Demand-Aware Network Design with Minimal Congestion and Route Lengths

Chen Avin , Kaushik Mondal, Stefan Schmid

INFOCOM 2019

Motivation

Motivation

- Traditionally, networks topologies designs are demand oblivious:

Motivation

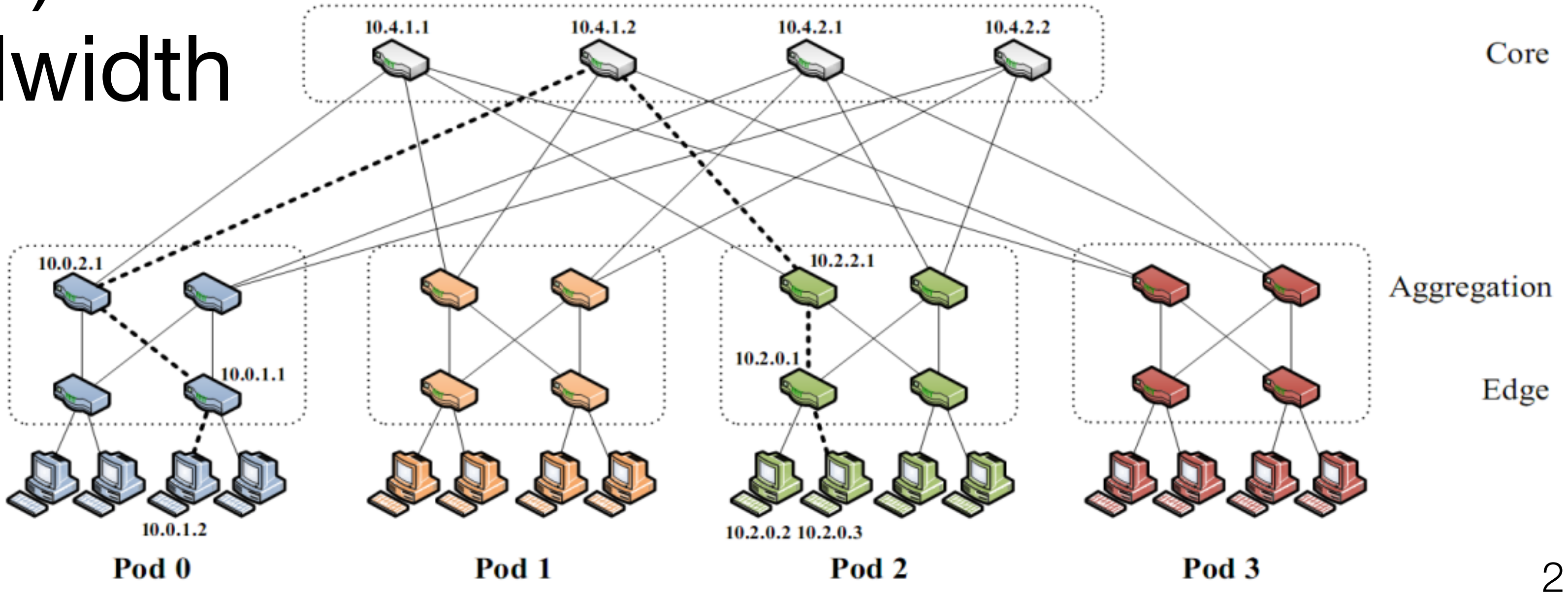
- Traditionally, networks topologies designs are demand oblivious:

Motivation

- Traditionally, networks topologies designs are demand oblivious:
  - optimized for the "worst-case": all-to-all communications

