Class 5

Chernoff Bounds! MGFs!

Randomized Routing on the Hypercube!

1 Warm-Up

Suppose you roll a 6-sided coin n times. Use a Chernoff bound to bound the probability that you see more than $\frac{1+\delta}{6} \cdot n$ threes, where $\delta \in (0, 1)$. What bound do you get as a function of n?

Announcements

- HW2 due Friday!
- HW3 out now if you want to get an early start!

- Add/drop deadline Friday!

Recap of minilectures

- Moment-generating functions: $M_X(t) = E[\ e^{tX}\ ]$
- Chernoff bounds! Say that $X = \sum_{i=1}^{n} X_i$ where $X_i \in \{0,1\}$ i.i.d.
  - Multiplicative form: $\Pr\left[\ X \geq (1+\delta)E[X]\right] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{E[X]}$
  - Simplification: $\Pr\left[\ X \geq (1+\delta)E[X]\right] \leq \exp\left(-\frac{\delta^2 E[X]}{3}\right)$ for $\delta \in (0,1]$

    or: $\Pr\left[\ X \geq c \cdot E[X]\right] \leq \exp(-cE[X])$ for $c \geq 6$

  …
  - Alternatively (follows from Hoeffding): $\Pr[X \geq E[X] + t] \leq \exp\left(-\frac{2t^2}{n}\right)$

Question: Is there ever a situation where one Chernoff bound is quantitatively (asymptotically) better than another?

Which Chernoff bound to use?

- Lots of the time you get the same asymptotic answer.
- Not always: e.g., If E[X] is small, to bound $\Pr\left[\, X \geq (1+\delta)E[X]\right]$….

$$\Pr\left[\, X \geq (1+\delta)E[X]\right] \leq \exp\left(-\frac{\delta^2 E[X]}{3}\right)$$

$$\Pr[X \geq E[X] + t] \leq \exp\left(-\frac{2t^2}{n}\right)$$

$$\Pr[X \geq E[X] + \delta E[X]] \leq \exp\left(-\frac{2\delta^2 E[X]^2}{n}\right)$$

Which Chernoff bound to use?

- Lots of the time you get the same asymptotic answer.
- Not always: e.g., If E[X] is small, to bound $\Pr\left[\, X \geq (1+\delta)E[X]\right]$....

$$\Pr\left[\, X \geq (1+\delta)E[X]\right] \leq \exp\left(-\frac{\delta^2 E[X]}{3}\right)$$

If $E[X] \ll n$, this is better

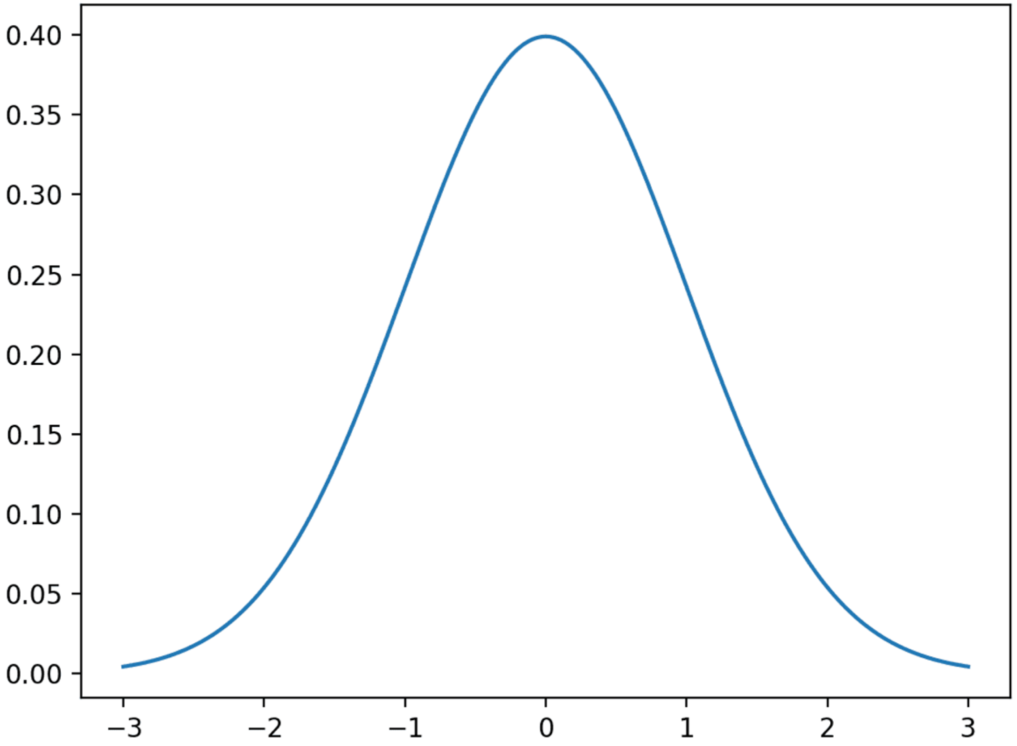
$$\Pr[X \geq E[X] + t] \leq \exp\left(-\frac{2t^2}{n}\right)$$

$$\Pr[X \geq E[X] + \delta E[X]] \leq \exp\left(-\frac{2\delta^2 E[X]^2}{n}\right) = \exp\left(-2\delta^2 E[X]\left(\frac{E[X]}{n}\right)\right)$$

- Can't hurt to try them all until you get the hang of it...

Recap of mini-lectures

- The main point: Sums of (enough, reasonable) independent random variables concentrate like Gaussians!



Questions?

1 Warm-Up

Suppose you roll a 6-sided coin n times. Use a Chernoff bound to bound the probability that you see more than $\frac{1+\delta}{6} \cdot n$ threes, where $\delta \in (0, 1)$. What bound do you get as a function of n?

Suppose that $0 < \delta < 1$. Suppose you roll a 6-sided die n times. What bound do you get on the probability that you see more than $(1+\delta)n/6$ threes? (For the asymptotic notation, think of $\delta \to 0$ and $n \to \infty$.)

0

$\leq \delta$

$\leq O(\delta^2/n)$

$\leq \exp(-\Omega(\delta^2/n))$

$\leq \exp(-\Omega(\delta^2 n))$

Suppose that $0 < \delta < 1$. Suppose you roll a 6-sided die n times. What bound do you get on the probability that you see more than $(1+\delta)n/6$ threes? (For the asymptotic notation, think of $\delta \to 0$ and $n \to \infty$.)

✓✓ 0

$\leq \delta$	0
$\leq O(\delta^2/n)$	0
$\leq \exp(-\Omega(\delta^2/n))$	0
$\leq \exp(-\Omega(\delta^2 n))$	0

Suppose that $0 < \delta < 1$. Suppose you roll a 6-sided die n times. What bound do you get on the probability that you see more than $(1+\delta)n/6$ threes? (For the asymptotic notation, think of $\delta \to 0$ and $n \to \infty$.)

