PARALLEL PROCESSING

Mohammed Alabdulkareem

kareem@ksu.edu.sa

Office 2247



- Sieve algorithm for constructing the list of all prime numbers in the interval [1, n]
 - Initialize X[1]=1, all remaining X[2]..X[n]=0
 - At each step the next unmarked number m in the list is prime and all its multiples are marked none prime beginning with $\mathsf{X}[m^2]$
 - The algorithm stops when $m^2 > n$
 - All unmarked numbers are prime.



1	2 1- 2	3	4	5	6	7	8	9 10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
		3 n=3		5		7		9	11		13		15		17		19		21		23		25		27		29	
	2	3		5 m=5		7			11		13				17		19				23		25				29	
	2	3		5		7 n−7			11		13				17		19				23						29	



P	Curn	ent P	rime	Index					
	1	2	П		n				

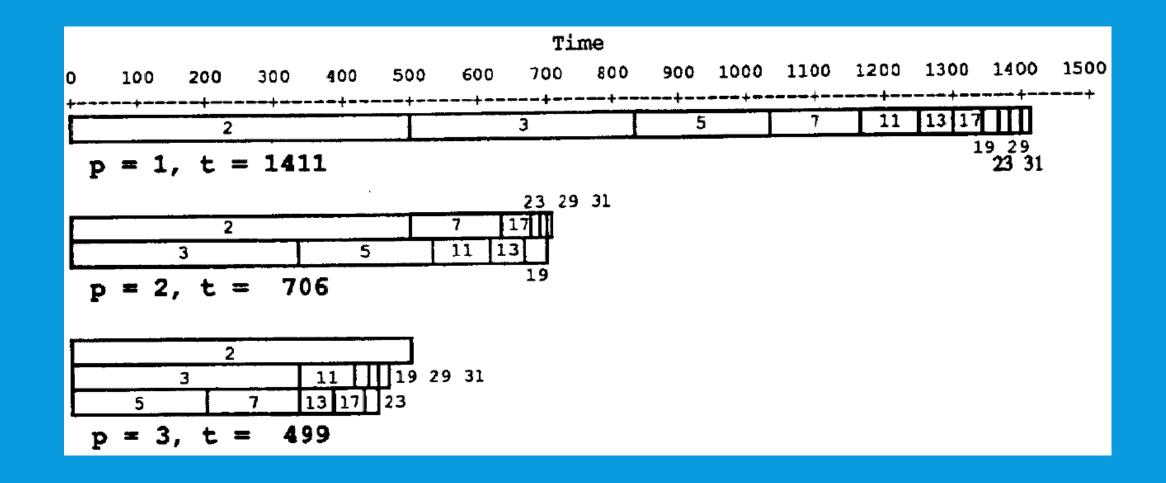


- Shared memory parallel algorithm
 - Shared current_prime, list _of_numbers
 - Index is local to each processor
 - An idle processor check the shared memory for the next prime *m* and updates the current_prime accordingly.
 - Using its local index , mark the multiples of m



