CMSC 714 paper reading round 1

LogP
A Practic Parallel

Jing Xie

Mar 11, 2021

Culler, David E., et al. "LogP: A practical model of parallel computation." Communications of the ACM 39.11 (1996): 78-85.

Problems

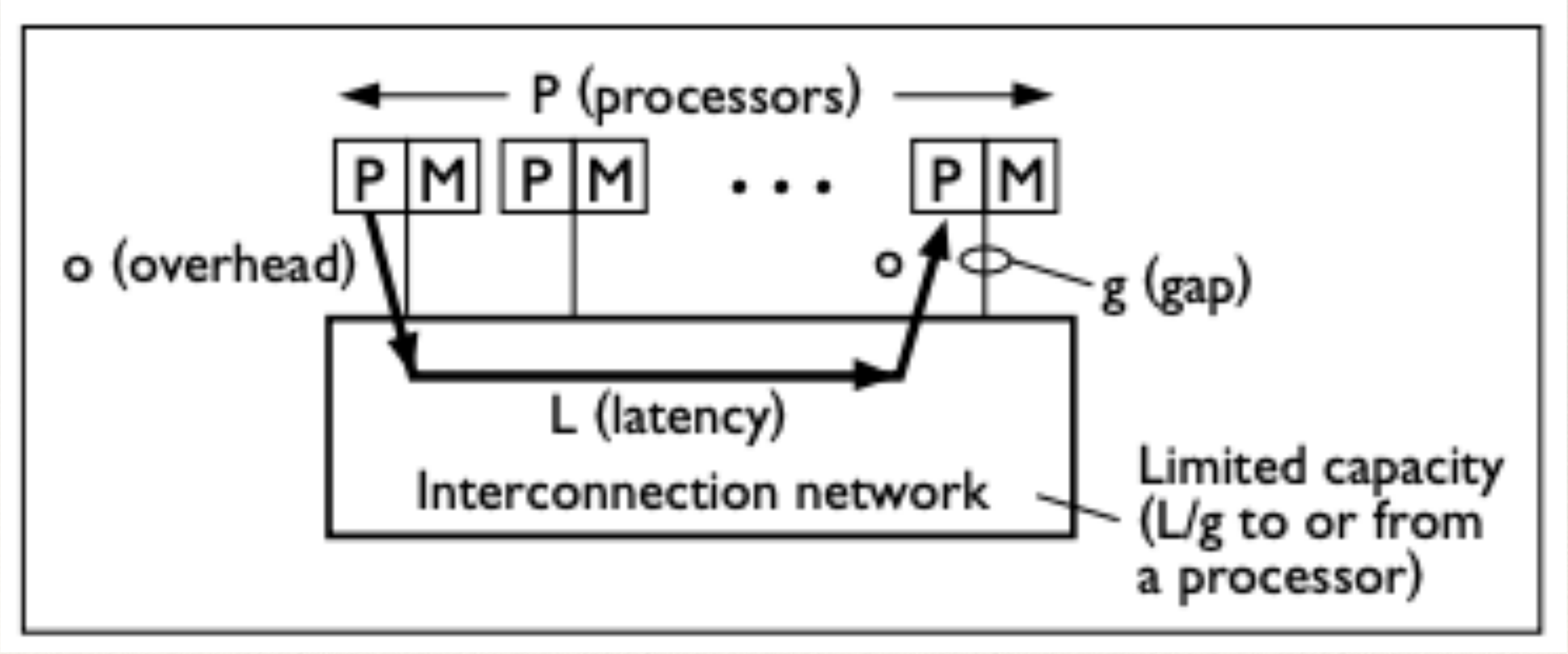
- Many current parallel models are impractical
  - Artificial factors.
- PRAM
  - A collection of synchronous processors.
  - communicate with global random access memory
  - Processors can access any memory cell in unit time.
  - Variation of PRAM make it more practical
- Network models lack robustness

Motivations

- BSP (bulk synchronous parallel model)
  - A radical variant of PRAM
  - capture key performance bottlenecks
  - drop off the details
- Apparent architectural convergence
- Phenomenal increase of microprocessor performance and capacity.
  - x100/x1000 64-bit off-the-shelf processors.
  - A large number of data on each processor.
- Network topology lags far behind.
  - high latency, overhead of communication, limited bandwidth
- Physical interconnect underlying a program is different.

LogP model

- A model of a distributed-memory multiprocessor.
- Specify the performance factors of interconnection network without using the structure of it.



LogP model

- Assumptions:
- To deal with variant latency in asynchronous processor
  - L is the upper bound
- All messages are of a small and fixed size
- Network has a finite capacity.
- Not attempt to model local computations.