✓ 0

$\leq \delta$

0

$\leq O(\delta^2/n)$

0

$\leq \exp(-\Omega(\delta^2/n))$

0

$\leq \exp(-\Omega(\delta^2 n))$

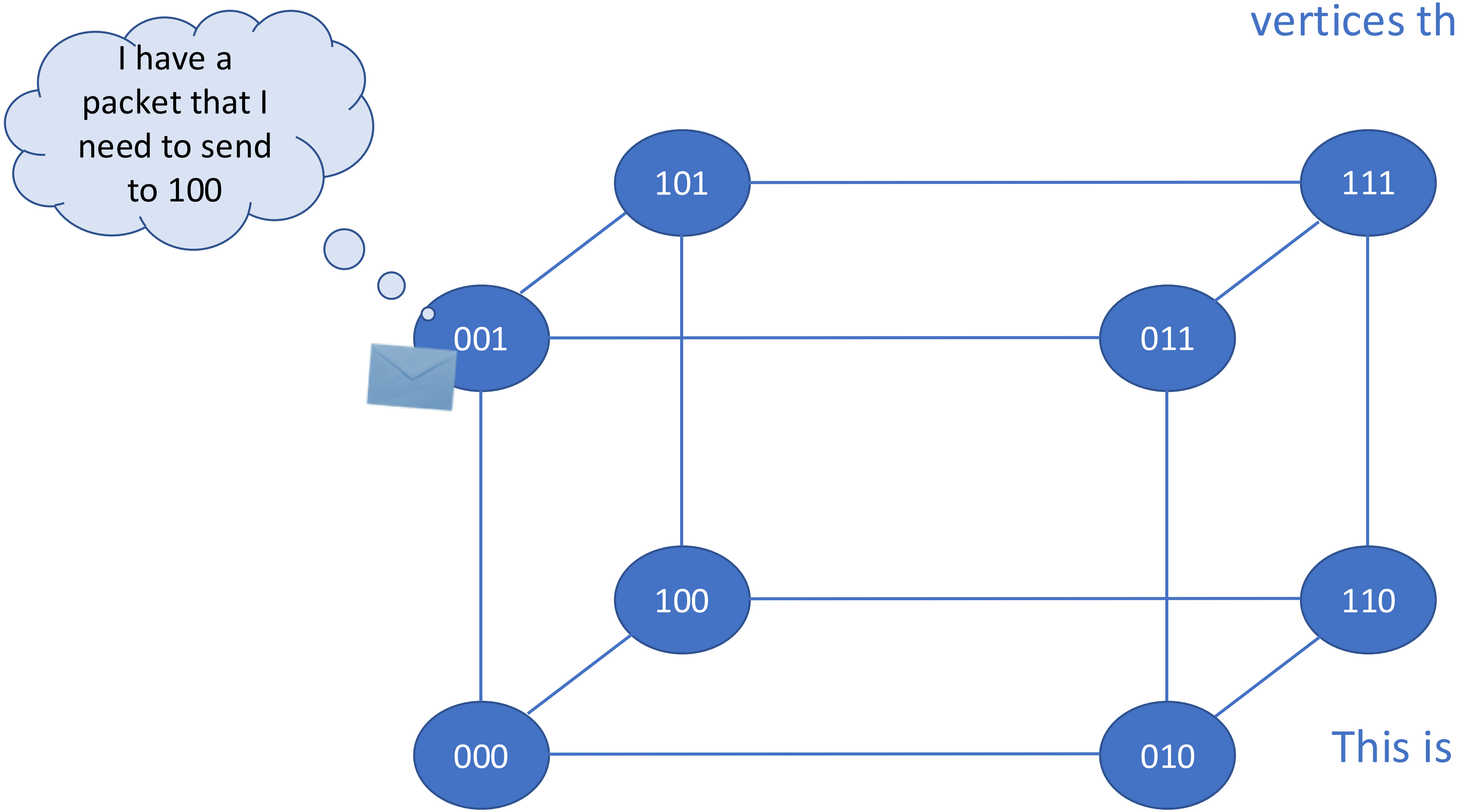
0

Randomized Routing

on the hypercube

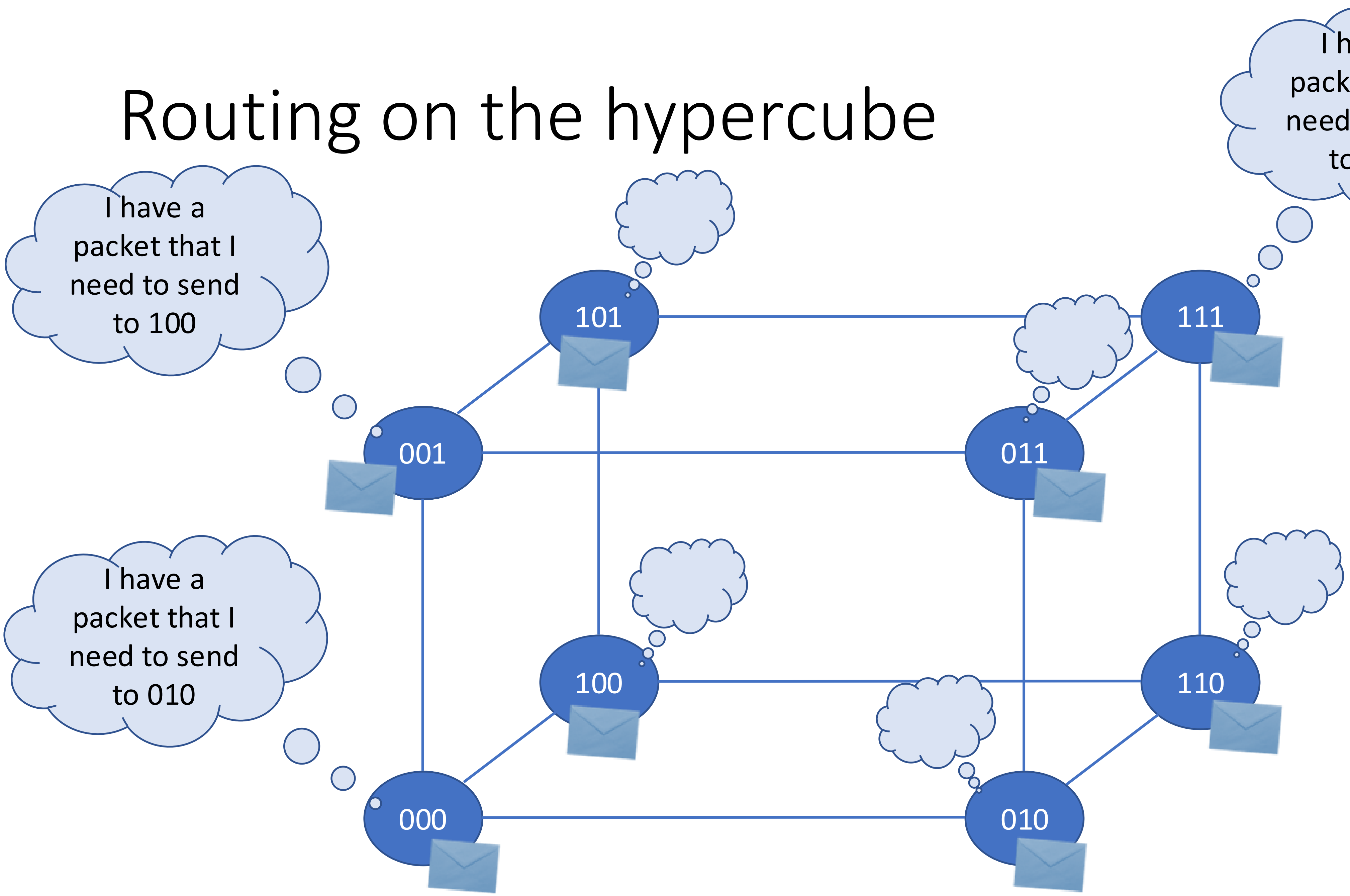
Routing on the hypercube

Definition: The n-dimensional hypercube is a graph with vertices $\{0,1\}^n$ and edges between any two vertices that differ by a single bit.



This is the 3-dimensional hypercube.

Routing on the hypercube



I have a packet that I need to send to 000

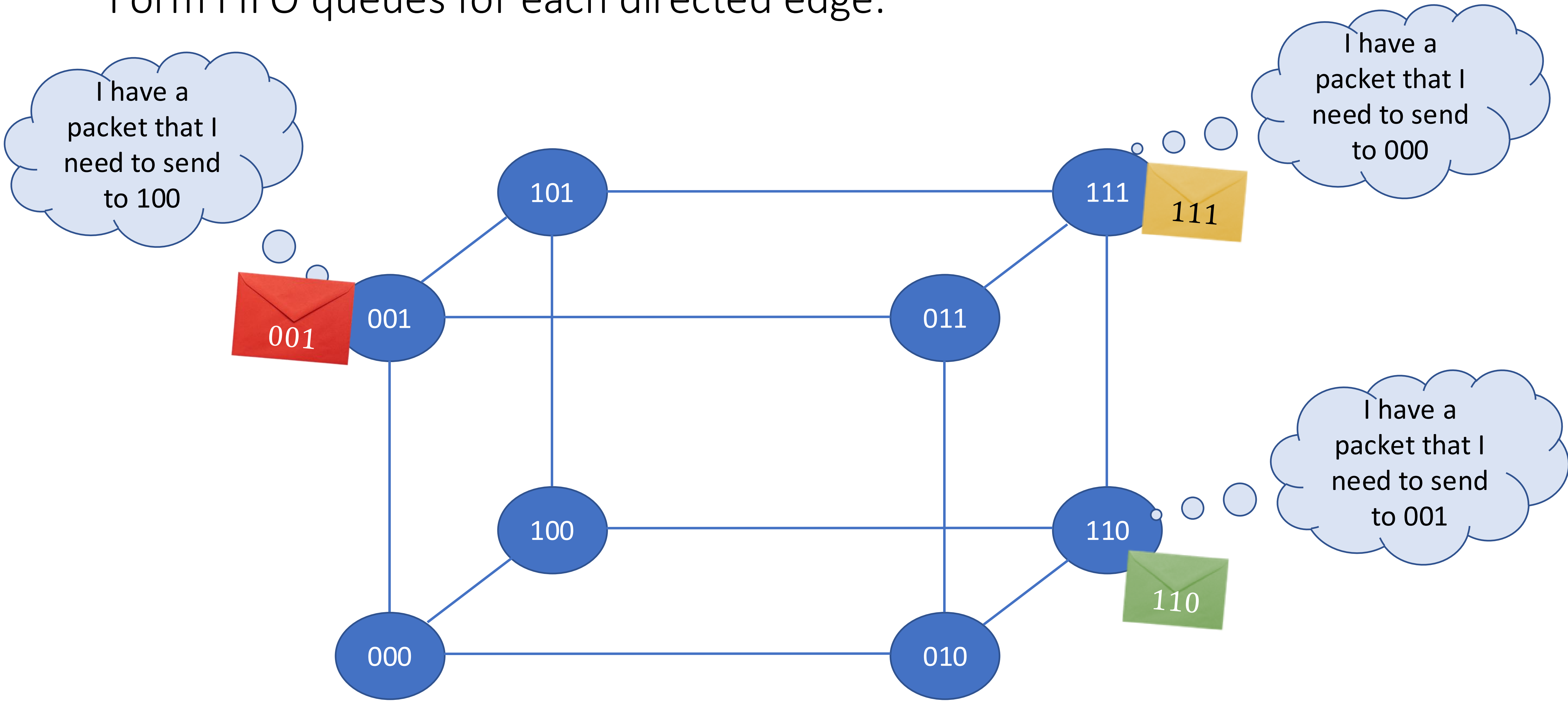
I have a packet that I need to send to 100

I have a packet that I need to send to 010

Say each node i wants to send a packet to node $\pi(i)$, where $\pi : \{0,1\}^n \rightarrow \{0,1\}^n$ is a permutation.

