Parallel Smith-Waterman Algorithm

- Local Sequence Alignment -

myson @ postech.ac.kr

CSE700-PL @ POSTECH

Outline

- Smith-Waterman algorithm
- How to parallelize?
- Parallel Smith-Waterman algorithm
  - A wavefront algorithm
  - A scalable algorithm
  - Variants of a wavefront algorithm

Smith-Waterman Algorithm

- Local sequence alignment
- Sequence similarity comparison

ex) S=ATAGCT and
T=GATATGCA

```
S  =  -ATA-GCT
T  =  GATATGCA
```

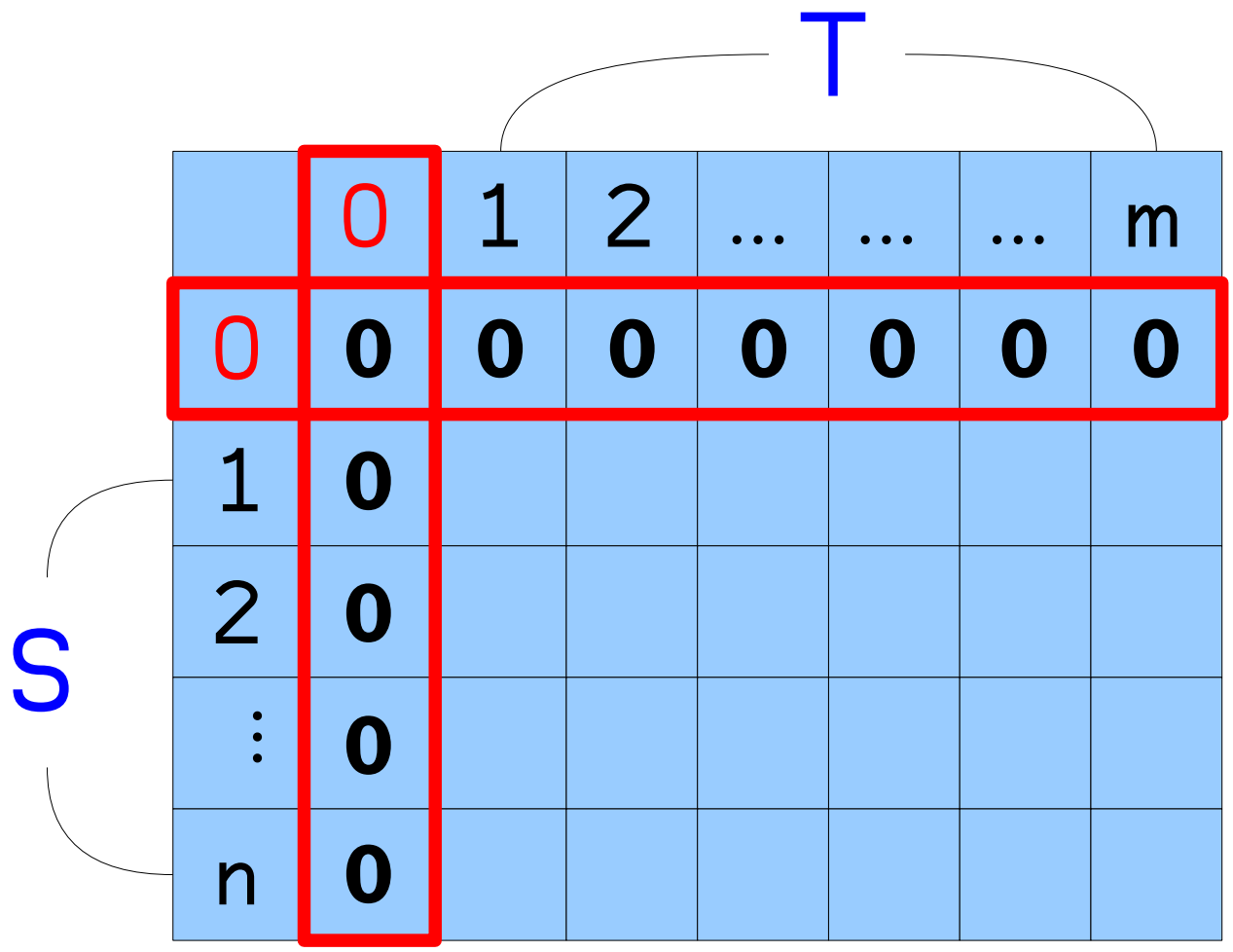
Smith-Waterman Algorithm (cont.)