P ₁ Index	P ₂ Index	P _p Index						
Shared Memory	Current Prime							
1	2	n						

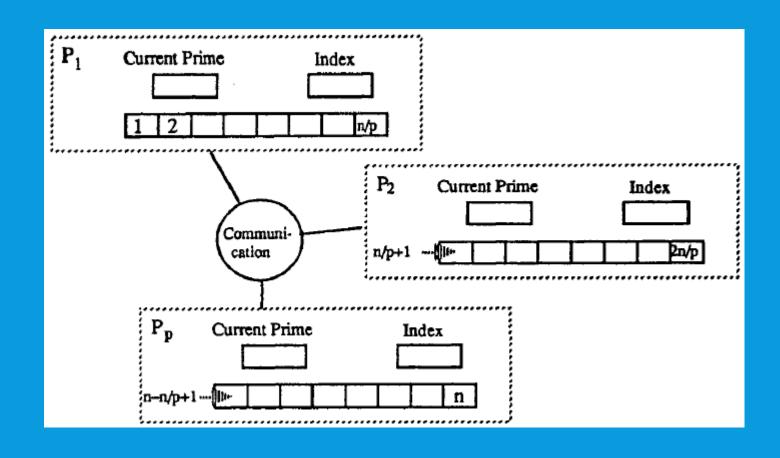






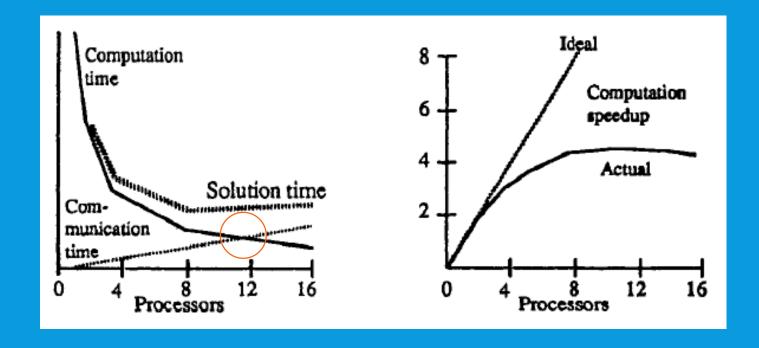
- Data Parallel Approach.
- *p* is the number of processors.
- The bit-vector representing the *n* integers is divided into *p* equal-length segments.
- The condition here is $p < \sqrt{n}$ to make sure all prime numbers are in the first segment.
- The first processor p1 will work as coordinator. It find the next prime and broadcast it to all other processors to mark the multiples as not prime.
- The total time = communication time + computation time.
- Communication time is time spent on transmitting the selected primes to all processors.
- Computation time is time spent by each processor marking their sub lists.





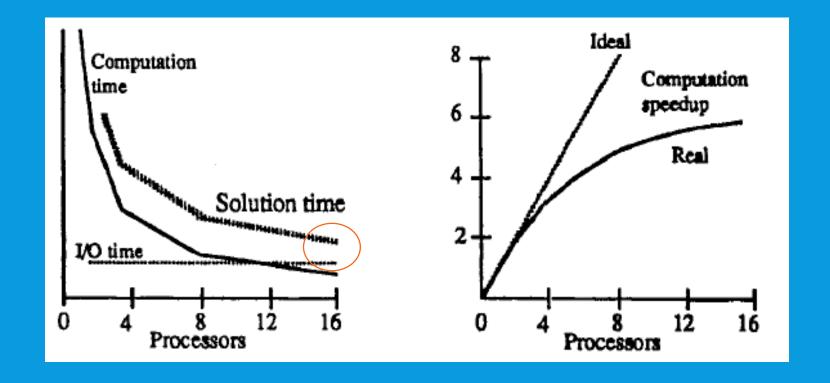


Effect of communication time





Effect of constant I/O Time

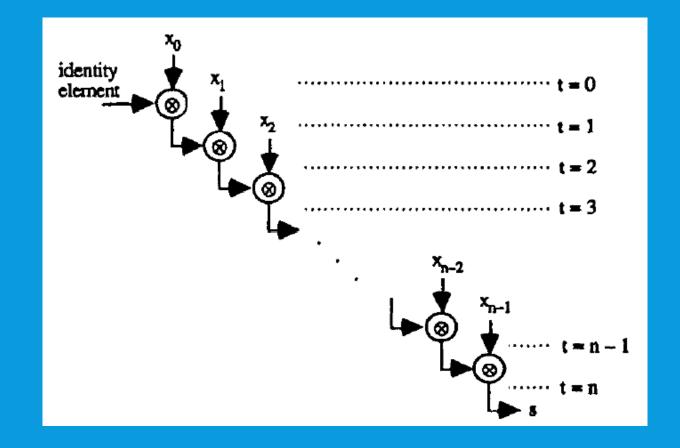




- Semigroup computation (reduction or fan-in)
- Let \otimes be an associative binary operator; i.e., $(x \otimes y) \otimes z = x \otimes (y \otimes z)$ for all x, y, $z \in S$. A semigroup is simply a pair (S, \otimes) , where S is a set of elements on which \otimes is defined.
- Given a list of *n* values *x*0, *x*1, . . . , *xn*-1, compute *x*0 \otimes *x*1 \otimes . . . \otimes *xn*-1
- Common examples for the operator \otimes include +, ×, \wedge , \vee , \oplus , \cap , \cup , max, min.
- The parallel algorithm can compute chunks of the expression using any partitioning scheme.
- The chunks must be combined in left-to-right order.