Only one packet on an edge at a time (in each direction)

Form FIFO queues for each directed edge.



I have a packet that I need to send to 100

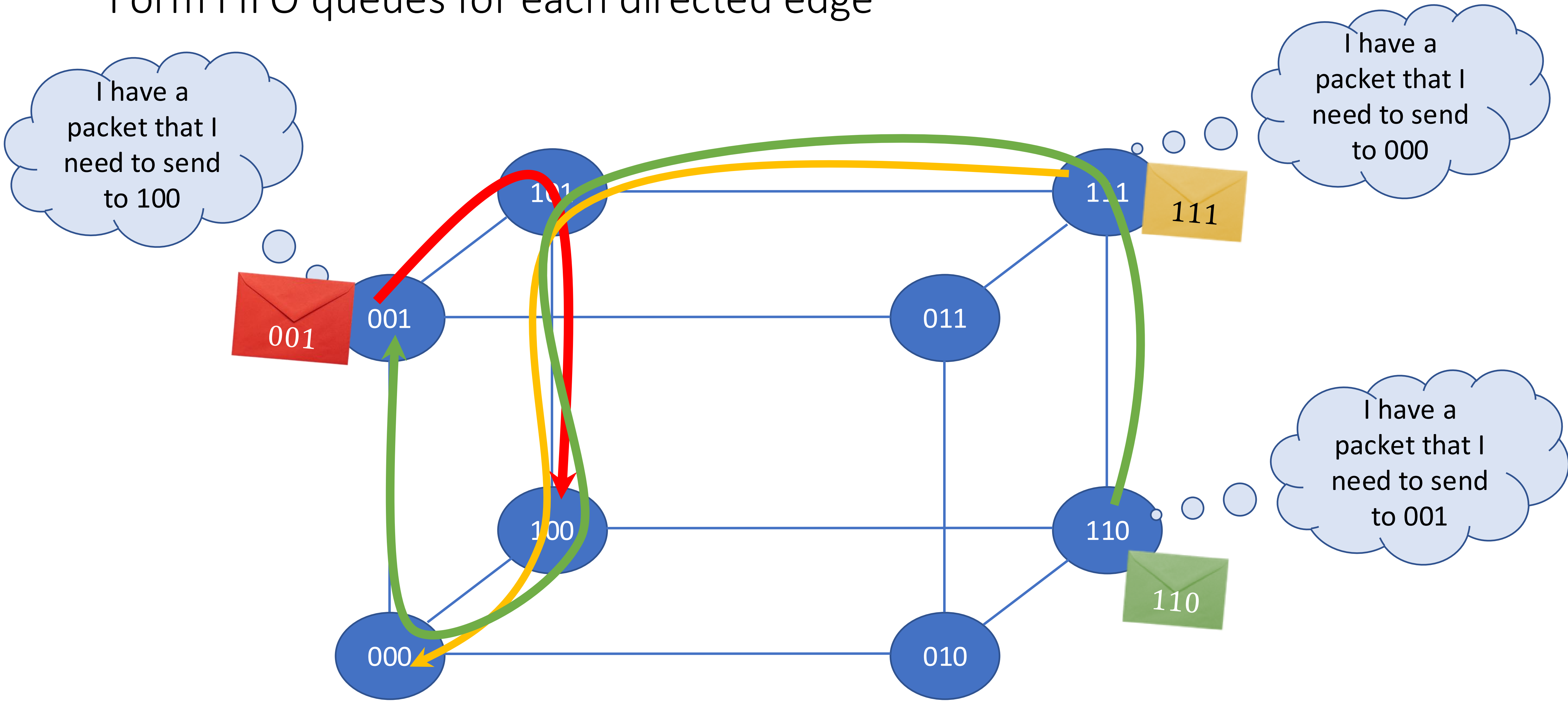
I have a packet that I need to send to 000

I have a packet that I need to send to 001

t=0

Only one packet on an edge at a time (in each direction)

