Outline

Multithread Parallelism (MP)

1

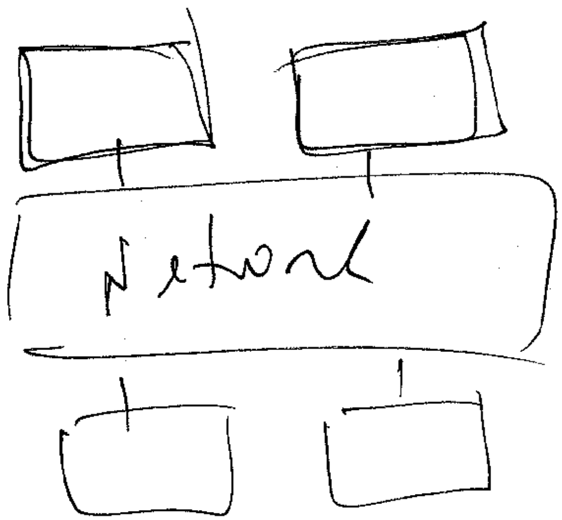
Some basics
- Granularity
- Amdahl's Law
- Metrics (speedup, efficiency, redundancy)

Multi-core (one thread spans an engine, multiple engines)
- Tightly coupled vs. Loosely coupled
- Interconnect
- Cache Coherency
- Memory Consistency

Examples
- Early: cm*, HEP, cosmic cube
- Recent: cell, power, niagara, larrabee
- Tomorrow: ACMP

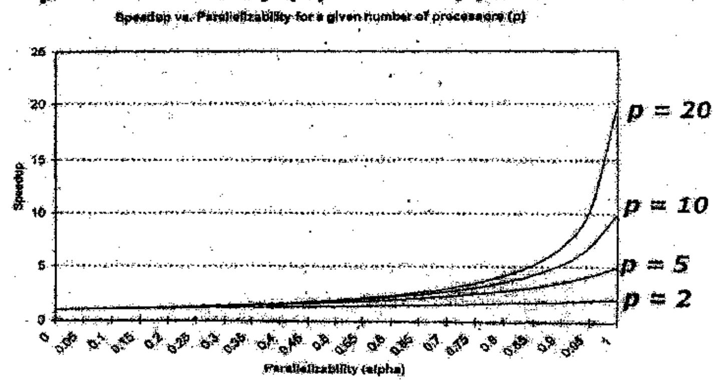
Granularity of Concurrency

- Intra-instruction (Pipelining)
- Parallel instructions (SIMD, VLIW)
- Tight-coupled MP
- Loosely-coupled MP