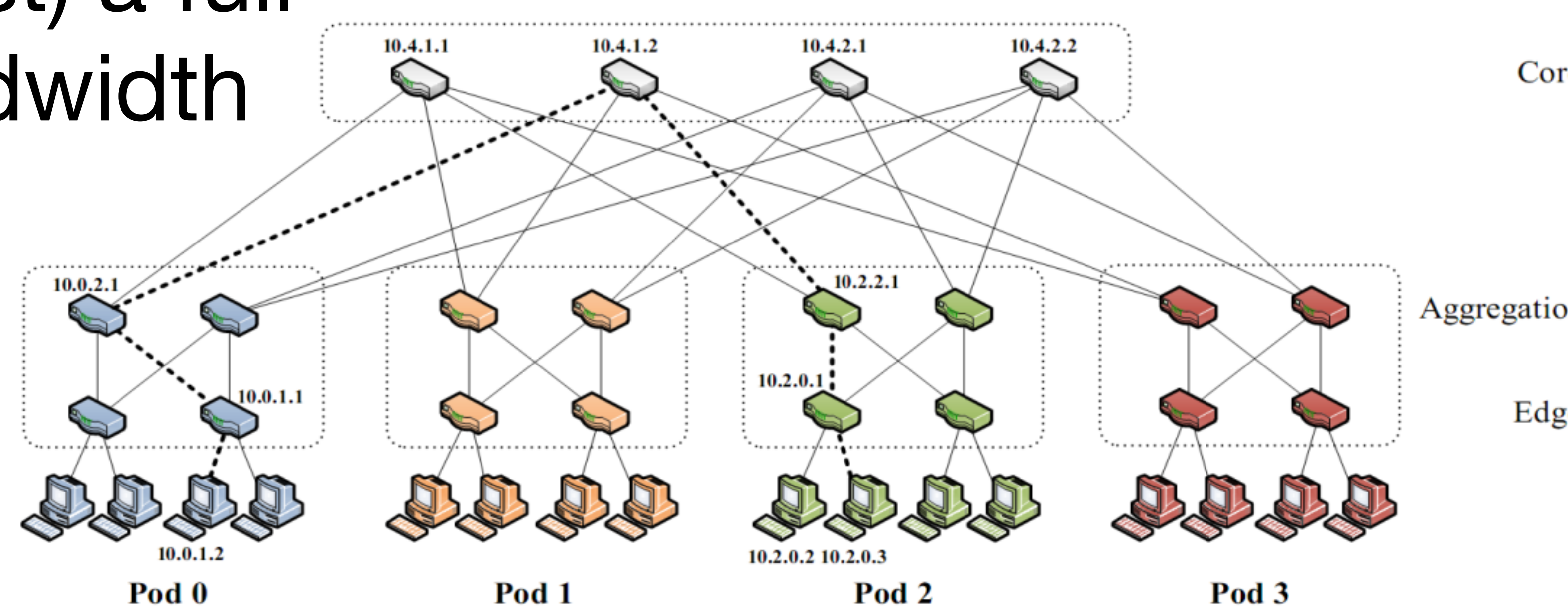
Motivation

- Traditionally, networks topologies designs are demand oblivious:
  - optimized for the "worst-case": all-to-all communications
    - e.g., Fat-tree topologies provide (almost) a full bisection bandwidth



Motivation

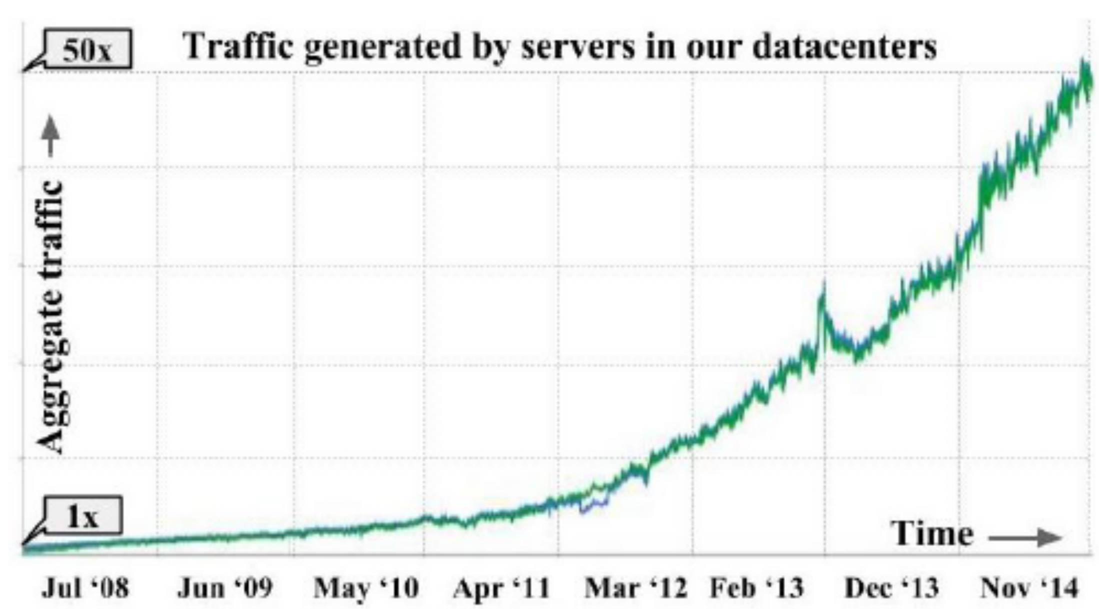
- Traditionally, networks topologies designs are demand oblivious:
  - optimized for the “worst-case”: all-to-all communications
    - e.g., Fat-tree topologies provide (almost) a full bisection bandwidth
- But…