1. Making a similarity matrix
2. Identifying the optimal alignment

–	–	G	A	T	A	T	G	C	A
–	0	0	0	0	0	0	0	0	0
A	0	0	−1	0	−1	0	0	0	−1
T	0	0	0	−2	0	−2	0	0	0
A	0	0	−1	0	−3	−1	−1	0	−1
G	0	−1	0	0	−1	−2	−2	0	0
C	0	0	0	0	0	0	−1	−3	−1
T	0	0	0	−1	0	−1	0	−1	−2

Similarity Matrix

- |S| = n, |T| = m



Similarity Matrix (cont.)

#define *Match* (-1)
#define *Mismatch* 1
#define *Gap* 2

Score(i, j) = (S[i] == T[j] ? *Match* : *Mismatch*)

A[i][j] = min (0,
A[i-1][j] + *Gap*,
A[i][j-1] + *Gap*,
A[i-1][j-1] + **Score**(i, j))

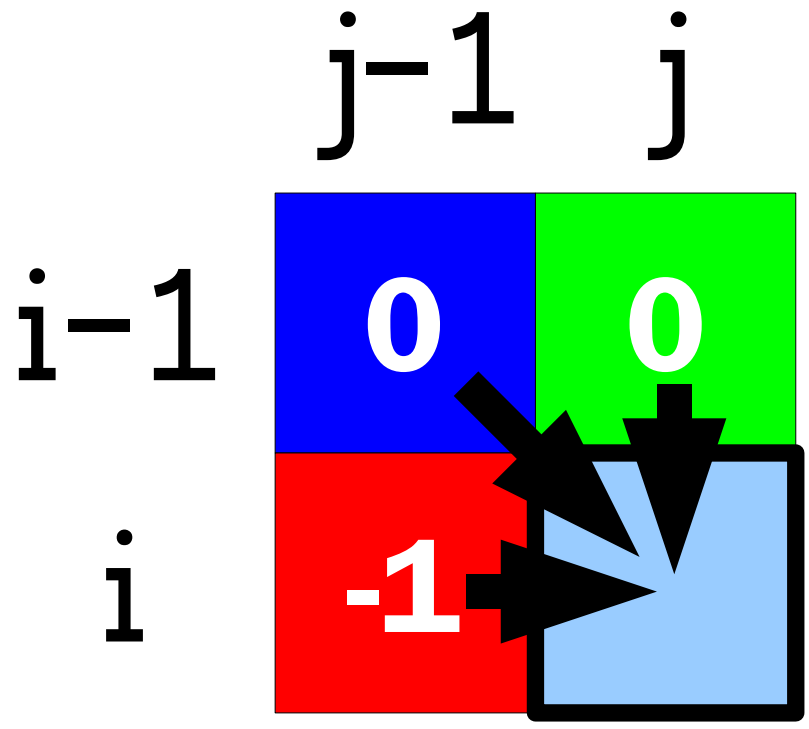
Similarity Matrix (cont.)