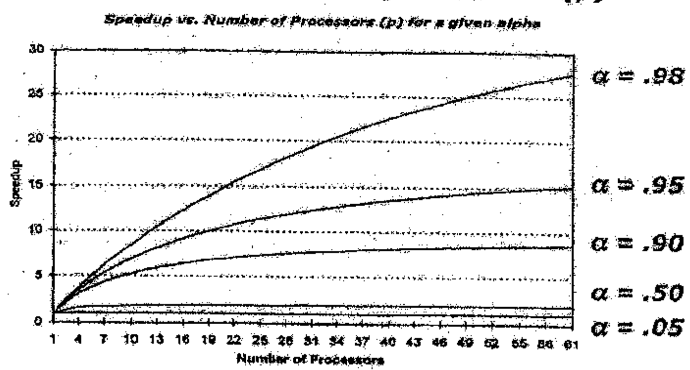


Amdahl's Law

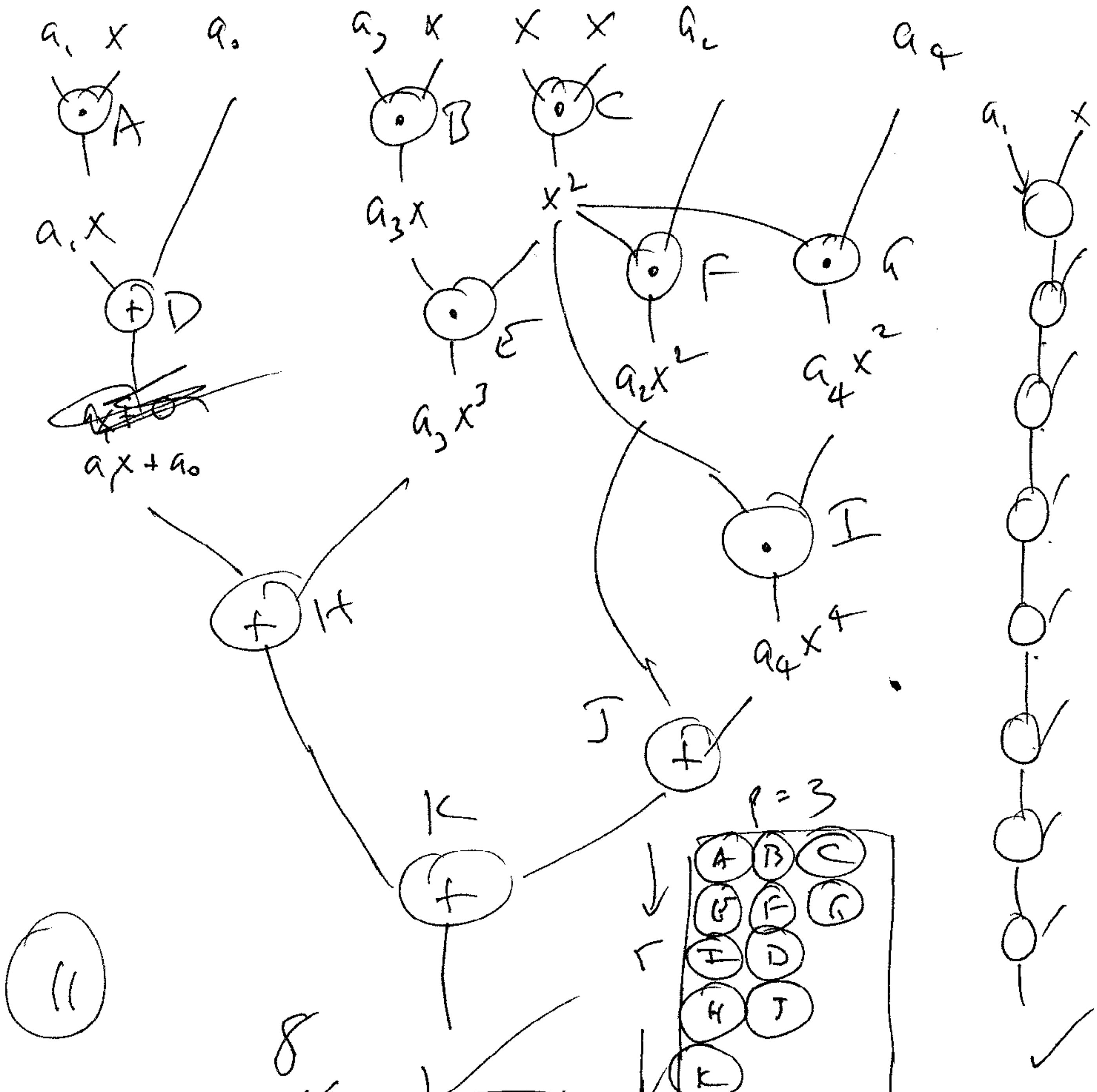
* *Speed-up as a function of the parallelizability (α) of the application*



* *Speed-up of an application as we add more and more processors (p)*



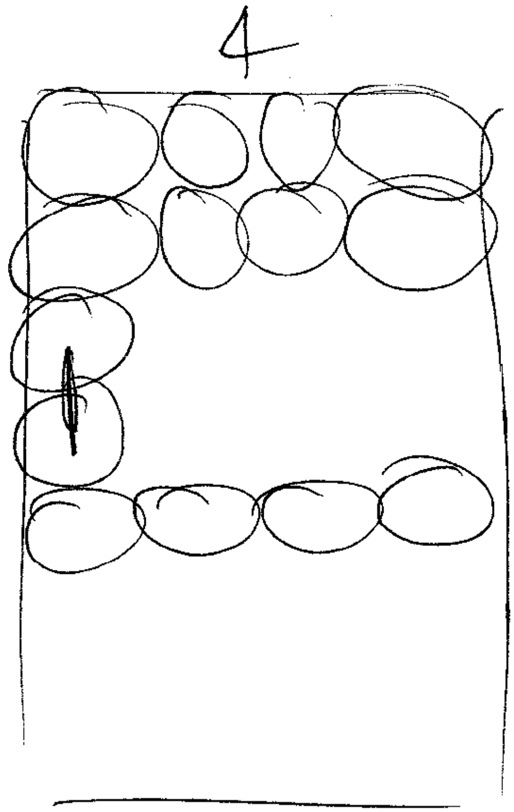
$$x(a_4 x^3 + a_3 x^2 + a_2 x) + a_1) + a_0$$