Motivation

Motivation

- Traffic is growing fast

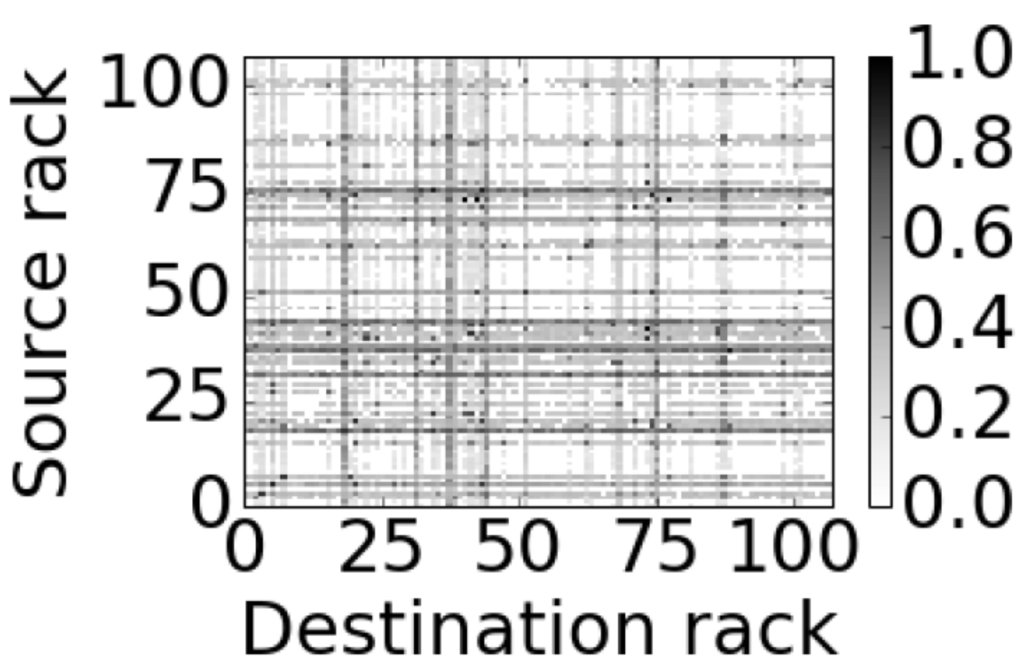


Aggregate Server Traffic in Google datacenter

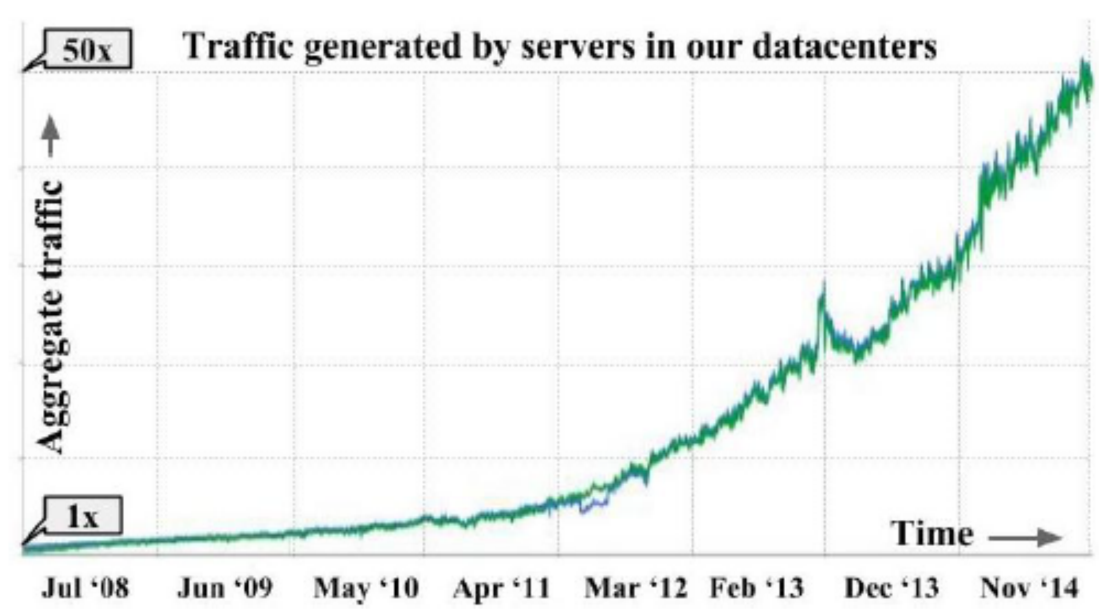
Jupiter rising @ SIGCOMM 2015

Motivation

- Traffic is growing fast
- Real communication patterns are sparse and feature structure (?!)



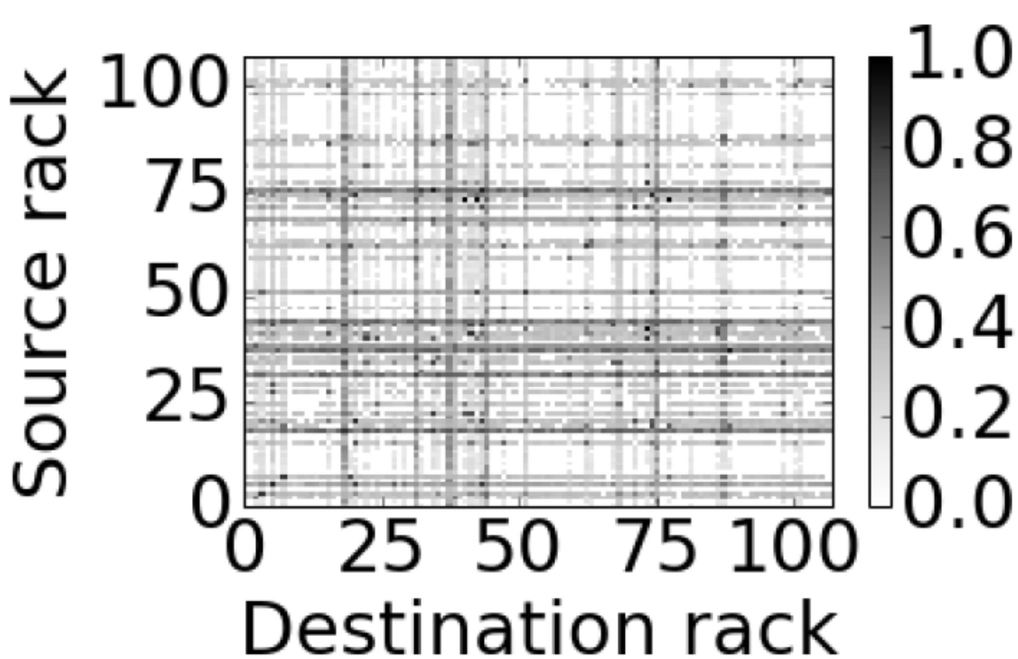
Heatmap of rack-to-rack traffic ProjecToR @ SIGCOMM 2016



