Packet Routing in a Parallel Computer

Communication Network:

- Nodes - processors, switches; unique identifier between 1 and N.
- edges - communication links.
- All communication proceeds in a sequence of synchronous steps.
- Each link can carry a packet, in a step.
- During a step, a processor can send at most one packet to each of its neighbors.

A *permutation routing* problem: each node is the source and destination of exactly one packet.

Assume that each node contains a queue for each edge leaving the node.

A routing algorithm must specify a queueing discipline.

Oblivious routing: The route of a packet depends on its source and destination alone, and not not on the source and destination of any other packet in the system.

Deterministic algorithms: lower bound

Theorem 1. *For any N-node network with maximum degree d, and any deterministic oblivious packet routing algorithm, there is a permutation that requires $\Omega(\sqrt{N/d^3})$ steps.*

The Hypercube

The n-cube: $N = 2^n$ nodes.

Let $x = (x_1, \ldots, x_n)$ be the number of node x in binary.

Nodes x and y are connected by an edge iff their binary representations differ in exactly one bit.

Bit-wise routing: correct bit i in the i-th transition – route has length at most n.

Randomized routing algorithm

Two-phase routing algorithm:

1. Send packet to a randomly chosen destination.
2. Send packet from the random place to real destination.

Path: bit-wise routing.

Any greedy queuing method – if some packet can traverse an edge one does; e.g., FIFO.

Theorem 2. *The two-phase routing algorithm routes an arbitrary permutation on the n-cube in $O(\log N)$ steps with high probability (w.h.p.).*

Proof.

We focus first on phase 1. We bound the routing time of a given packet v_i.

Let $\rho_i = e_1, \ldots, e_k$ be the edges traversed by v_i in phase 1.

The number of steps taken by v_i is equal to the length of ρ_i, which is at most n, plus queueing delay in the intermediate nodes.

Lemma 1. *If a packet leaves the path of another packet it cannot return to that path in the same phase.*

Proof. Leaving the path at the i-th transition implies different i-th bit, this bit cannot be changed again in that phase. $\square$

Lemma 2. *Let S be the set of packets (other than v_i) whose routes pass through at least one of the edges in ρ_i. Then the delay incurred byv_i is at most $|S|$.*

Let the r.v. $H_{ij} = 1$ if ρ_i and ρ_j share at least one edge, and 0 otherwise. The total delay incurred by v_i is at most $\sum_{j=1}^{N} H_{ij}$.

For an edge e in the hypercube, let the r.v. $X(e)$ denote the number of packets that traverse edge e in this phase. Then for route $\rho_i = (e_1, e_2, \ldots, e_k)$,

$$\sum_{j=1}^{N} H_{ij} \leq \sum_{l=1}^{k} X(e_l)$$

For the $N-1$ packets other than v_i let $F_l^j = 1$ iff packet j traversed edge e_l, else $F_l^j = 0$.

$$E[X(e_l)] = \sum_{j=1}^{N-1} \Pr(F_l^i = 1)$$

Since traversing e_l "fixes" the l-th bit, a packet can cross that edge only in its lth transition.

$$\sum_{l=1}^{k} X(e_l) \leq n/2$$

.

Chernoff bound

By the Chernoff bound, it follows that

$$\Pr(\sum_{j=1}^{N} H_{ij} \geq 3n) \leq 2^{-3n} = 1/N^3$$

Thus with probability at least $1 - 1/N^2$, every packet finishes phase 1 in $4n$ or fewer steps.

Phase 2 is by symmetry.

Every packet reaches its destination in $8n$ steps w.h.p.

□