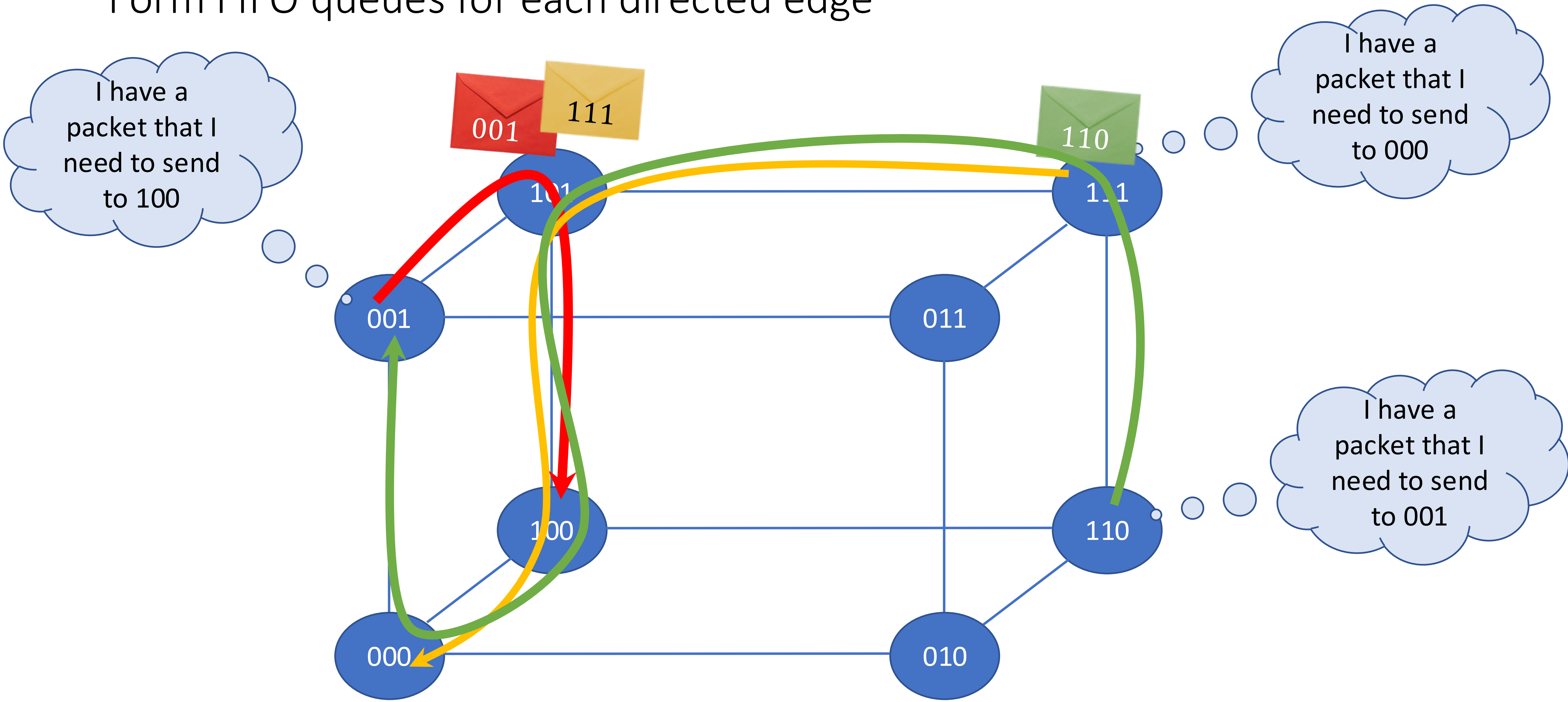
Form FIFO queues for each directed edge



t=1

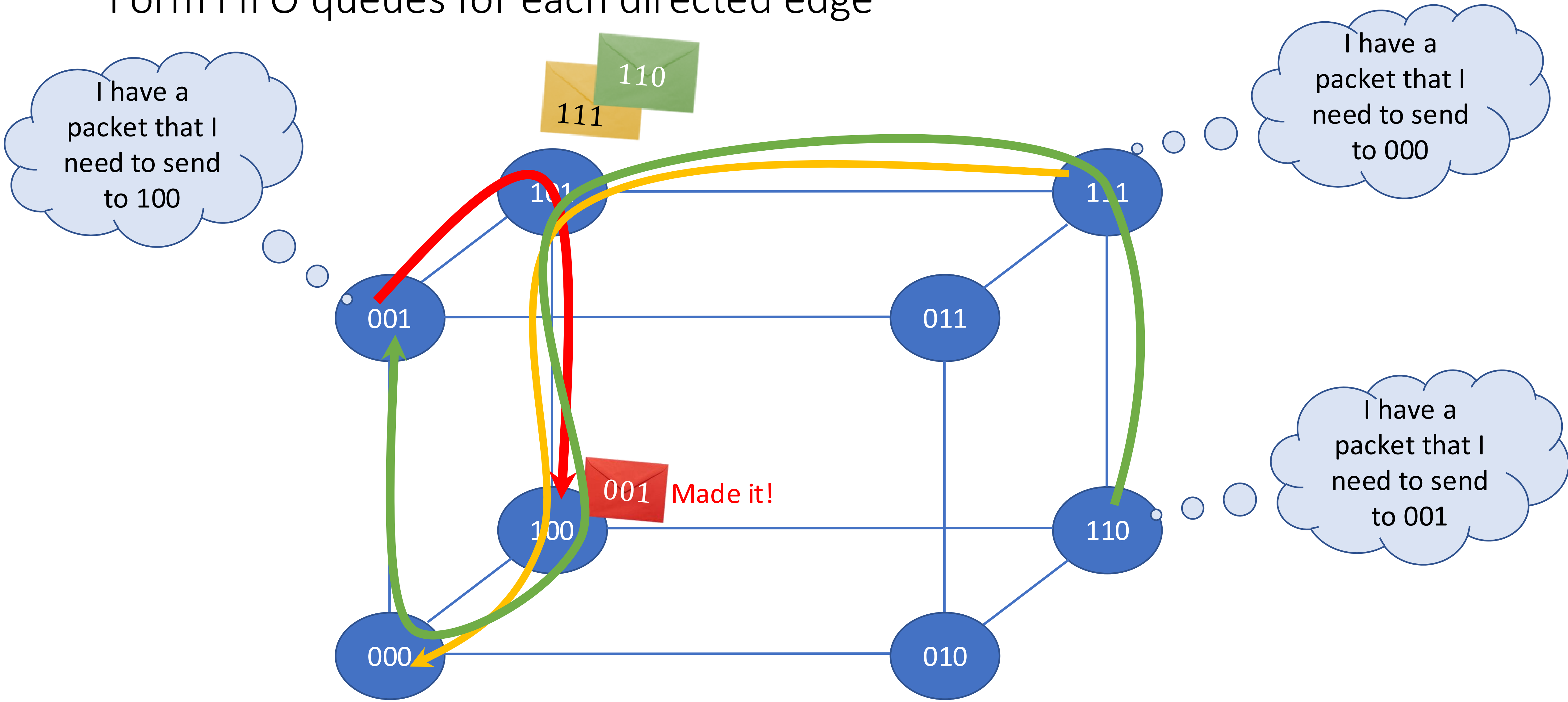
Only one packet on an edge at a time (in each direction)

Form FIFO queues for each directed edge



Only one packet on an edge at a time (in each direction)

Form FIFO queues for each directed edge



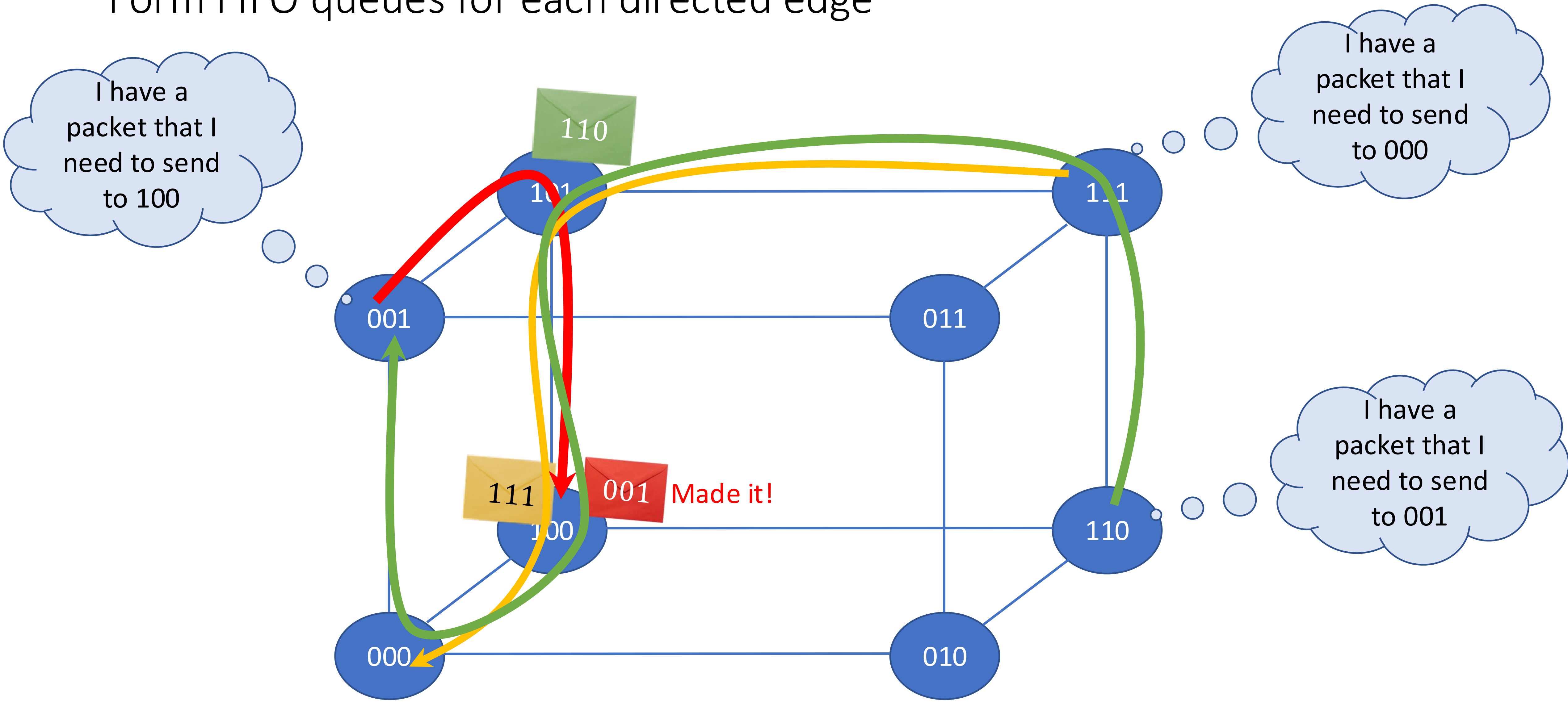
I have a packet that I need to send to 100

I have a packet that I need to send to 000

I have a packet that I need to send to 001

Only one packet on an edge at a time (in each direction)