A[i][j] = min (0,

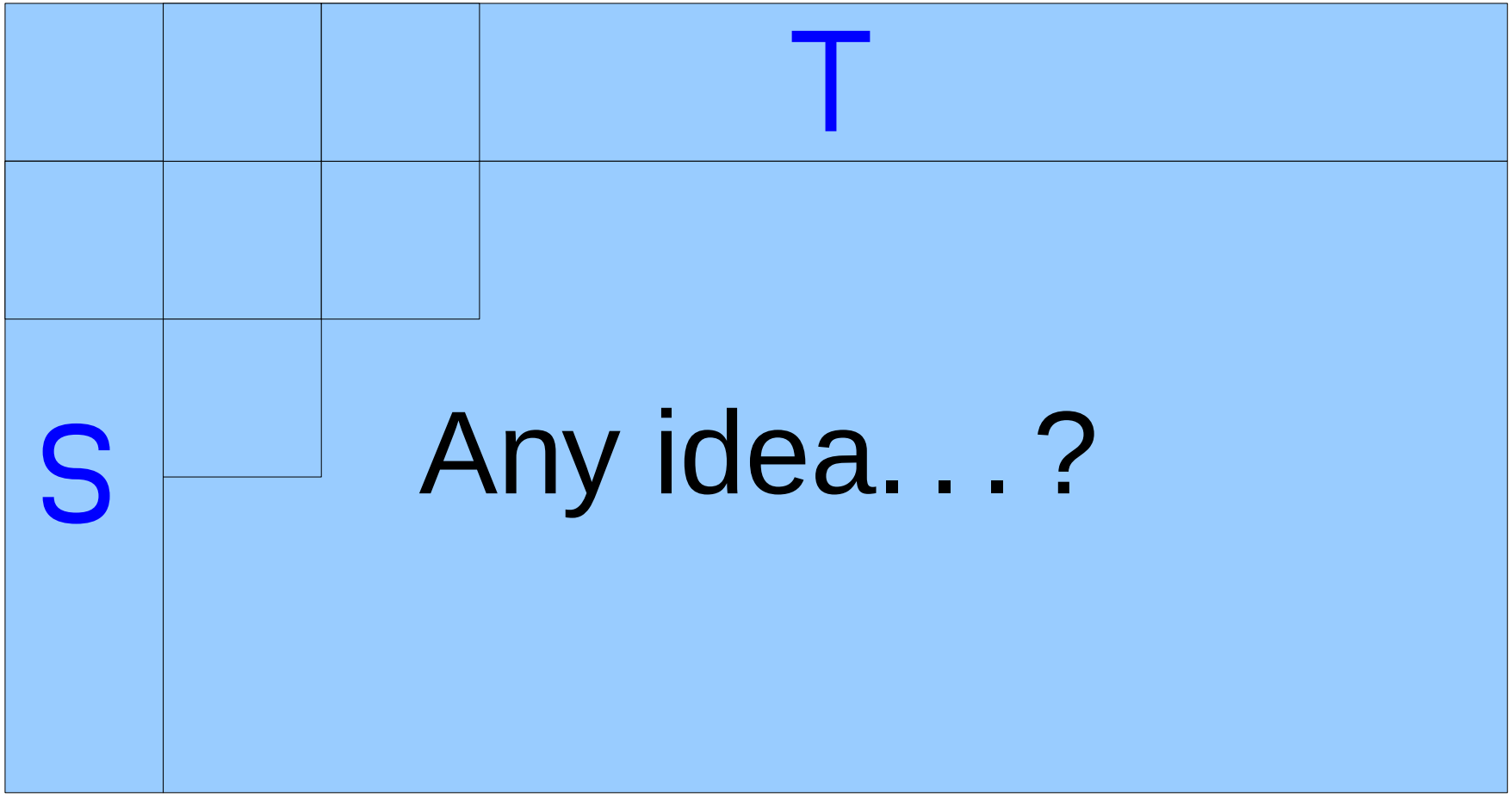
A[i-1][j] + *Gap*,

A[i][j-1] + *Gap*,

A[i-1][j-1] + **Score**(i, j))

