LogP in Cloud computing

LogP in Cloud computing

- Made for clusters.

- Differences between Cloud and Cluster environment
  - Latency
  - Overhead
  - Gap

- Does it work?

Why may not work?

- Cloud vs Cluster
    - Imprecise to make variables.
    - A same request for a cloud infrastructure can have different configuration results.
    - Latency, overhead and gap may vary depending on the configuration, traffic or the use of shared processors.

Test development

- Use of Open MPI library

- Measuring ‘L’,’o’ and ‘g’

O ⇨Use an non-blocking messages method proposed by “Measuring MPI Send and Receive Overhead and Application Availability in High Performance Network Interfaces”

L ⇨(Total communication time - overheads)

g ⇨(Time for two messages - 2*Time for single message)

Possible solution

If our hypothesis be proved true, we propose a possible solution:

- Probabilistic algorithm

- LogP would become L(o^s)(g^t)P:
    - s: probabilistic part that refers to usage of the processor IO
    - t: probabilistic part that refers to usage of the network

L(o^s)(g^t)P

- L: average latency of the network. (considering the usage of each part of the network)
- o: overhead of the processor.

- s: probable stack message size
    - Consider splitted processors with other programs/algorithms.

- g: gap between messages in the network.

- t: probable traffic in the network

    - Consider network traffic caused by other programs/algorithms.

Testing the solution

First tests:

- Same pattern used in the cloud LogP test.

- Same algorithms and data.

- Analyse the results with the proposed model.

What to consider

- First hypothesis be proved false.
    - Disconsider the proposed solution.
- Second solution be inaccurate.
- Complexity of the proposed solution.

Algorithm used to test (Matrix multiplication)

Block-Striped Decomposition algorithm with square matrices

$$A(n \times n) \ . \ B^T(n \times n) \ = \ C^T(n \times n)$$

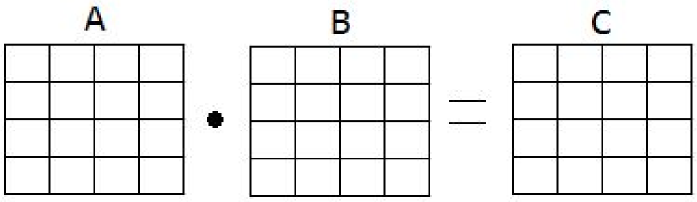
- Each subtask at each iteration holds a row from A and a column from B and calculates an element from the C matrix

- n^2 subtasks, so $P < n^2$

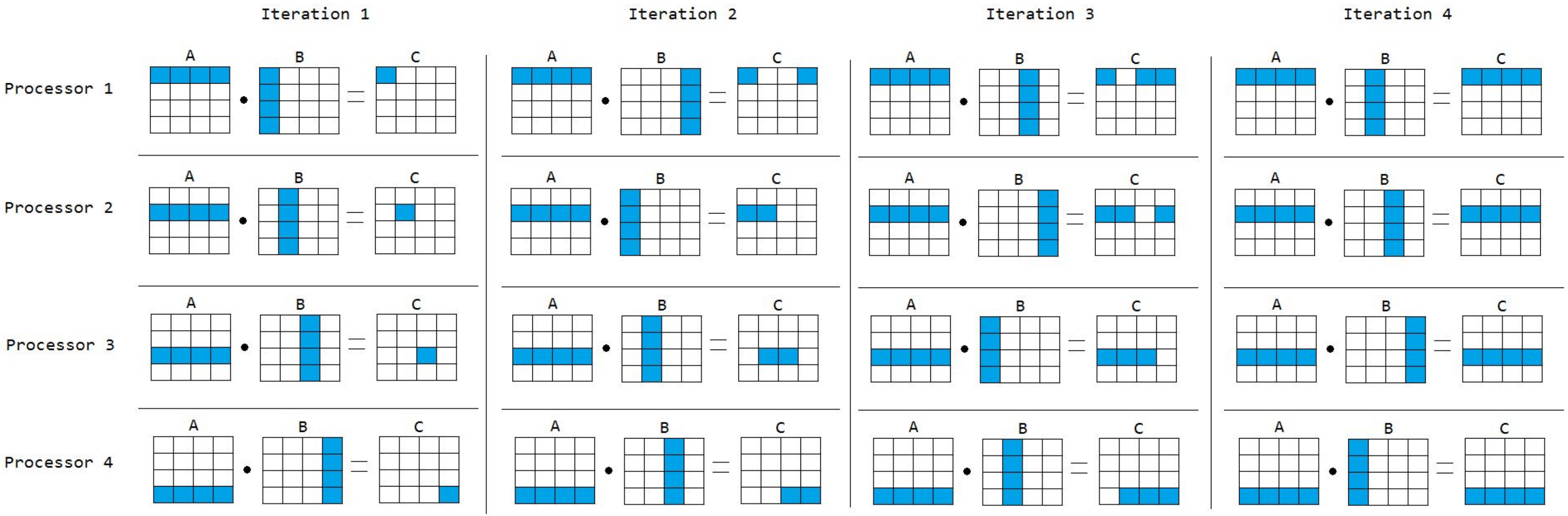
Block-Striped Decomposition example with P = n

- Each processor will execute n subtasks and calculate a row from C matrix

- At every iteration
    - The processors will calculate the product of a row from A with a column from B, resulting an element from C
    - Each processor i will send the column from B to its neighbor (i+1) mod n



Block-Striped Decomposition example with P = n



$$O(nL + 2no + n^2g)$$

Why choosing this algorithm?

- Variable number of processors up to n^2.
- Deterministic number of operations and iterations.
- Clear analysis since the algorithm is simple.

Bibliography

Culler, David; Karp, Richard; Patterson, David; Sahay, Abhijit; Schauser, Klaus Erik; Santos, Eunice; Subramanian, Ramesh; Von Eicken, Thorsten. (July 1993). LogP: Towards a realistic model of parallel computation.

ALEXANDROV, A., IONESCU, M., SCHAUSER, K. E., AND SCHEIMAN, C. LogGP: Incorporating Long Messages into the LogP model - One step closer towards a realistic model for parallel computation. In 7th Annual Symposium on Parallel Algorithms and Architectures (May 1995).

Douglas Doerfler; Ron Brightwell. Measuring MPI Send and Receive Overhead and Application Availability in High Performance Network Interfaces

Parallel Matrix Algorithms. Retrieved from https://www3.nd.edu/~zxu2/acms60212-40212/Lec-06-1.pdf