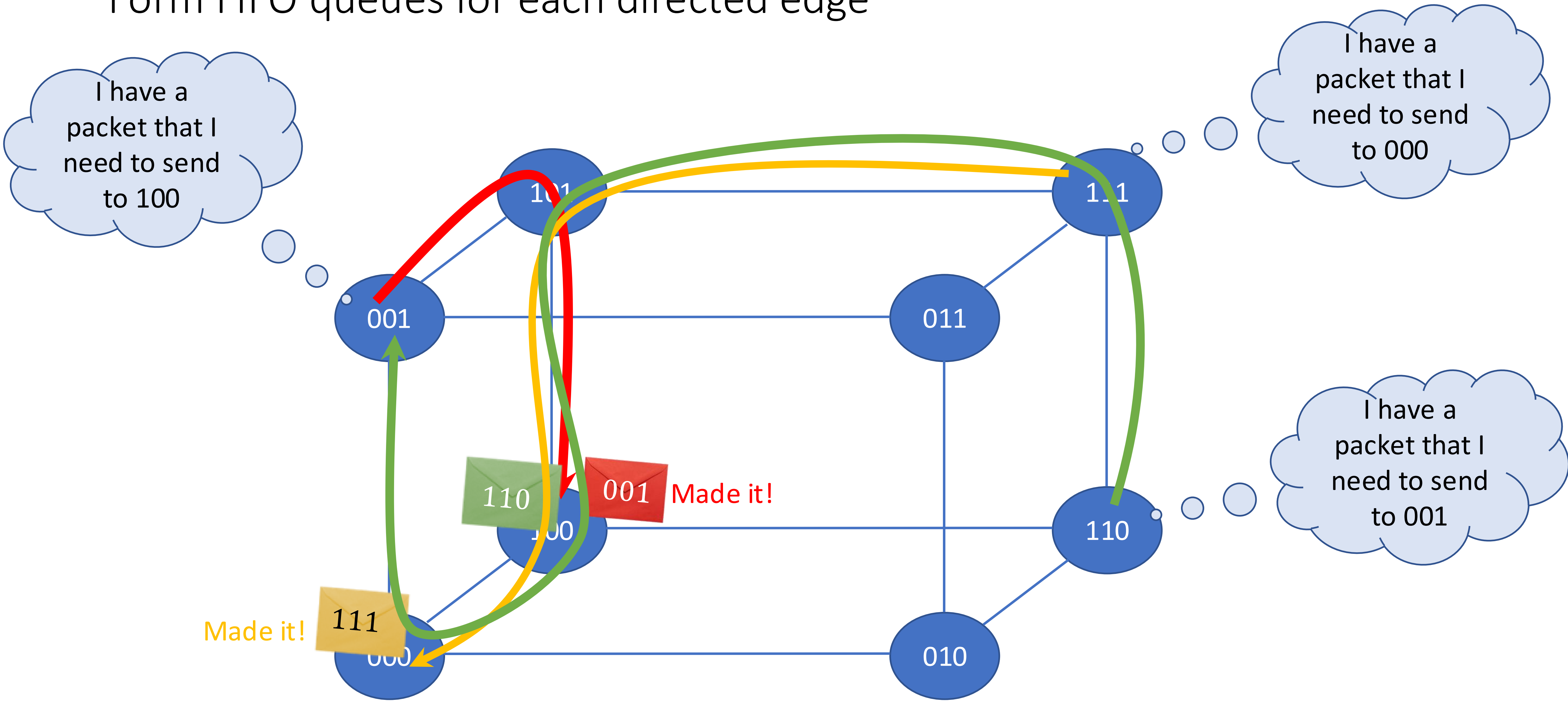
Form FIFO queues for each directed edge



t=4

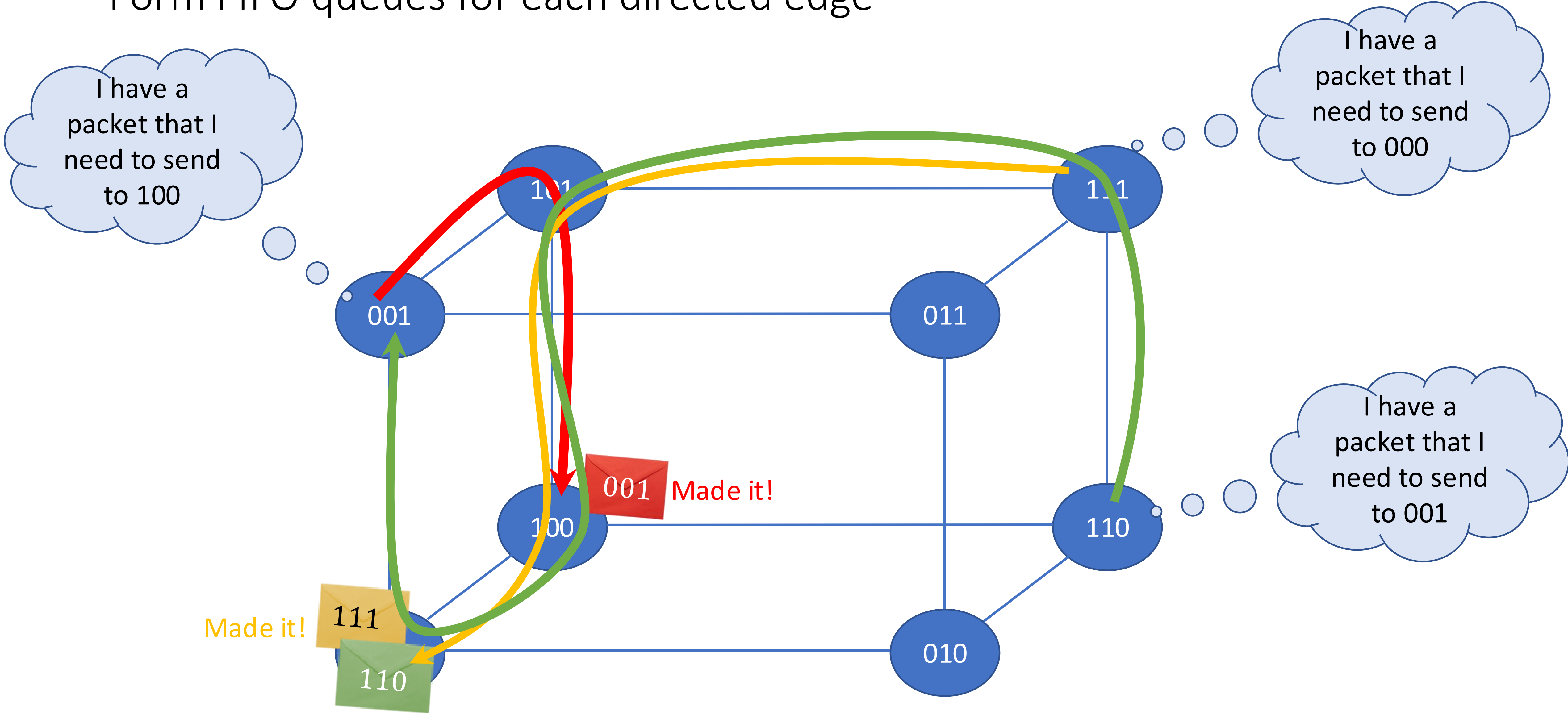
Only one packet on an edge at a time (in each direction)

Form FIFO queues for each directed edge



Only one packet on an edge at a time (in each direction)

Form FIFO queues for each directed edge



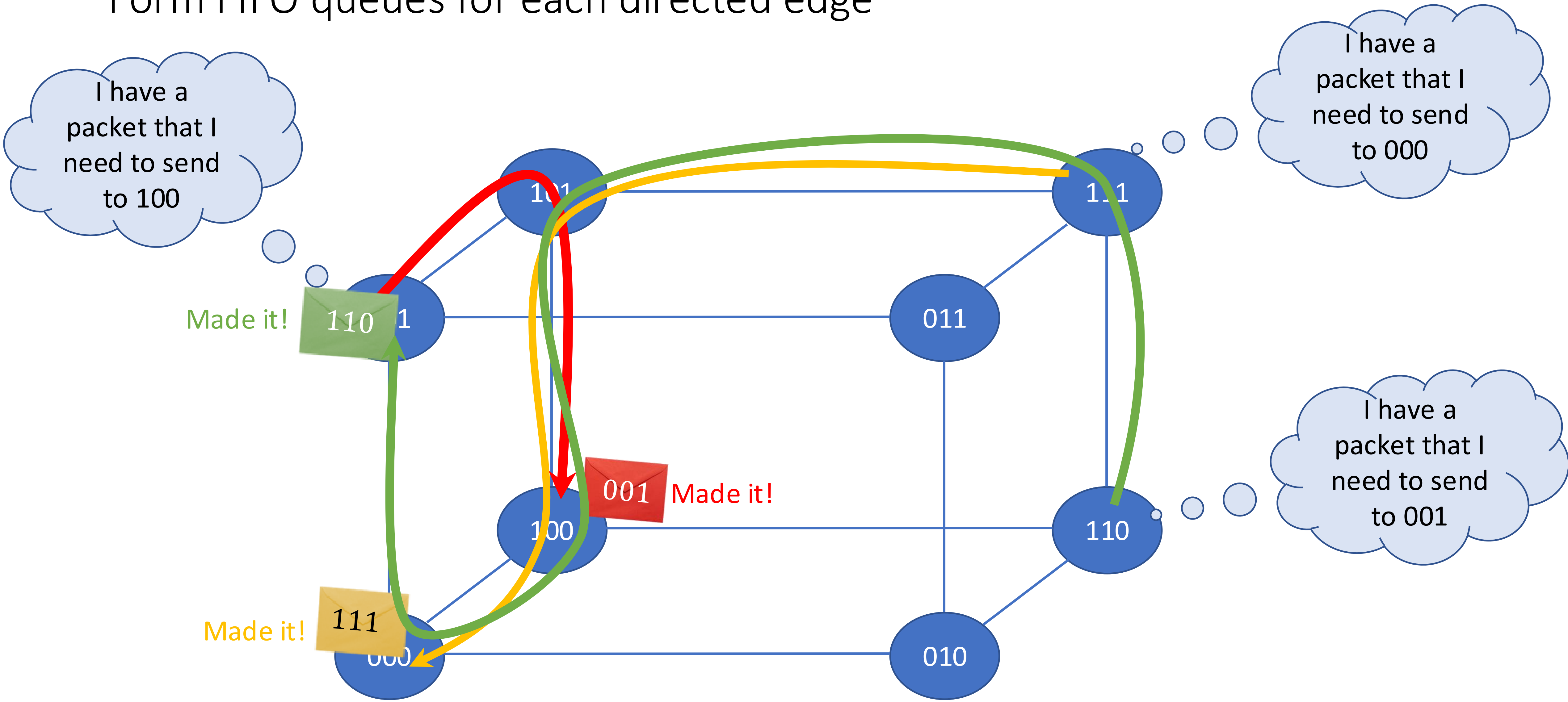
I have a packet that I need to send to 100

I have a packet that I need to send to 000

I have a packet that I need to send to 001

Only one packet on an edge at a time (in each direction)