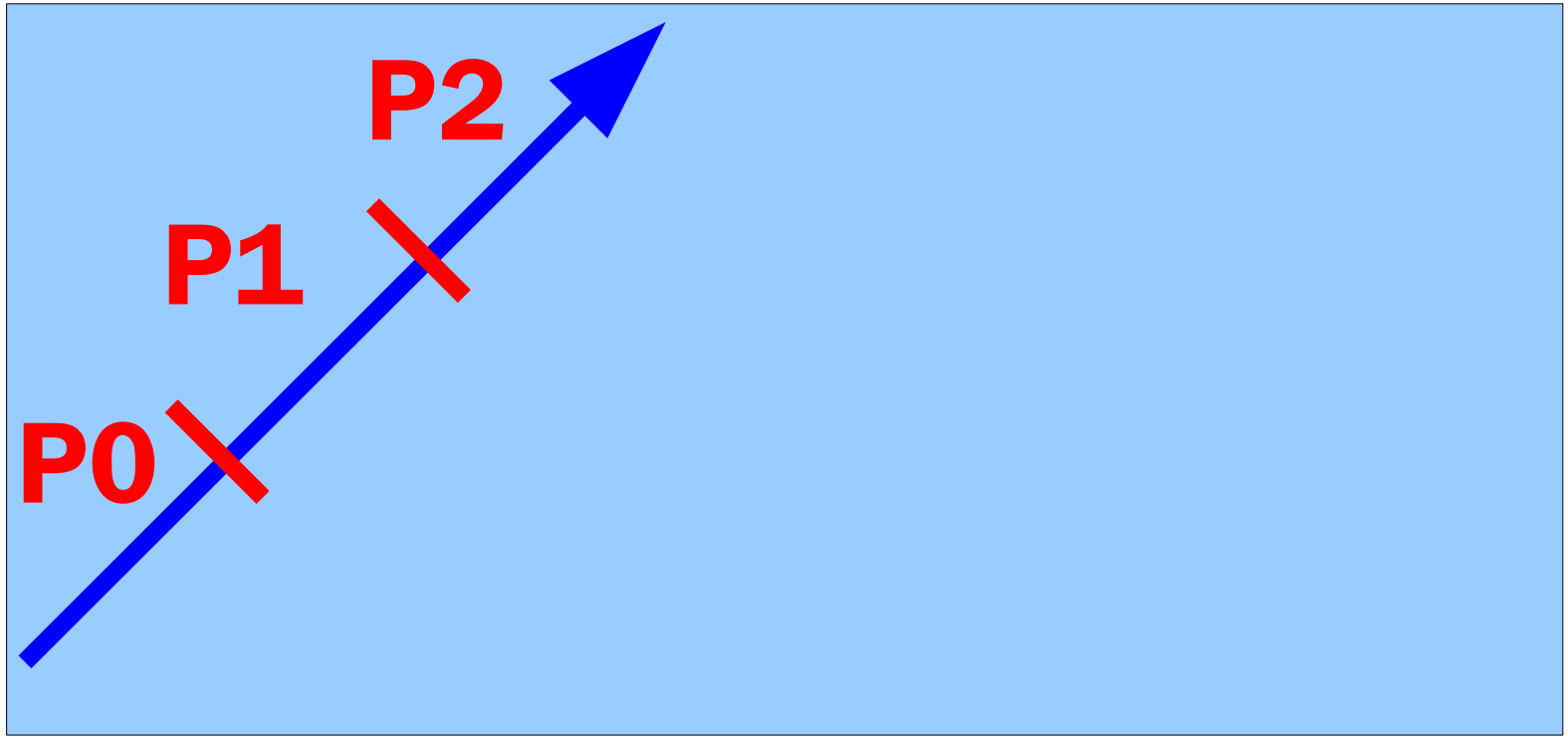


How to parallelize?