$S_3 = \frac{~~11~~\ 8}{5} = \boxed{2.2}$

$\boxed{1.6}$

4/14



VECTORIZABLE

50%	50%	
15	30	= 45
3	30	33
0	30	= 30

ONE HOUR

$$S_p(A) = \frac{T_1(A)}{T_p(A)} \quad \frac{60}{30} = 2$$

$$S_p = \frac{T_1}{T_1} = \frac{\alpha T_1 + (1-\alpha) T_1}{T_p}$$

$$T_p = \frac{\alpha T_1}{p} + \frac{(1-\alpha) T_1}{1}$$

$$S_p = \frac{T_1}{T_p} = \frac{\cancel{T_1}}{\alpha \frac{\cancel{T_1}}{p} + (1-\alpha)\cancel{T_1}}$$

$$S_p = \frac{1}{\frac{\alpha}{p} + (1-\alpha)}$$

Metrics

Speed-up: $\frac{8}{5}$

Efficiency: $\frac{1.8}{3.5} = \frac{8}{15}$

Utilization: $\frac{11}{15}$

Redundancy: $\frac{11}{8}$

MP/8

Tightly-coupled vs Loosely-coupled

Tightly coupled (i.e., Multiprocessor)

- Shared memory
- Each processor capable of doing work on its own
- Easier for the software
- Hardware has to worry about cache coherency, memory contention

DSM

coprocessor

Loosely-coupled (i.e., Multicomputer Network)

- Message passing
- Easier for the hardware
- Programmer's job is tougher

Interconnection networks

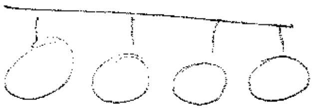
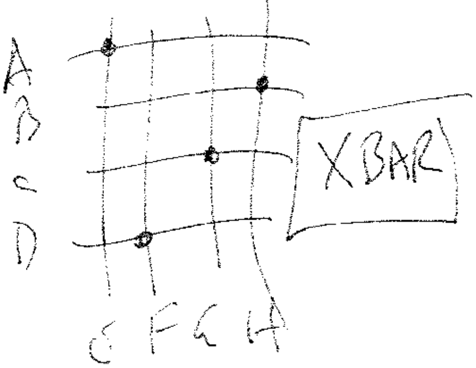
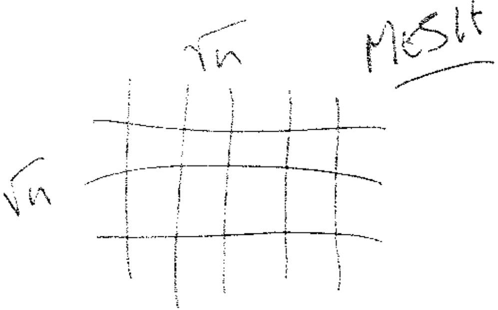
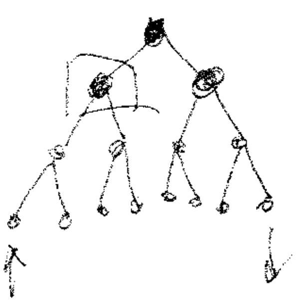
- *Cost*
- *Latency*
- *Contention*

