6.895
11/3/03
L16.1

Everyone talk to teaching staff this week re project (if you have not already) - esp. Singapore!

"Ideal" Parallel Computer (Slide 2-3)

Problem: # wires $= \Theta(N^2)$ bad
degree $= \Theta(N)$ bad
diameter $= \Theta(1)$ good

Look at random routing or perm. routing, else hotspot could make any network look bad.

Desire: low-degree networks (Slide 4)
Linear array: $\Theta(N)$ diameter
2D mesh: $\Theta(\sqrt{N})$ diameter
Tree: $\Theta(\lg N)$ diameter

Thm. Bounded degree $\Rightarrow \Omega(\lg N)$ diameter.

| Dist | 0 | 1 | 2 | 3 | k |
|---|---|---|---|---|---|
| # nodes $\le$ | 1 | $d$ | $d(d-1)$ | $d(d-1)^2$ | $\Theta(d^k)$ |

$\Theta(d^k) \ge N \Rightarrow k = \Omega(\log_d N)$. $\boxtimes$

Tree has low diameter, but is it a good routing network? No: congestion.

Def. Minimum bisection width = min # edges that must be removed to partition network in half (to within 1).

BW (tree) = 1
BW (array) = 1
BW (2D mesh) = $\sqrt{N}$
BW (3D mesh) = $N^{2/3}$

Thm. $N$ messages sent at random from $N$ procs.
$E[\text{Routing time}] = \Omega(N/BW + \text{diameter})$

Pf. Expect $\Theta(N)$ messages to cross BW wires. Each wire ships $\le 1$ msg in unit time $\Rightarrow$
Time $\ge \Theta(N)/BW$.
Also Time $\ge$ diameter $\boxtimes$

6.895
11/3/03
L16.2

## Hypercube (Slides 5-6)

Binary rep of node
$\langle b_{d-1}, b_{d-2} \dots b_0 \rangle$
connected to
$\langle \overline{b_{d-1}}, b_{d-2}, \dots, b_0 \rangle$
$\langle b_{d-1}, \overline{b_{d-2}}, \dots, b_0 \rangle$
$\vdots$
$\langle b_{d-1}, b_{d-2}, \dots, \overline{b_0} \rangle$.

Two nodes connected if Hamming distance = 1.
↑ # bit positions in which they differ.

Routing on hypercube.

10111010 → 01101110

Flip any bit that's wrong by routing on that dimension. Bitwise XOR of current msg location and dest. Init: 11010100 → 00000000.

Diameter = $\lg N$ ⟵ Time = $\Omega(\lg N)$
Degree = $\lg N$
BW = $N/2$
$\Theta(N \lg N)$ wires.

## Cube-connected cycles (Slide 7)

$N = n \lg n$ nodes.
Degree = $\Theta(1)$ ($\pm$, depending on whether wires are duplex)
Diameter = $\Theta(\lg N)$
BW = $\Theta(n) = \Theta(N/\lg N)$ since $\lg N = \lg n + \lg\lg n$
$= \Theta(\lg n)$

6.895
11/3/03
L16.3

## Butterfly (FFT) Network (Slide 8-9)

$n$ inputs, $n$ outputs. (Direct network vs. indirect)
$N = n \lg n$ nodes
$\Theta(1)$ degree
Diameter $= \Theta(\lg N)$ «little tricky if not I or O»
$BW = \Theta(n) = \Theta(N/\lg N)$.

Same as CCC, but authors didn't realize!

### Routing on butterfly (Slide 21)

- Just like hypercube, but uses a specific order of dimensions.

$\left\{\begin{array}{l} \text{dest} = 0 \Rightarrow \text{go up} \\ \quad\quad\;\; 1 \Rightarrow \text{go down} \end{array}\right\}$ or $\left\{\begin{array}{l} \text{XOR} = 0 \Rightarrow \text{straight} \\ \quad\quad\;\;\; 1 \Rightarrow \text{cross} \end{array}\right\}$

- CBT rooted at each input (slide 22)
- " " " " output (slide 23).

### Decomposing a butterfly (slides 10-13)

Remove "major cycles" $\Rightarrow$ 2 $n/2$-input butterflies.

Remove "minor" cycles $\Rightarrow$ 2 $n/2$-input butterflies (Slides 14-20)

### Packet routing

source $x_{d-1} x_{d-2} \cdots x_0 \rightarrow$ dest $y_{d-1} y_{d-2} \cdots y_0$

Route major to minor

$x_{d-1} x_{d-2} \cdots x_0$
$y_{d-1} x_{d-2} \cdots x_0$
$y_{d-1} y_{d-2} \cdots x_0$
$\vdots$
$y_{d-1} y_{d-2} \cdots y_0$

$d = \lg n$ steps.