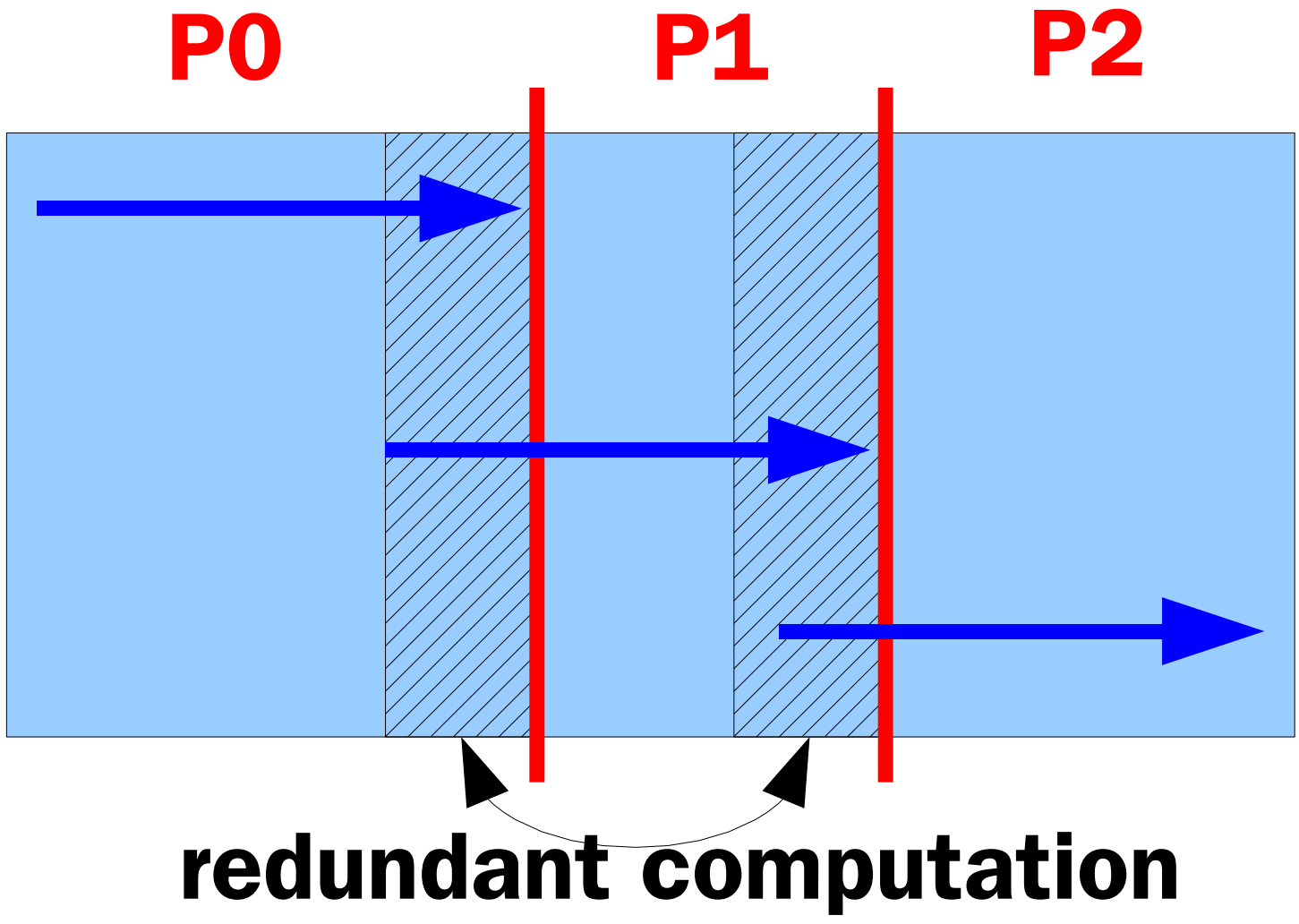
Parallel SW - Algorithms

- A wavefront algorithm