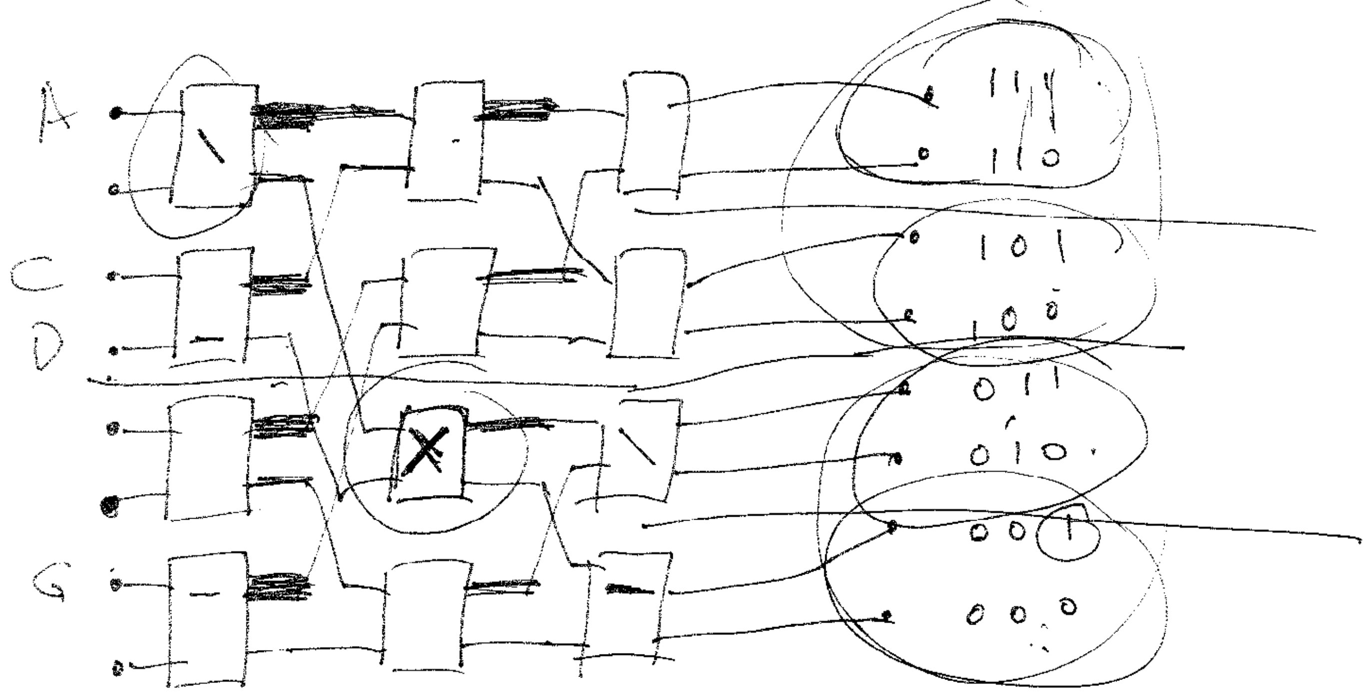
Cache Cohererency

- *Snoopy*
- *Directory*

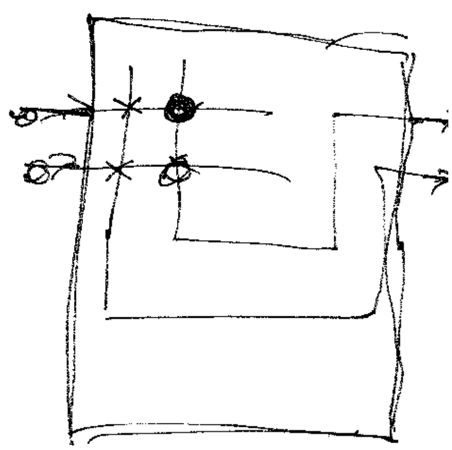
Memory Consistency

- *Sequential Consistency and Mutual Exclusion*

Network	Cost	Latency	Contention
	$O(n)$	$O(1)$	BAD
	$O(n^2)$	$O(1)$	!
	~~IB~~	~~[illegible]~~ ~~$O(\log n)$~~	IB
Ω NETWORK DUNCAN LAWRIE TRAC ~~BANYAN TREE~~	IB	$O(\log n)$	IB
	~~$O(n^2)$~~ $O(n)$ 3n	$O(\sqrt{n})$	
		$O(\log n)$	
		$O(\log n)$	
		$O(n)$	



$n \to n^2$



$A \to 001$

$D \to 010$

Latency → $\log_k n$

64

Cost $\frac{n}{k} \times \log_k n \times k^2$

	L	C
2		
4		
6		
64	1	

11/10