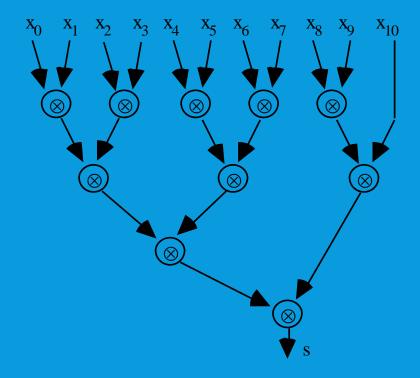


Semi-group computation on one processor





Semi-group parallel computation



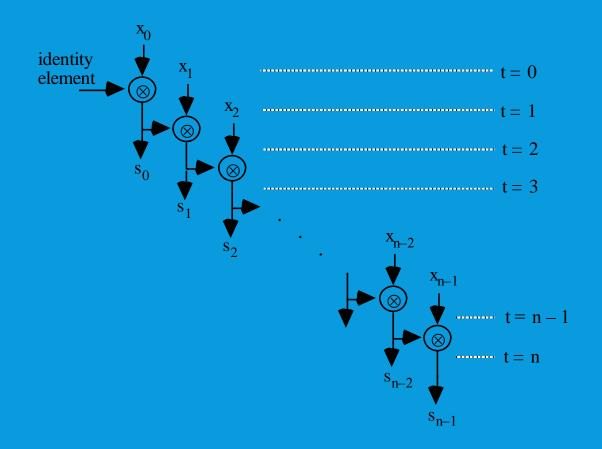


Prefix Sum

i=		1							8	9
A[]=	3	5	2	-4	6	10	4	-5	3	2

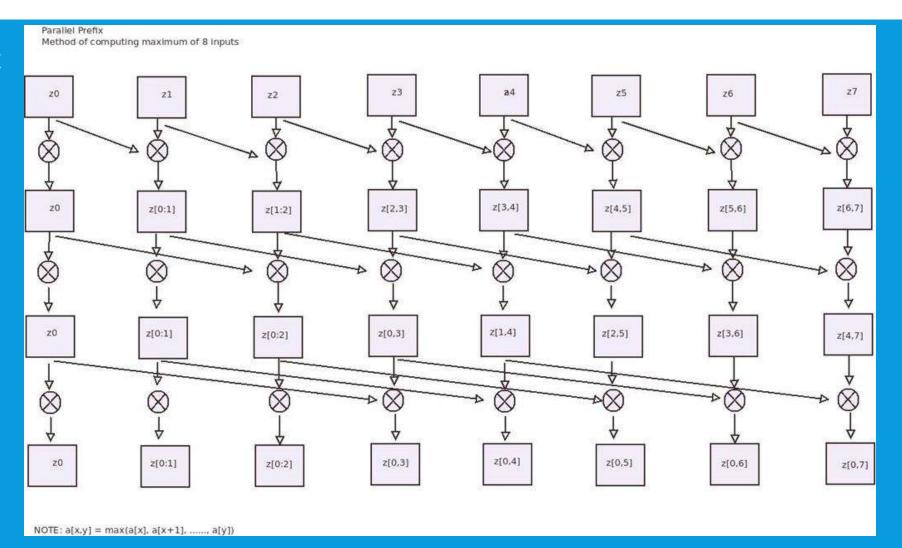


Parallel prefix





Parallel prefix



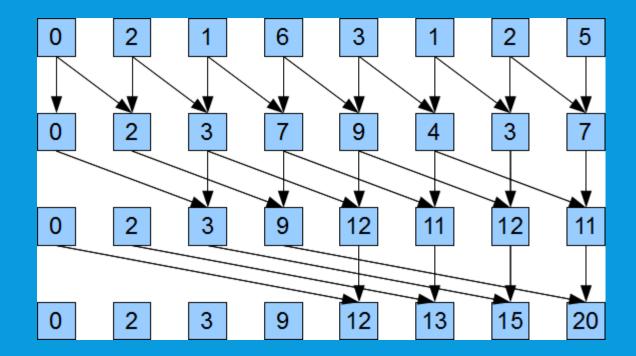


Parallel prefix

for
$$d = 0$$
 to $\log n - 1$ do
for $i = 0$ to $n - 1$ by 2^{d+1} do in parallel
 $a[i + 2^{d+1} - 1] := a[i + 2^d - 1] + a[i + 2^{d+1} - 1]$



- Parallel prefix
- simultaneously evaluating all of the prefixes of the expression.