Form FIFO queues for each directed edge



I have a packet that I need to send to 100

I have a packet that I need to send to 000

I have a packet that I need to send to 001

Goal for today

- Minimize number of timesteps needed to route all 2^n packets in the n-dimensional hypercube.
  - Can't do better than n steps, since maybe 0000000 wants to go to 1111111.
  - Can we do $O(n)$?
- Subject to some assumptions:
  - Assume that π is worst-case.
  - Assume none of the packets/nodes needs to know anything about any of the other packets/nodes.
  - Definition: A routing scheme is **oblivious** if the route that packet i takes to node $\pi(i)$ only depends on i and $\pi(i)$.
- Fun exercise: Any *deterministic*, oblivious routing scheme requires at least $2^{n/2}/n$ steps for some permutation π. ☹
- We will get a *randomized* oblivious routing scheme that runs in $O(n)$ steps in the worst case! ☺

Bit-fixing path

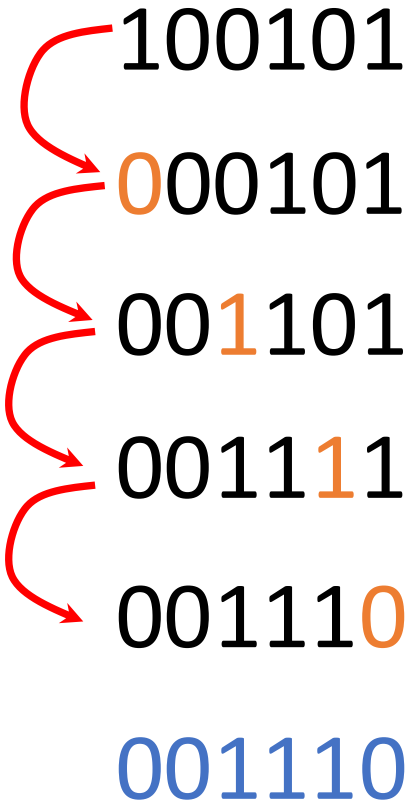
- Definition by example: Suppose you want to send 100101 to 001110.

100101

001110

Bit-fixing path

- Definition by example: Suppose you want to send 100101 to 001110.



100101

000101

001101

001111

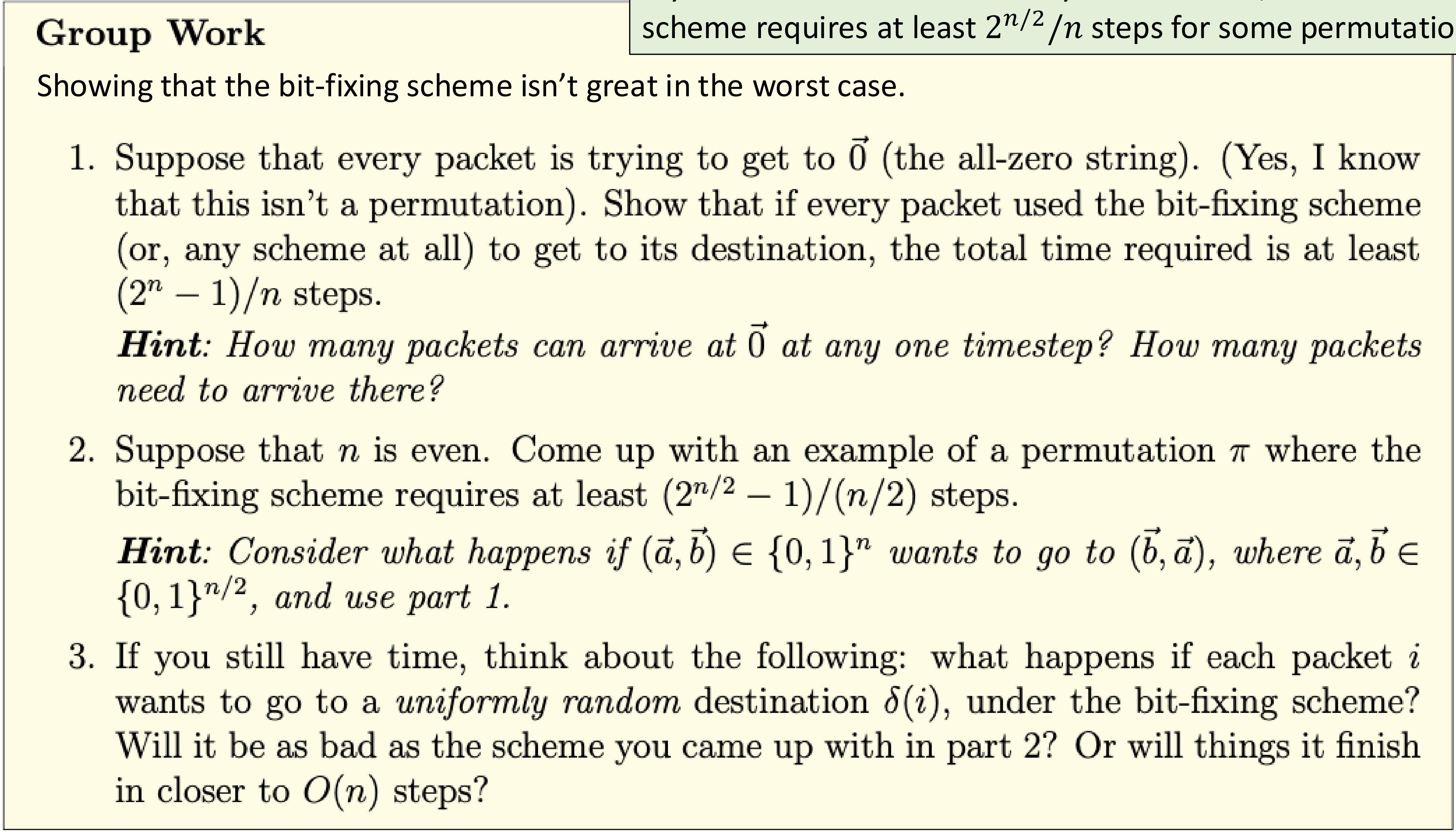
001110

001110

Is the bit-fixing scheme a good routing scheme?

- Turns out, **NO.**
- You'll prove that in the next group work!
- **Spoiler alert:** it will be useful in designing a good routing scheme!

If you have time: Show that any *deterministic*, oblivious routing scheme requires at least $2^{n/2}/n$ steps for some permutation π.



Group Work

Showing that the bit-fixing scheme isn’t great in the worst case.

1. Suppose that every packet is trying to get to $\vec{0}$ (the all-zero string). (Yes, I know that this isn’t a permutation). Show that if every packet used the bit-fixing scheme (or, any scheme at all) to get to its destination, the total time required is at least $(2^n - 1)/n$ steps.

   Hint: *How many packets can arrive at $\vec{0}$ at any one timestep? How many packets need to arrive there?*

2. Suppose that n is even. Come up with an example of a permutation π where the bit-fixing scheme requires at least $(2^{n/2} - 1)/(n/2)$ steps.

   Hint: *Consider what happens if $(\vec{a}, \vec{b}) \in \{0,1\}^n$ wants to go to $(\vec{b}, \vec{a})$, where $\vec{a}, \vec{b} \in \{0,1\}^{n/2}$, and use part 1.*

3. If you still have time, think about the following: what happens if each packet i wants to go to a *uniformly random* destination $\delta(i)$, under the bit-fixing scheme? Will it be as bad as the scheme you came up with in part 2? Or will things it finish in closer to $O(n)$ steps?

Solutions to group work

1. At each step, at most n packets can get to 0000000. There are $2^n - 1$ packets that need to get there, so it takes at least $\frac{2^n-1}{n}$ steps.

Solutions to group work

1. At each step, at most n packets can get to 0000000. There are $2^n - 1$ packets that need to get there, so it takes at least $\frac{2^n-1}{n}$ steps.
2. Consider the permutation $\pi: (\vec{a}, \vec{b}) \mapsto (\vec{b}, \vec{a})$.
   - The bit-fixing map will route $(\vec{a}, \vec{0}) \mapsto (\vec{0}, \vec{a})$ by first routing $\vec{a}$ to $\vec{0}$, for any $\vec{a}$.
   - Apply part 1.

- Takeaway: the bit-fixing scheme is not great in the worst case.
- But what about if the destinations are random?...