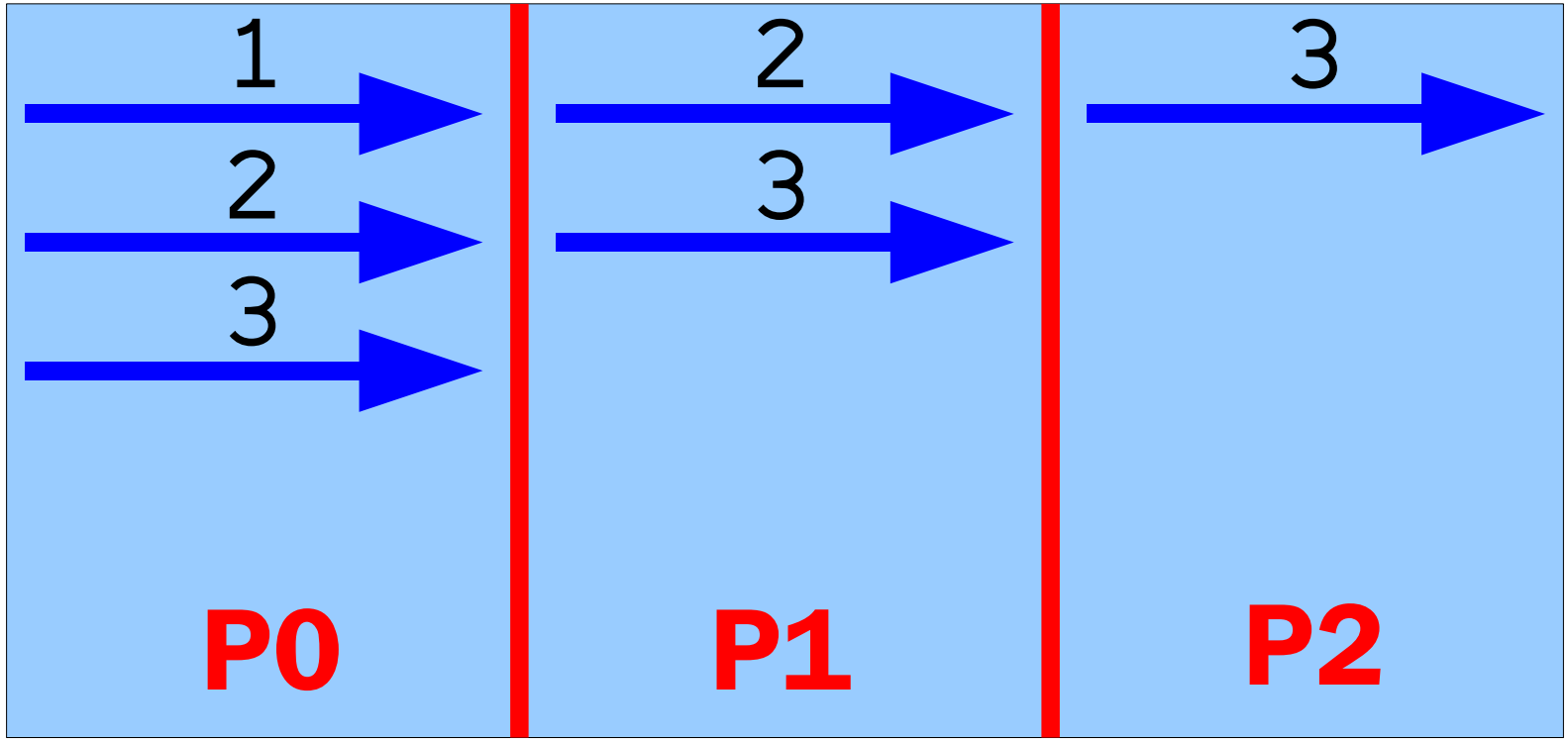
Parallel SW - Algorithms (cont.)

