

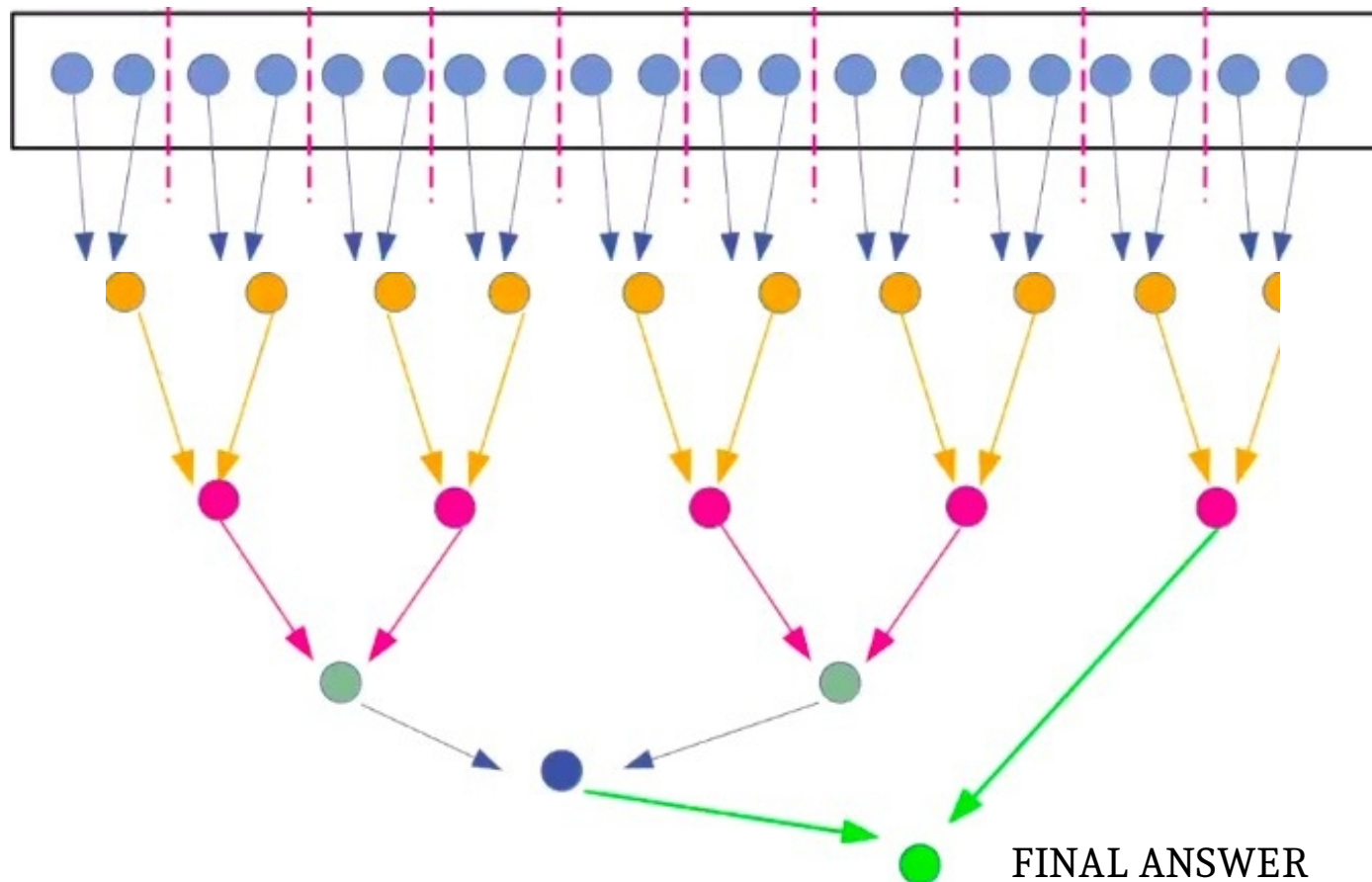
Why are we building parallel systems?

- Single processor performance
 - Transistor density
 - IC
 - Speed
 - ILP
- Speed increases, power increases, heat increases
- Chips too hot to be reliable. Limits to heat dissipation
- How to exploit the increasing number of transistor density?
 - Add parallelism
 - Go multicore

Why we need to write parallel programs?

- Programs written for conventional single core systems cannot exploit presence of multiple cores
- We can run multiple instances of a program on a multicore system.
BUT this is of less help
- e.g. running multiple instances of gaming program is not desirable !
 - We want the game to run faster !
- Rewrite serial programs to parallel programs so that they use multiple cores
- Rewrite or translate serial code
 - Translate => can we auto-translate?
 - Till now researchers have had limited success in writing programs that convert the serial program to parallel program
 - It is not straightforward to convert serial to parallel by identifying par-constructs
- Rather, step back and devise entirely new algorithm


Example: sum of array elements



Example

- Sum of all elements of an array
- Tough for the translator program to discover the procedure
- Certain softwares can identify common serial constructs and efficiently parallelise them
- HOWEVER to apply the principles on ever increasing complex serial programs is difficult
- Therefore, we cannot simply continue to write serial programs.
- We MUST write parallel programs to exploit the power of multiple cores

How do we write parallel programs

- Basic idea is of partitioning the work to be done among the cores
 - Task parallelism
 - Data parallelism
- Ex: Prof. P, TAs: A, B, C, D 
 - Have to check 100 answer books with 5 questions each
 - Data parallel => P, A, B, C, D each take 20 copies to check
 - Task parallel => P, A, B, C, D each take 1 question each to check