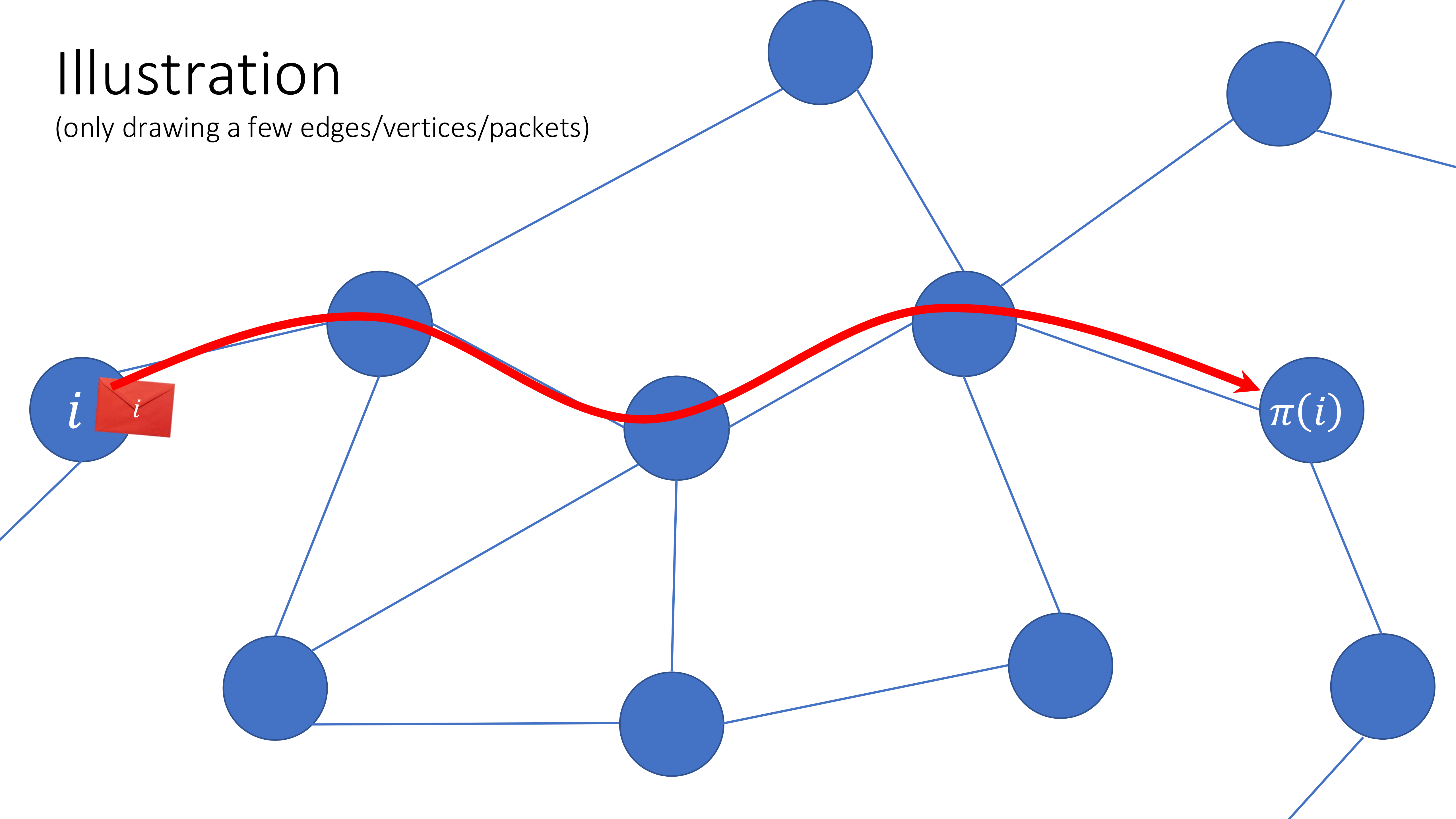
Lemma

- Let $D(i)$ denote the delay for packet i.
    - Number of timesteps that i spends waiting.
- Let $P(i)$ denote the path taken by packet i.
    - $P(i)$ is just a set of directed edges.
- Let $N(i)$ be the number of other packets j so that $|P(j) \cap P(i)| \neq 0$
    - Number of other packets who ever want to use any directed edge in i's path.

- Then $D(i) \leq N(i)$.

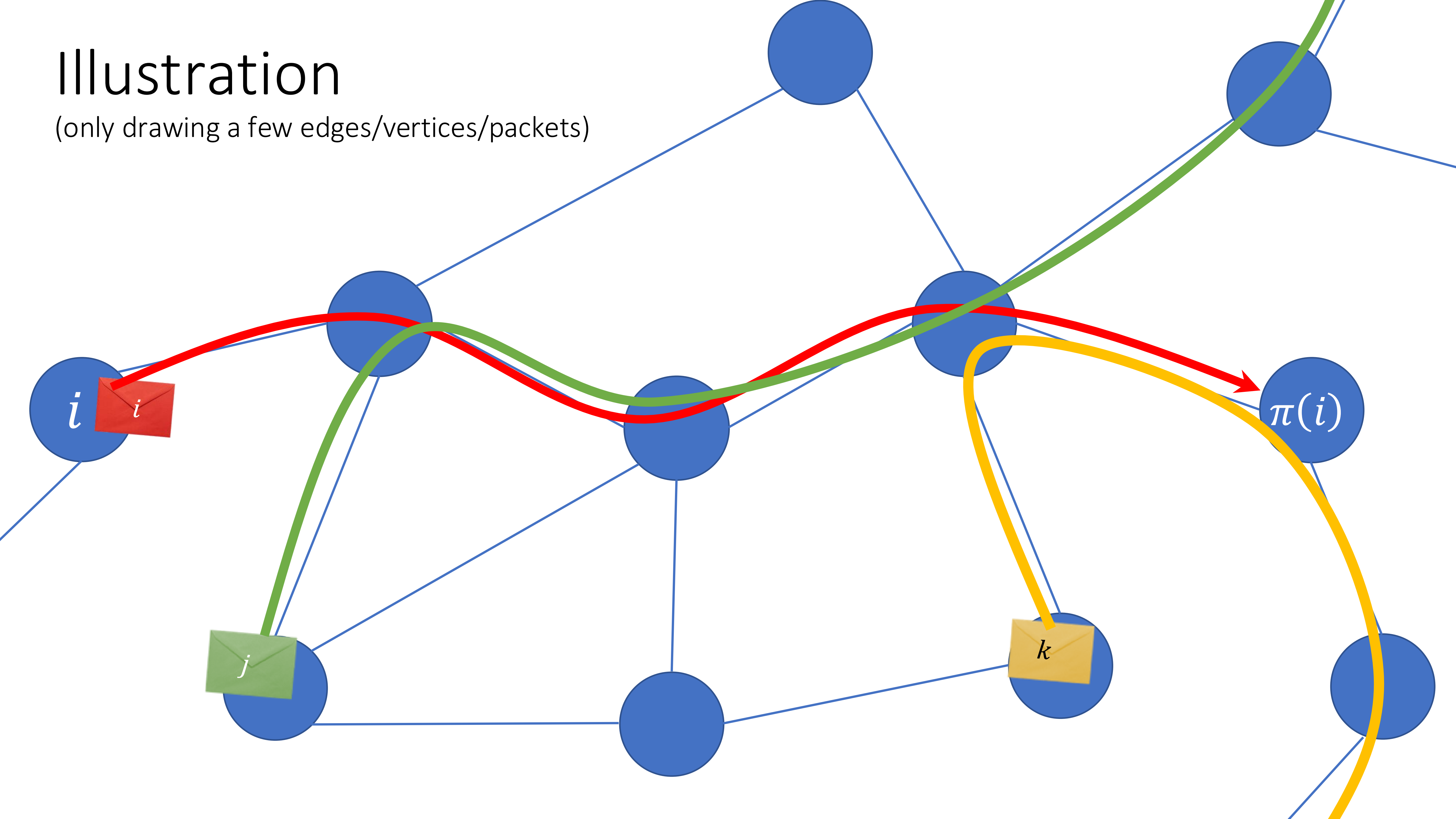
Illustration

(only drawing a few edges/vertices/packets)