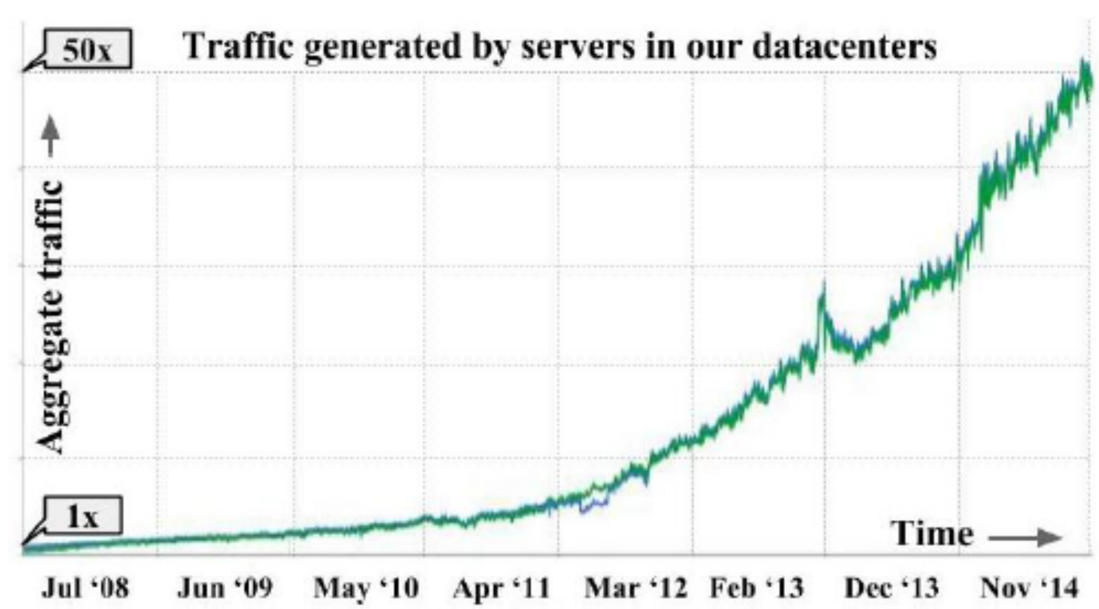
Aggregate Server Traffic in Google datacenter
Jupiter rising @ SIGCOMM 2015

Motivation

- Traffic is growing fast
- Real communication patterns are sparse and feature structure (?!)
- Can be exploited if demand is known: *demand-aware design*



Heatmap of rack-to-rack traffic ProjecToR @ SIGCOMM 2016



Aggregate Server Traffic in Google datacenter
Jupiter rising @ SIGCOMM 2015

Demand-Aware Design?

Application Layer

Transport Layer

Network Layer

Link / Physical Layer

Networks Capable of Change.
Jennifer Rexford.
Infocom 2019 Keynote.

Demand-Aware Design?