- Packet Routing :
 - A packet resides on processor *i* and need to be sent to processor *j* .
 - The packet may need to pass through intermediate processor.
 - The problem is complicated when a processor has more than one packet to be routed at the same time.
 - One-to One communication.



Broadcast:

- Disseminate the value α to all processors as fast as possible.
- One-to-all communication.

• Multicast:

- Disseminate the value a to some processors.
- One-to-many.



Sorting:

Given a list of values x1, x2, ..., xn rearrange the values such that

x1, x2, ..., xi, xj, ..., xn where $xi \le xj$ for all i and j

In our examples we will sort records based on key value, we will just sort the keys.

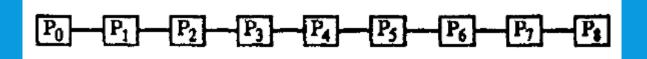


- D the diameter is the longest of the shortest distances between any two processors.
- d is the maximum node degree the number of links at each processor.

- Some simple architectures:
 - 1. Linear array of processors
 - 2. Binary tree of processors
 - 3. Two-dimensional mesh of processors
 - 4. Multiple processors with shared variables

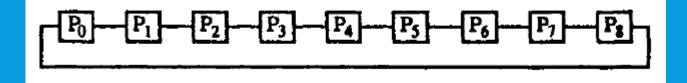


- Linear array of processors
- The diameter of linear array of p-processors is D=p-1
- The maximum node degree (number of links) d=2.



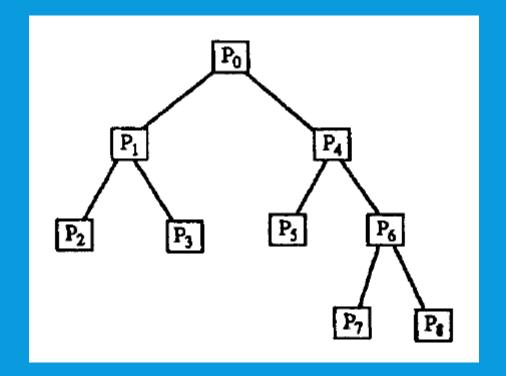


- The ring linear array of processors
- The diameter of linear array of p-processors is $D = \left\lfloor \frac{p}{2} \right\rfloor$
- The maximum node degree (number of links) d=2.



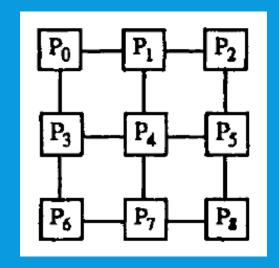


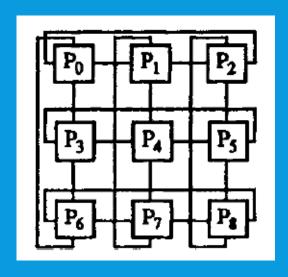
- Binary tree of processors (the tree is balanced the leaf levels differ by at most 1)
- The diameter of binary tree of p-processors is D=2[log p] and the degree d=3





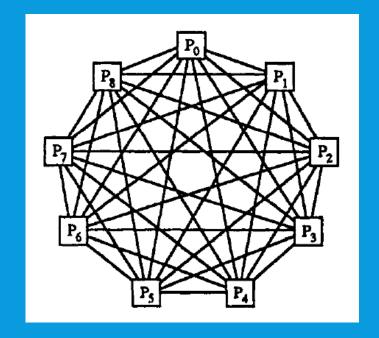
- 2D mesh of processors
- The diameter of 2D mesh of p-processors is $D=2\sqrt{p}$ -2 and d=4
- What about tours?







- Shared memory can be represented as complete graph.
- The diameter of shared memory multiprocessor is D=1 and d=p-1
- Think about the cost?





• We will see how to implement some of the simple computations on the simple types of architectures.