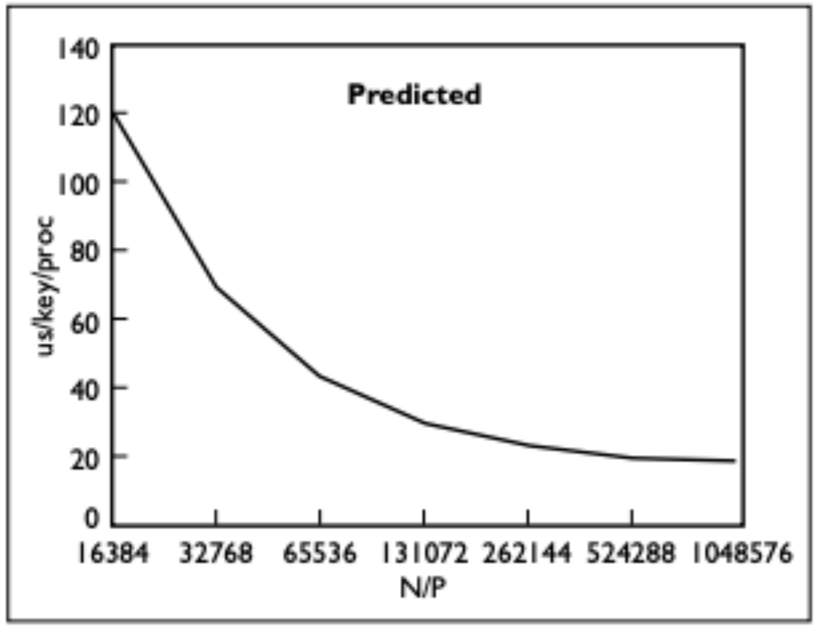
LogP model

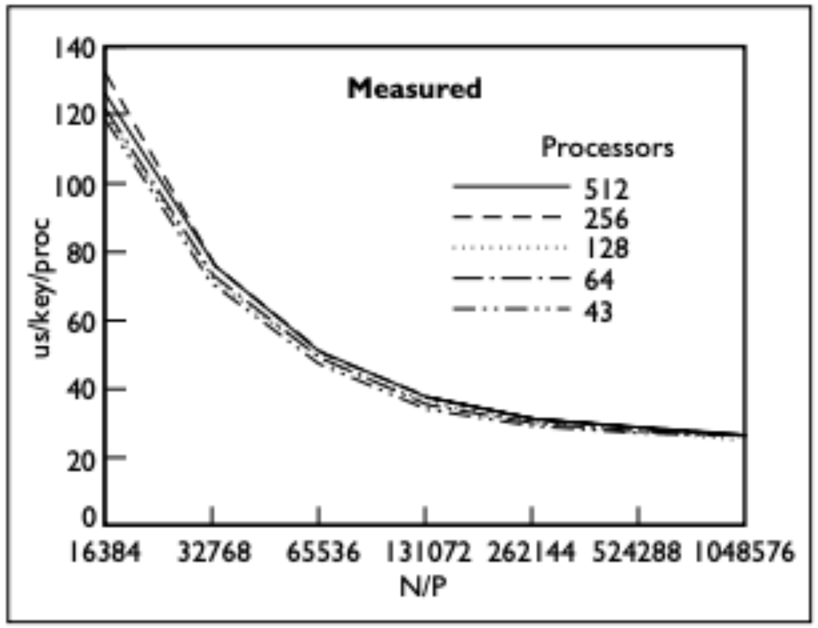
- Parameters in LogP are not equally important. Trivial params are discarded.
  - A trade-off between capturing more execution characteristics and providing a reasonable framework for analysis.
- Loopholes that other models permit are discouraged
- Encourage techniques that work well in practice
  - Reduced the amount of communication
  - Careful scheduling of communications.

Utility of the LogP model

- Benefits:
- Solutions to basic theoretical problems under LogP are better than other solutions under traditional models.
- Designing according to the performance result under LogP models helps improve the quality of solutions.
- It's possible to accurately predict the performance on real machines.

Utility of the LogP model





- Predicted and measured times over problem N/P that is the total number of keys per processor.
- Predicted performance vs. measurement.
- LogP helps identify deficiencies.

Model on real machines

- Task: transmission of an M-bit message in an unloaded or lightly loaded network.
- Total message communication time:

$$T(M, H) = T_{snd} + \lceil \frac{M}{w} \rceil + Hr + T_{rcv}$$

-

Machine	Network	Cycle ns	w bite	$T_{snd} + T_{rcv}$ cycles	r cycles	avg. // (1024 Proc.)	L (M = 160) (1024 Proc.)
nCUBE/2	Hypercube	25	1	6400	40	5	6760
TMC CM-5	4-ary Fat-tree	25	4	3600	8	9.3	3714
IBM SP-2	Banyan	25	8	2100	5	9.3	1560
Meiko CS-2	4-ary Fat-tree	14	8	2700	20	9.3	3050
Intel Paragon	2d Mesh	7	16	4300	7-10	21	4450
Cray T3D	3d Torus	7	16	35	3	10	145
Dash	2d Torus	30	16	30	2	6.8	53
J-Machine	3d Mesh	31	8	16	2	12.1	60
Monsoon	Butterfly	20	16	10	2	5	30
nCUBE/2 (AM)	Hypercube	25	1	1000	40	5	1360
CM-5 (AM)	4-ary Fat-tree	25	4	132	8	9.3	246
Meiko CS-2 (AM)	4-ary Fat-tree	14	8	230	20	9.3	570
Intel Paragon (AM)	2d Mesh	7	16	540	7-10	21	750

Summary

- The communication network is abstracted into three parameters under LogP.
- Determine lower bound on parallel running time.
- Guide algorithm designs to be more efficient.