Application Layer ✓

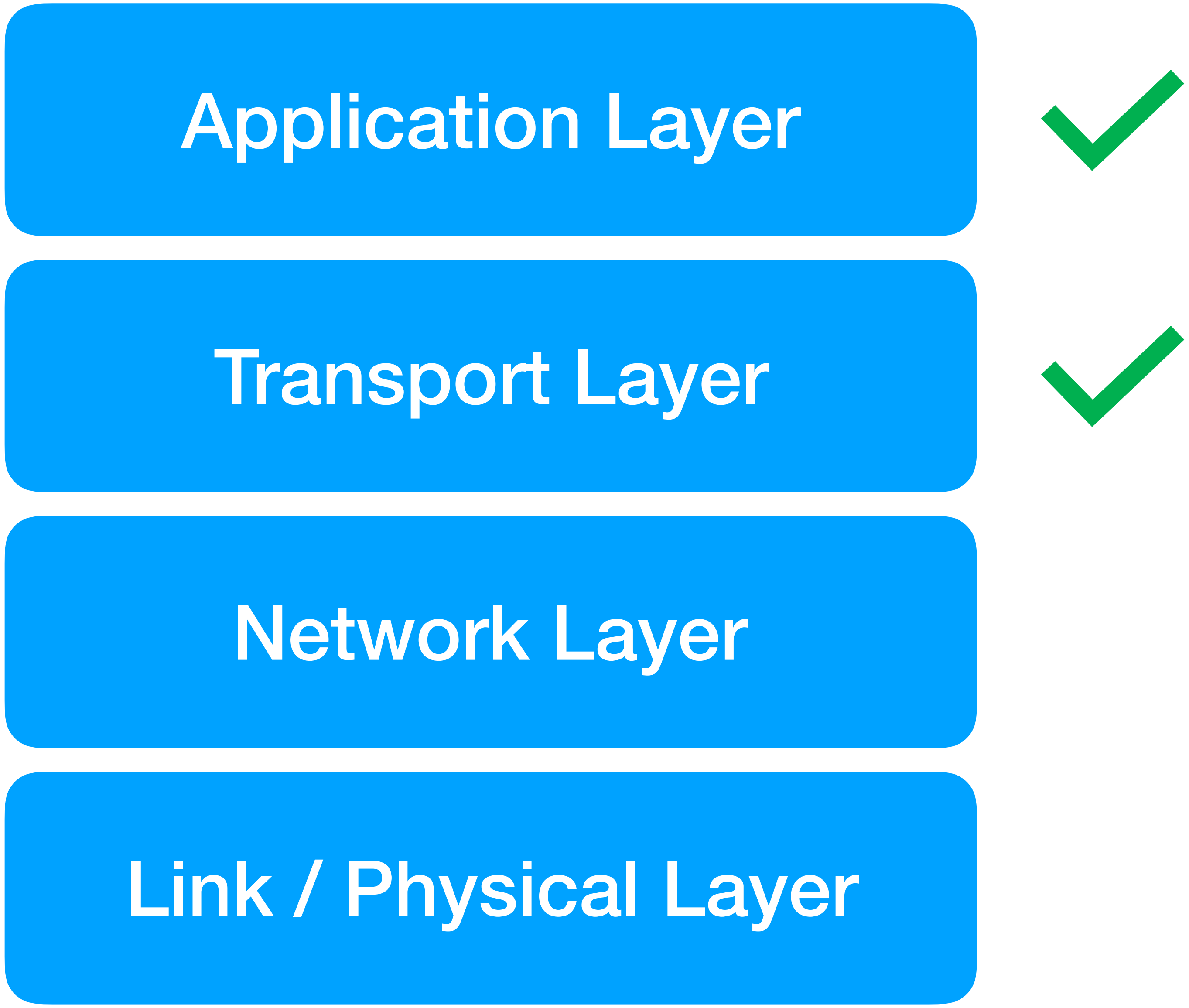
Transport Layer

Network Layer

Link / Physical Layer

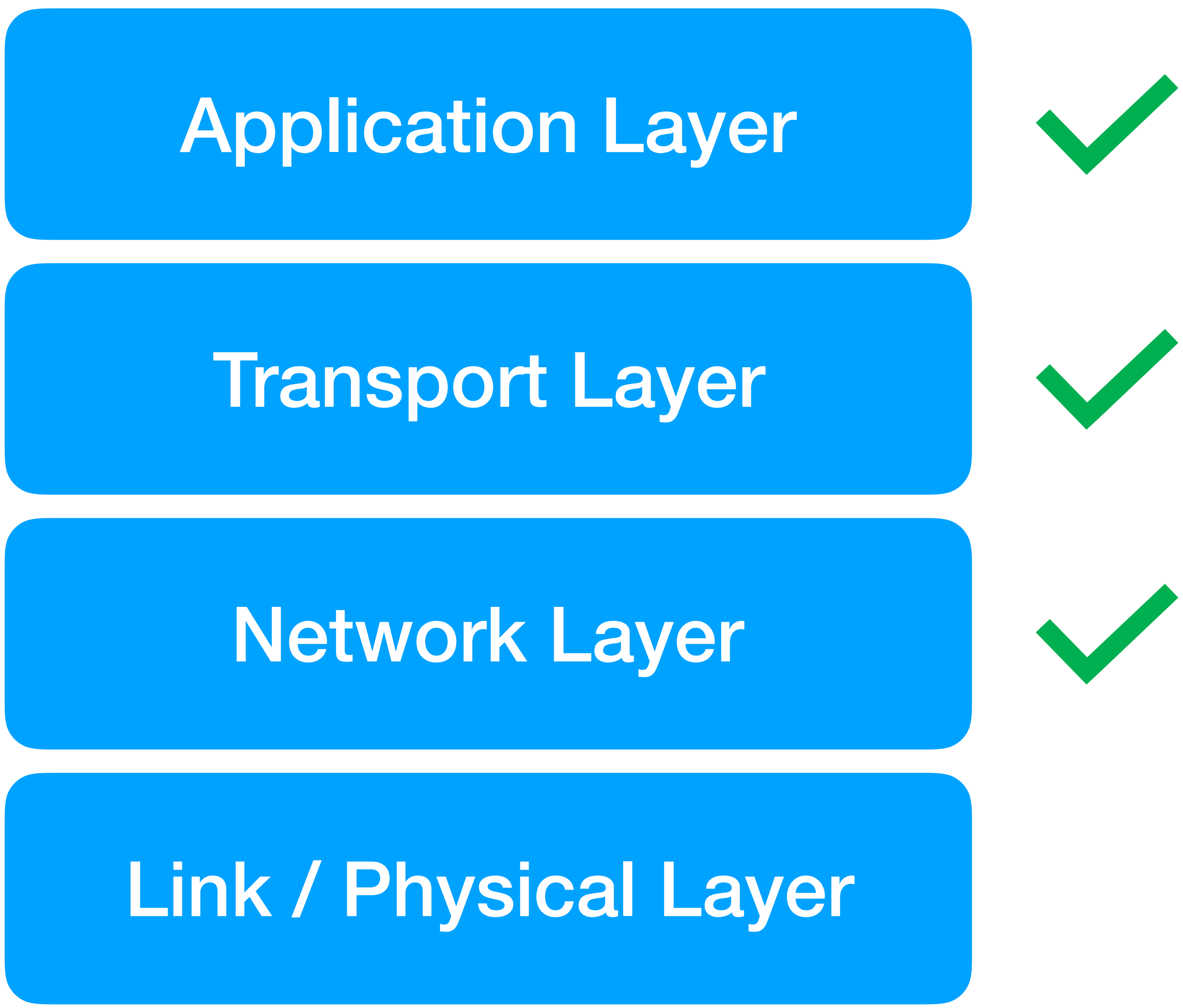
Networks Capable of Change.
Jennifer Rexford.
Infocom 2019 Keynote.

Demand-Aware Design?



