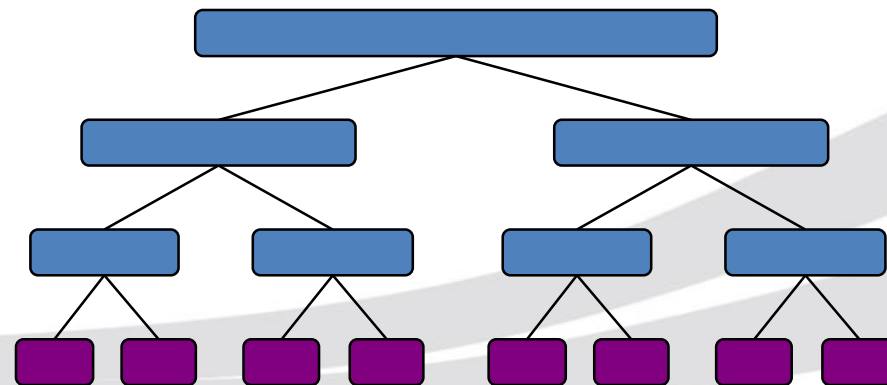


Thời gian chạy của thuật toán



- ❑ Chiều cao h của cây merge-sort là $O(\log n)$
 - Tại mỗi bước gọi đệ quy ta chia dãy cần sắp thành hai phần,
- ❑ Thời tổng thời gian làm việc trên các nút ở mức i nhiều nhất là $O(n)$
 - Chúng ta chia và trộn 2^i chuỗi có kích thước là $n/2^i$
 - Chúng ta gọi 2^i+1 lần đệ quy
- ❑ Vì vậy, tổng thời gian chạy của thuật toán mergesort là $O(n \log n)$

ĐSâu	#dãy	size
0	1	n
1	2	$n/2$
i	2^i	$n/2^i$
...



Hết