But, might have congestion!

$n$ packet on $n$-input butterfly.
What is worst-case perm?

$\sqrt{n}$ packets at sources $x_1 x_2 x_3 x_4 0000$
go to dests $0000\, x_1 x_2 x_3 x_4$.

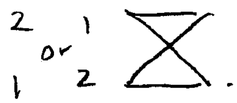
All go through $00000000$ halfway through network $\Rightarrow$ congestion $= \sqrt{n}$.

Beneš network. (slide 24-25)

Thm. Any $n$-perm can be routed (off-line) on an $n$-input Beneš with node-disjoint paths.

Pf. Induction on $n$.

Base. $N=2$.




Ind. case (Slides 26-35) $\boxtimes$

Corollary. An $n$-input Beneš network can simulate any $n$-node, degree-$d$ network in $O(d \lg n)$ time $\boxtimes$.

«But, butterfly is not so bad.»

Theorem. Consider the $N^N$ $N$-packet routing problems on an $N$-node ($n = \Theta(N/\lg N)$-input) butterfly. At least $N^N(1 - 1/N^{\Omega(1)})$ of these problems can be routed in $O(\lg N)$ time.

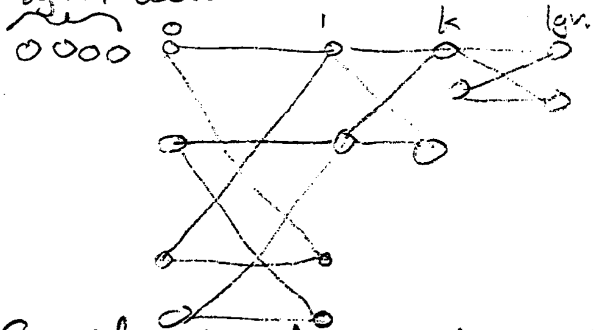
Proof. We'll do a congestion bound only that will lead to an $O(\lg^2 N)$-time result!

6.895
11/3/03
L16.5

WLOG, route packets to output, greedy input to output, straight to correct level.

Phase 1: straight
Phase 2
Phase 3

$\lg n$ packets




Phase 1 takes $O(\lg n)$ time.

Consider level-$k$ node $x$ during Phase 2:
# packets that can reach $x = 2^k \lg n$
(tree property, slide 23)

Prob. that given packet passes through node $x \le 2^{-k}$ (might not be able to reach $x$).

Consider any set of $r$ specific packets.
Prob they all pass through node $x$
$\le (2^{-k})^r = 2^{-kr}$ (independence)

Prob. that $\ge r$ packets pass through node $x$
$$\le \underbrace{\binom{2^k \lg n}{r}}_{\text{\# ways of choosing } r \text{ packets}} \underbrace{2^{-kr}}_{\text{prob they all go through } x}$$

Note: This overcounts. If $r+\Delta$ packets pass through $x$, this event will be counted $\binom{r+\Delta}{r}$ times within the $\binom{2^k \lg n}{r}$ ways.

$$\le \left(\frac{e 2^k \lg n}{r}\right)^r 2^{-kr}$$

$$\boxed{\binom{a}{b} \le \left(\frac{ea}{b}\right)^b \text{ Deathbed}}$$

$$= \left(\frac{e \lg n}{r}\right)^r$$

$$\boxed{\text{Choose } r = 2e\lg n}$$

$$\le \left(\tfrac{1}{2}\right)^{2e\lg N}$$

$$\le N^{-2e}$$

$$\le 1/N^{5.4}$$

Prob. that <u>any</u> node has $\ge 2e\lg n$ packets

$$\le \underbrace{N}_{\#\text{ packets}} \cdot (1/N^{5.4})$$

$$\le N^{-4.4}$$

Bode's inequality:
Prob of union $\le \Sigma$

No indep. needed

$\therefore \ge N^N(1 - 1/N^{4.4})$ problems see $\le 2e\lg N$ congestion.

Hence, each level takes $O(\lg N)$ time $\times \lg N$ levels $= O(\lg^2 N)$ time

Phase 3 also takes $O(\lg N)$ time, since $O(\lg N)$ packets at each output. $\boxtimes$

<u>Corollary</u>.

$$E[\text{routing time}] = O(\lg N)\cdot(1 - 1/N^{4.4}) + O(N)\cdot 1/N^{4.4}$$
$$= O(\lg N). \boxtimes$$