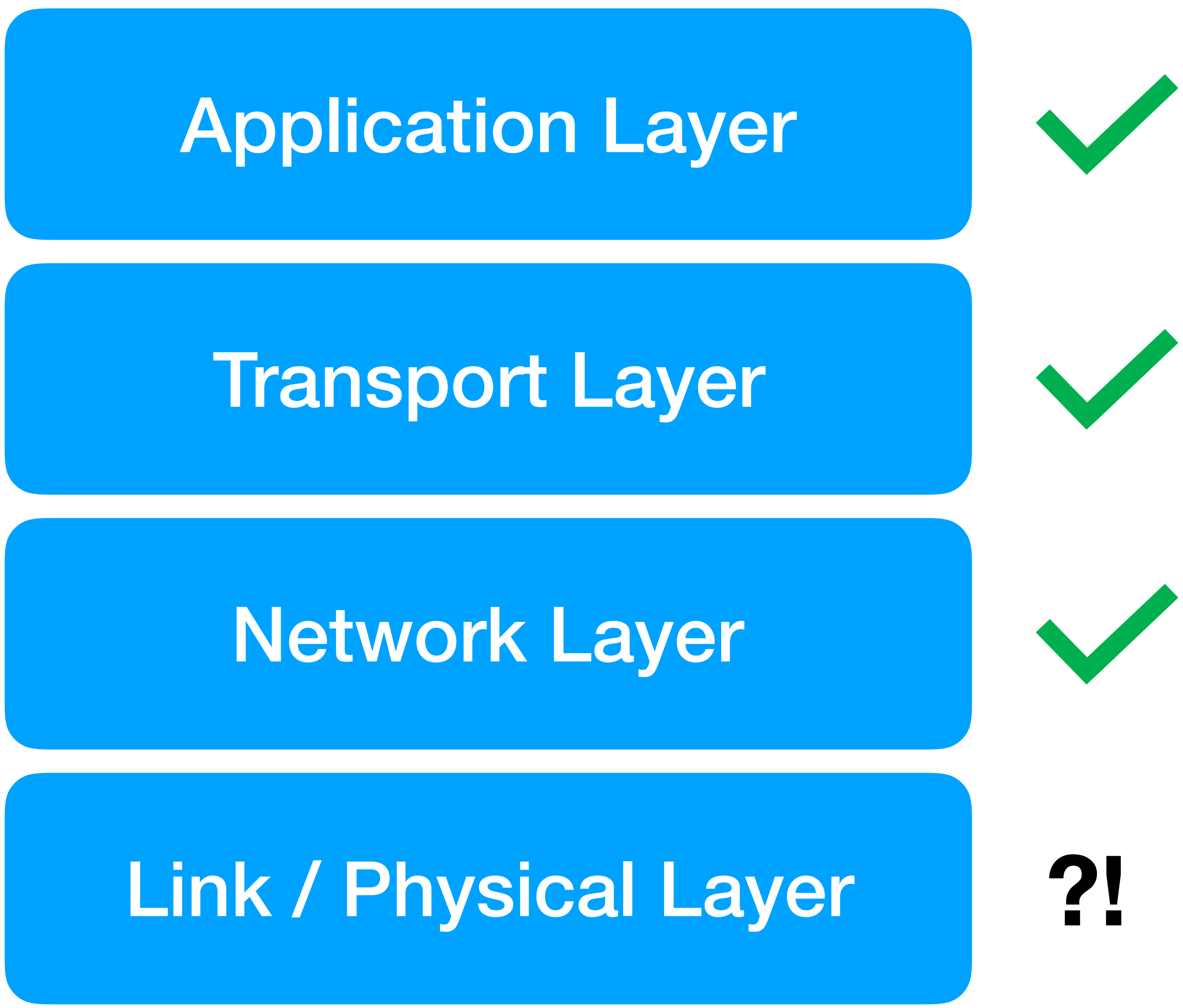
Networks Capable of Change.
Jennifer Rexford.
Infocom 2019 Keynote.

Demand-Aware Design?



Networks Capable of Change.
Jennifer Rexford.
Infocom 2019 Keynote.

Demand-Aware Design?



Networks Capable of Change.
Jennifer Rexford.
Infocom 2019 Keynote.

Demand-Aware Design?

- Application Layer ✓
- Transport Layer ✓
- Network Layer ✓
- Link / Physical Layer ?! ✓

Networks Capable of Change.
Jennifer Rexford.
Infocom 2019 Keynote.

Demand-Aware Design?

Application Layer ✓

Transport Layer ✓

Network Layer ✓

Link / Physical Layer **?!** ✓

Even in real time!

Networks Capable of Change.
Jennifer Rexford.
Infocom 2019 Keynote.

Goal

Goal

- Design a *demand aware*, *bounded degree* (scalable) networks with:

Goal

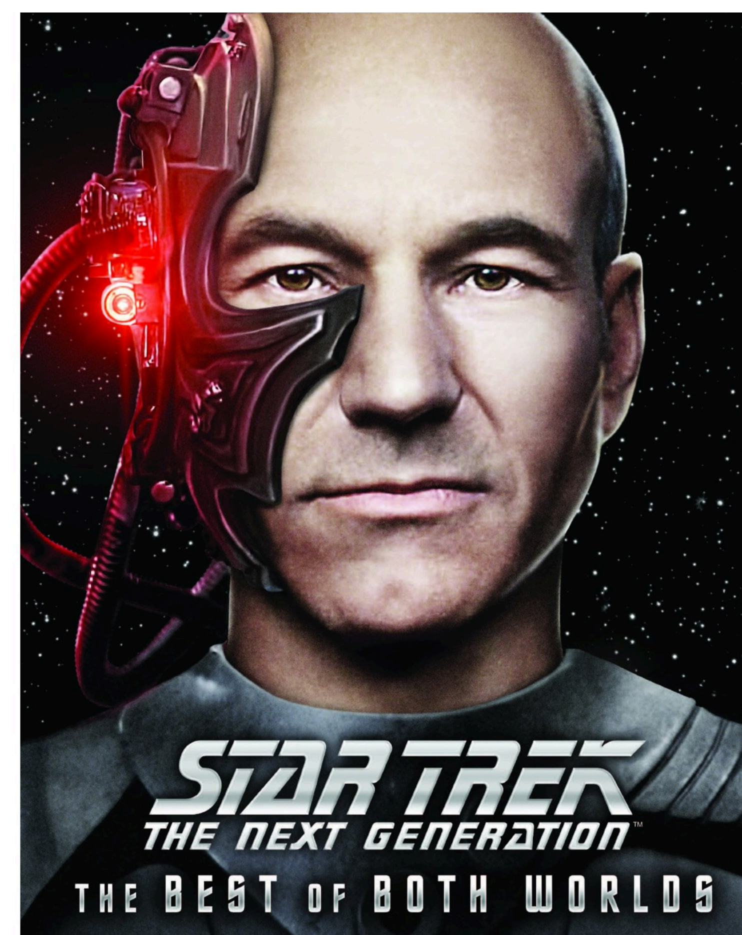
- Design a *demand aware*, *bounded degree* (scalable) networks with:
    1. Short *average route length (l)*, and

Goal

- Design a *demand aware*, *bounded degree* (scalable) networks with:
    1. Short *average route length (l)*, and
    2. *Low congestion (c)*:

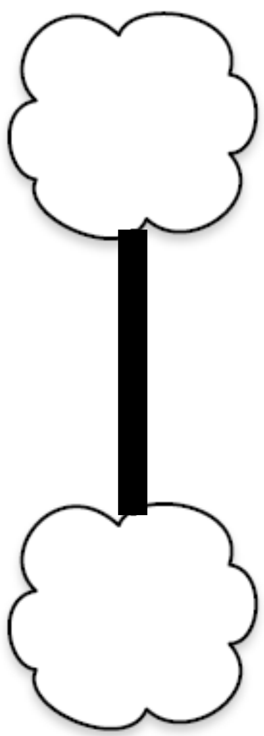
Goal

- Design a *demand aware*, *bounded degree* (scalable) networks with:
  1. Short *average* *route length (l)*, and
  2. *Low* *congestion (c)*:
- Both are important measures of efficiency



Challenges

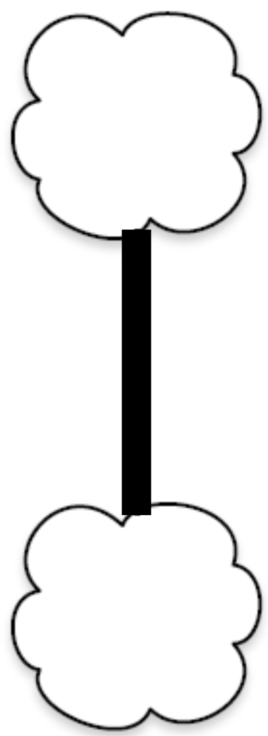
Challenges

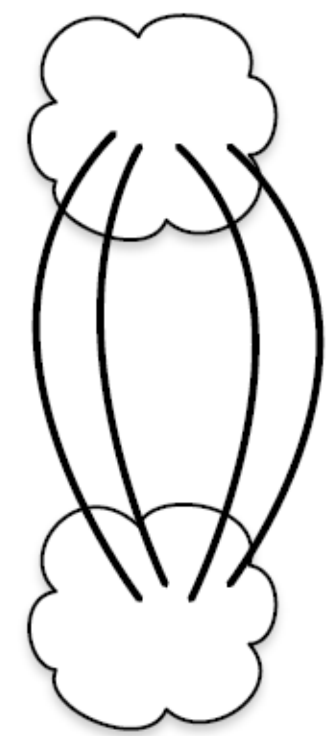


Short route length:
bottleneck

Chen Avin, Kaushik Mondal and Stefan Schmid
Demand-Aware Network Designs of Bounded Degree. DISC 2017