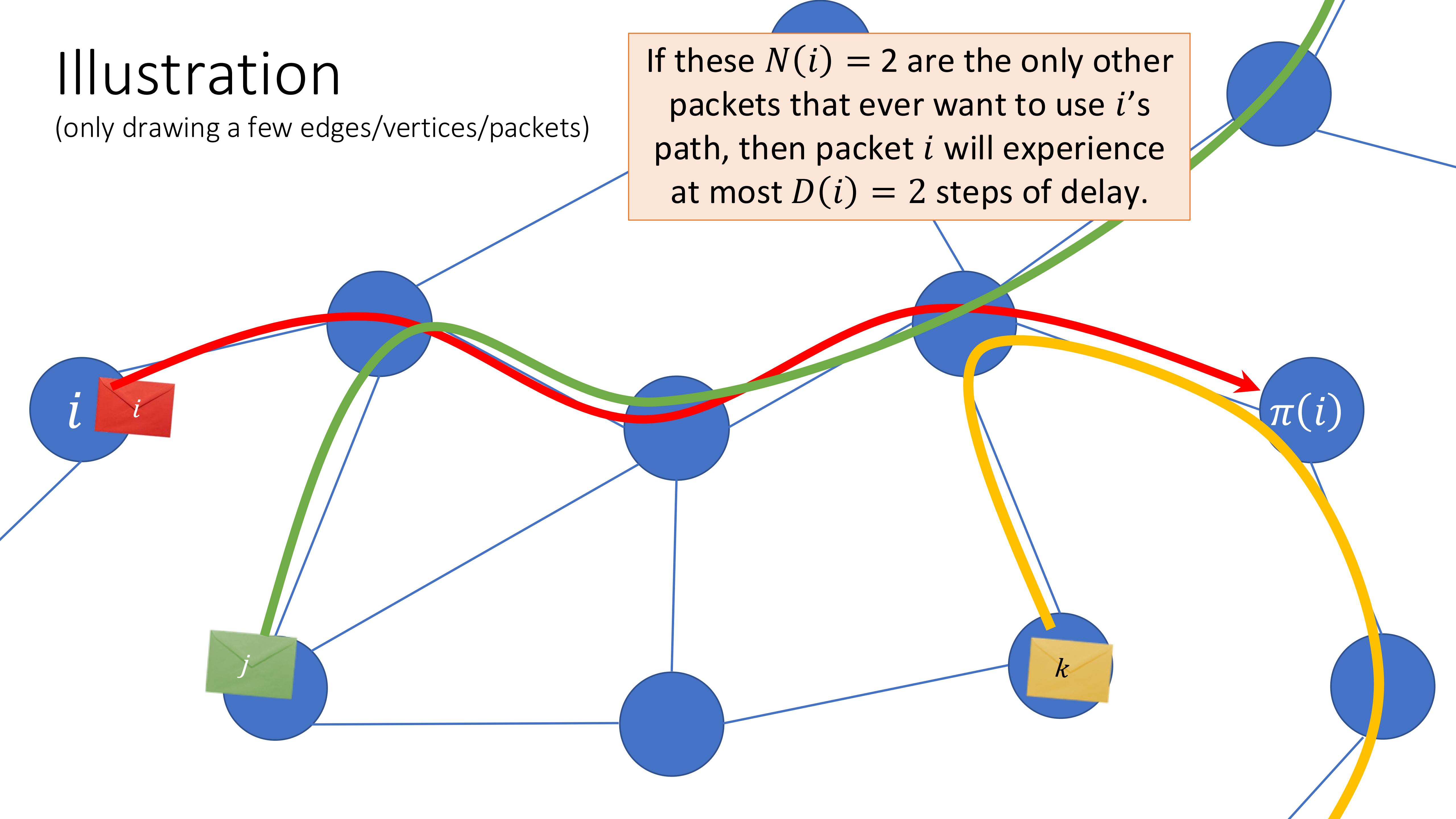
Illustration

(only drawing a few edges/vertices/packets)



Illustration

(only drawing a few edges/vertices/packets)



If these $N(i) = 2$ are the only other packets that ever want to use i's path, then packet i will experience at most $D(i) = 2$ steps of delay.

Lemma

See the lecture notes for the proof, or try to prove it yourself!

- Let $D(i)$ denote the delay for packet i.
  - Number of timesteps that i spends waiting.
- Let $P(i)$ denote the path taken by packet i.
  - $P(i)$ is just a set of directed edges.
- Let $N(i)$ be the number of other packets j so that $|P(j) \cap P(i)| \neq 0$
  - Number of other packets who ever want to use i's path.

- Then $D(i) \leq N(i)$.

Strategy

- Let $\delta: \{0,1\}^n \to \{0,1\}^n$ be a *random* map.
  - Not necessarily a permutation

- Show that, with high probability over δ, $N(i)$ (number of other packets intersecting i's path) is small when trying to route each j to $\delta(j)$.

- Conclude that $D(i)$ (the delay for packet i) is small with high probability.

- Use this to design a randomized routing algorithm for a worst-case π ...

Group Work!

Coming up with a good routing algorithm

See agenda (handout or online) for more details about the questions and outlines/hints!!

Let $\delta: \{0,1\}^n \rightarrow \{0,1\}^n$ be a random map.

1. Fix a special i and let X_j be 1 if $P(i)$ intersects $P(j)$. Show $E[\sum_{j \neq i} X_j] \leq \frac{n}{2}$.
2. $P\left[\sum_j X_j > 10n\right] \leq$ ________?
3. Bound the probability that bit-fixing takes $> 11n$ timesteps to route all of the i’s to their destination $\delta(i)$.
4. Come up with a routing scheme for a worst-case permutation, that takes at most $22n$ steps whp.

Final Algorithm

1. Route packet i to a completely random node, $\delta(i)$ using the bit-fixing algorithm.
2. Route packet i from $\delta(i)$ to $\pi(i)$ using the bit-fixing algorithm.

So even though the bit-fixing algorithm is bad in the worst case, it's good for random destinations! We can use that to turn it into an algorithm that's good in the worst case (with high probability).

Group Work Solutions

1. Fix a special i and let X_j be 1 if $P(i)$ intersects $P(j)$. Show $E[\sum_{j \neq i} X_j] \leq \frac{n}{2}$.

Following the outline..

$$\mathbb{E}\left[\sum_{j \neq i} (\text{number of edges in } P(j))\right] = \sum_{j \neq i} \mathbb{E}(\text{number of edges in } P(j))$$

$$= \sum_{j \neq i} \frac{n}{2}$$

The number of edges in $P(j)$ is the number of bits that j and $\delta(j)$ differ on. In expectation, n/2.

$$\leq 2^n \cdot \frac{n}{2}$$

Group Work Solutions

1. Fix a special i and let X_j be 1 if $P(i)$ intersects $P(j)$. Show $E[\sum_{j\neq i} X_j] \leq \frac{n}{2}$.

Following the outline...

$$\mathbb{E}\left[\sum_{j\neq i}(\text{number of edges in } P(j))\right] \leq 2^n \cdot \frac{n}{2}$$

$$\mathbb{E}\left[\sum_{e}[\text{ number of P(j) that e is in}]\right] \leq 2^n \cdot \frac{n}{2}$$

$$\sum_{e}\mathbb{E}[\text{ number of P(j) that e is in}] \leq 2^n \cdot \frac{n}{2}$$

$$(\#\text{edges})\cdot\ \mathbb{E}[\text{ number of P(j) that e is in}] \leq 2^n \cdot \frac{n}{2}$$

Both are equal to $\mathbb{E}[\sum_j \sum_e \mathbf{1}[e \text{ is in } P(j)]]$

For a fixed edge *e*,

$$\mathbb{E}[\text{ number of P(j) that e is in}]$$

$$\leq \frac{2^n \cdot \frac{n}{2}}{\#\text{edges}} = \frac{1}{2}$$

Group Work Solutions

1. Fix a special i and let X_j be 1 if $P(i)$ intersects $P(j)$. Show $E[\sum_{j \neq i} X_j] \leq \frac{n}{2}$.

Following the outline…

For a fixed edge *e*,

$$\mathbb{E}[\text{ number of P(j) that e is in}] \leq \frac{1}{2}$$

$$\mathbb{E}\left[\sum_{j \neq i} X_j\right] \leq \mathbb{E}\left[\sum_{e \in P(i)} \sum_{j \neq i} \mathbf{1}[e \in P(j)]\right]$$

$$\leq \mathbb{E}\left[\sum_{e \in P(i)} (\#\text{ paths } P(j) \text{that } e \text{ is in})\right] \leq \sum_{e \in P(i)} \frac{1}{2} \leq \frac{n}{2}$$

Group Work Solutions: Problem 2

Use a Chernoff bound to bound the probability that $\sum_{j \neq i} X_j \geq 10n$

$$\mathbb{E}\left[\sum_{j \neq i} X_j\right] \leq \frac{n}{2}$$

$$\Pr\left[\sum_j X_j \geq 10n\right] = \Pr\left[\sum_j X_j \geq 20 \cdot \frac{n}{2}\right]$$

$$\leq \exp\left(-20 \cdot \frac{n}{2}\right) = \exp(-10\, n)$$

X_j is 1 if $P(i)$ intersects $P(j)$.

Group Work Solutions: Problem 3

Show that everyone gets where they are going in 11n steps, whp.

$$\Pr[\, i \text{ delayed by} \geq 10n \text{ steps} \,] \leq \Pr\left[\sum_{j \neq i} X_j \geq 10\, n \right]$$

By the lemma!

$D(i) \leq N(i)$

Delay for i

Number of j so that P(j) intersects P(i)

Aka $\sum_j X_j$

$$\leq \exp(-10\, n)$$

Union bound over all choices for i:

$$\Pr[\, \exists i \text{ delayed by} \geq 10n \text{ steps} \,] \leq 2^n \exp(-10\, n) \leq \exp(-9n)$$

If that doesn't happen, then everyone gets to where they are going in 11n steps. (at most n actually moving, and at most 10n waiting).

Group Work Solutions: Problem 4

Find a good algorithm!

1. Route packet i to a completely random node, $\delta(i)$ using the bit-fixing algorithm.
2. Route packet i from $\delta(i)$ to $\pi(i)$ using the bit-fixing algorithm.

Pr[this takes > 22n steps]

$\leq$ Pr[Step 1 takes > 11n steps] + Pr[Step 2 takes > 11n steps]

$\leq \exp(-9n) + \exp(-9n) = tiny.$

Recap

- We saw a randomized routing algorithm for the hypercube that runs in time O(n).

- Chernoff bounds are useful!
  - (Although it can often be non-trivial to figure out how/when to apply them!)

Next Time: Balls and Bins!