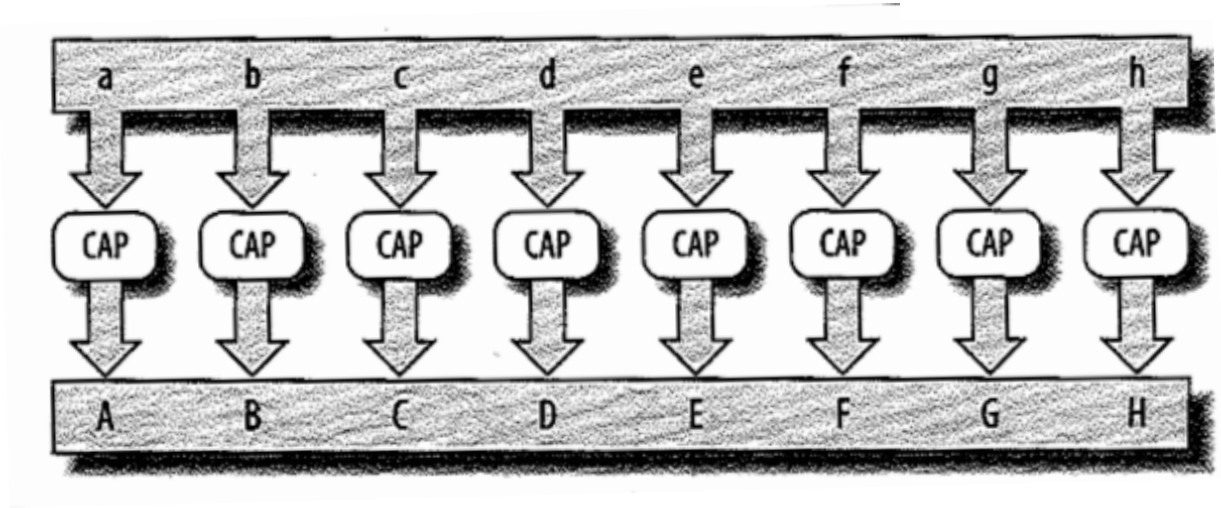
Kinds of parallelism

- Data parallelism
  - The same task run on different data in parallel
- Task parallelism
  - Different tasks running on the same data
- Hybrid data/task parallelism
  - A parallel pipeline of tasks, each of which might be data parallel
- Unstructured
  - Ad hoc combination of threads with no obvious top-level structure

Pictures in following slides due to James Reinders

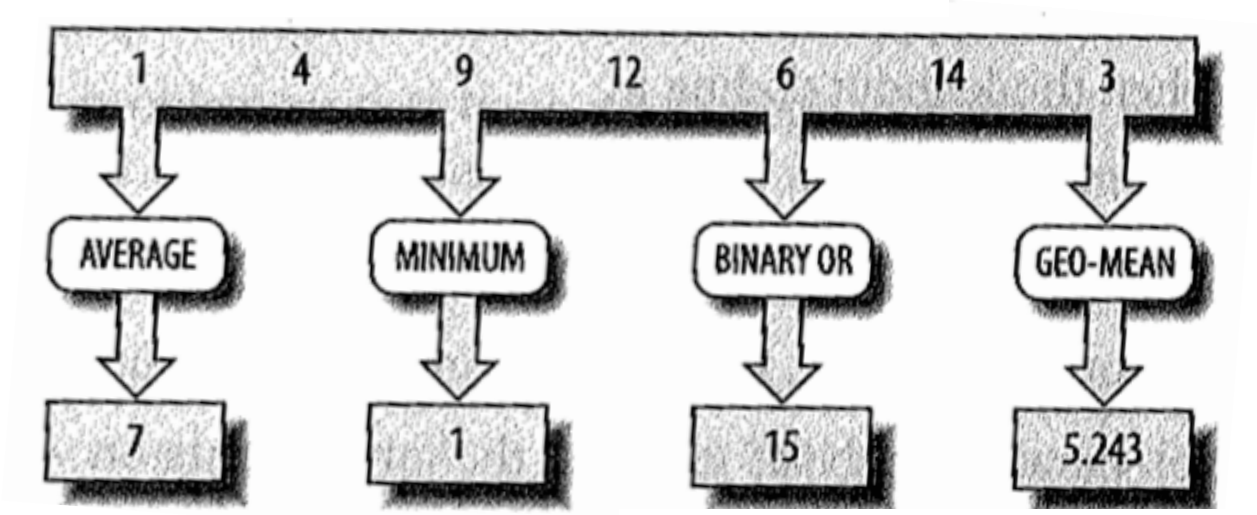
Data parallelism

- Example: convert all characters in an array to upper-case
  - Can divide parts of the data between different tasks and perform the tasks in parallel
  - Key: no dependencies between the tasks that cause their results to be ordered



Task parallelism

- Example
  - Several functions on the same data: average, minimum, binary or, geometric mean
  - No dependencies between the tasks, so all can run in parallel



Pipeline parallelism

- Output of one task is the input to the next
  - Each task can run in parallel
  - Throughput impacted by the longest-latency element in the pipeline

Pipeline load balancing

- Assign more than one computational process to each task
  - Combines data- and pipeline- parallelism