- A scalable algorithm



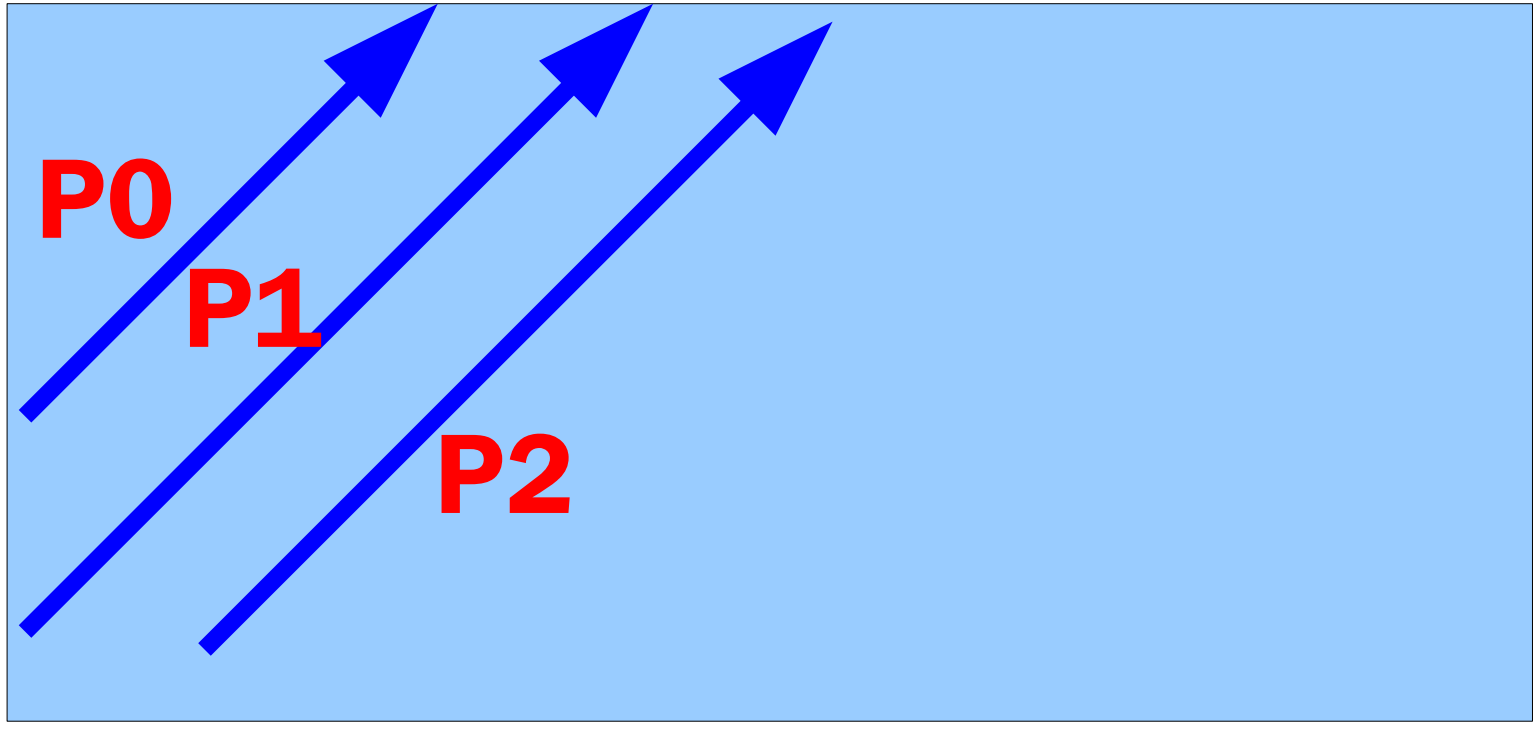
Parallel SW - Algorithms (cont.)

- Variants of a wavefront algorithm - 1



Parallel SW - Algorithms (cont.)

- Variants of a wavefront algorithm - 2



How to parallelize?

Any idea. . . ?

End

Any Questions. . . ?