Challenges



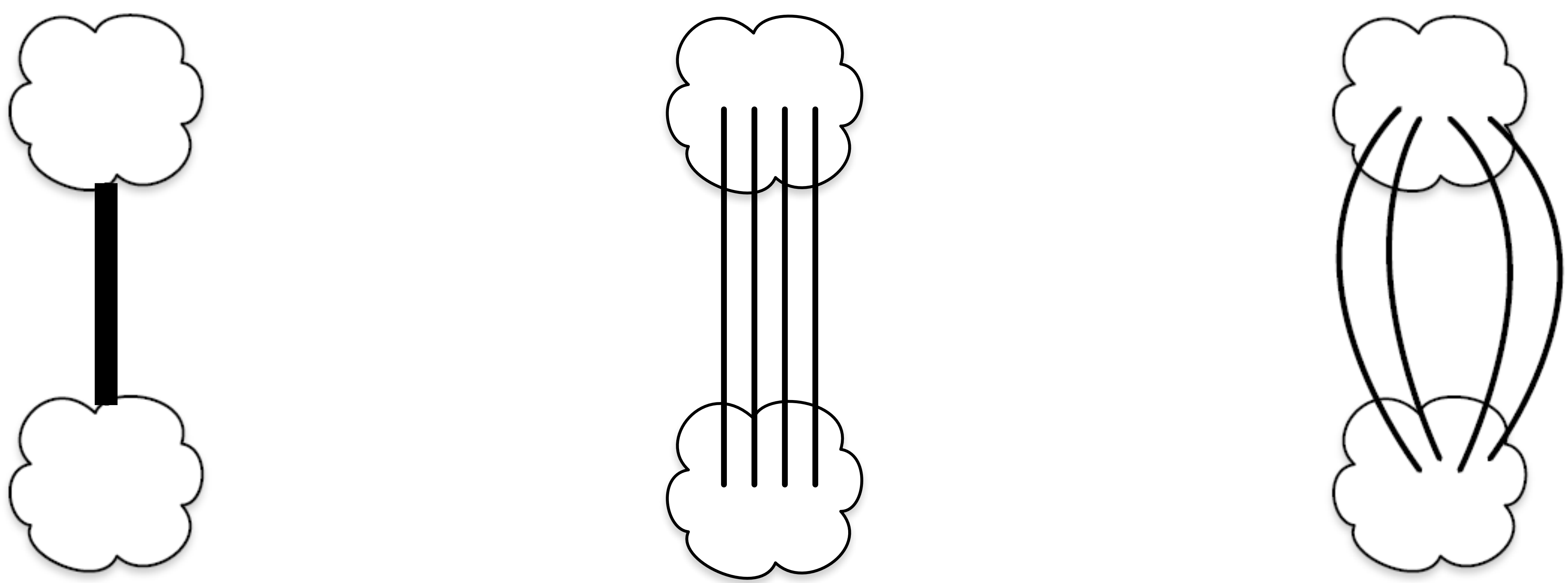


Short route length:
bottleneck

Low congestion:
high degree/ long routes

Chen Avin, Kaushik Mondal and Stefan Schmid
Demand-Aware Network Designs of Bounded Degree. DISC 2017

Challenges



Short route length:
bottleneck

Low congestion:
high degree/
long routes

Chen Avin, Kaushik Mondal and Stefan Schmid
Demand-Aware Network Designs of Bounded Degree. DISC 2017

Challenges: An Example

Goal: design an optimal network with bounded degree 3

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

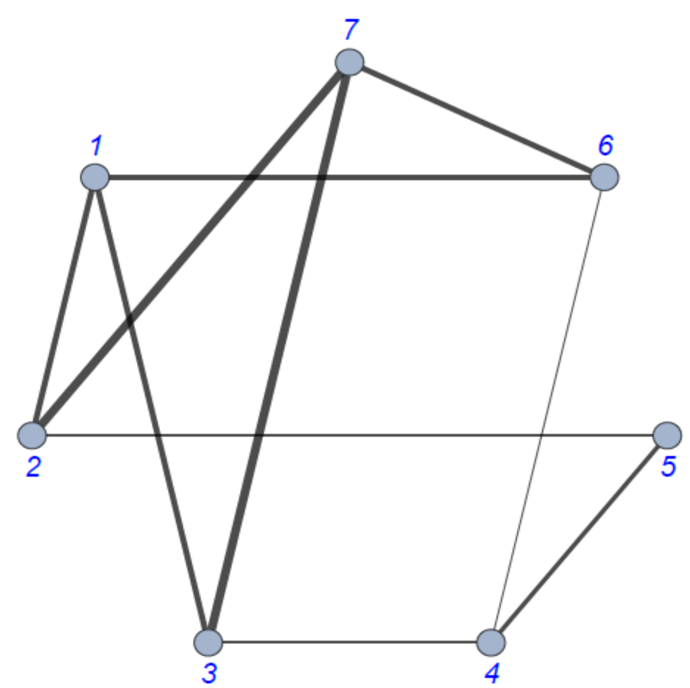
Demand Distribution

Challenges: An Example

Goal: design an optimal network with bounded degree 3

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

Demand Distribution



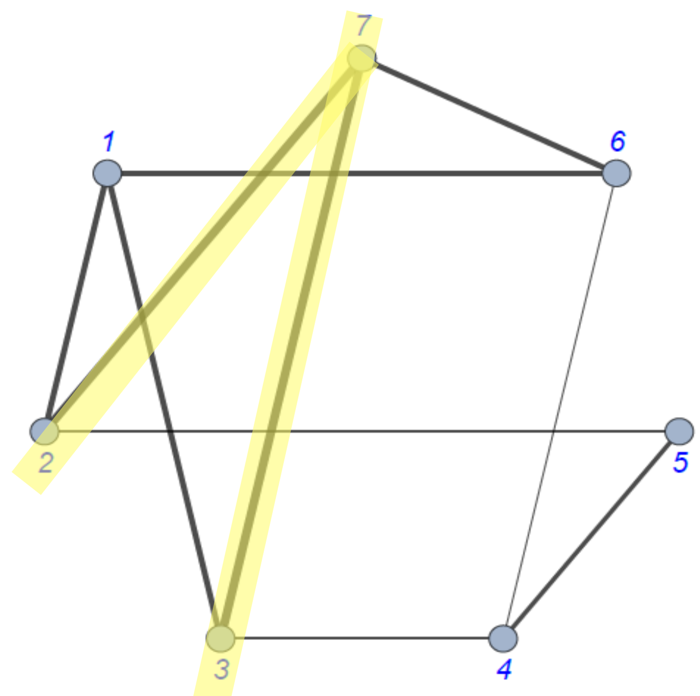
Route length is minimum, not congestion

Challenges: An Example

Goal: design an optimal network with bounded degree 3

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

Demand Distribution



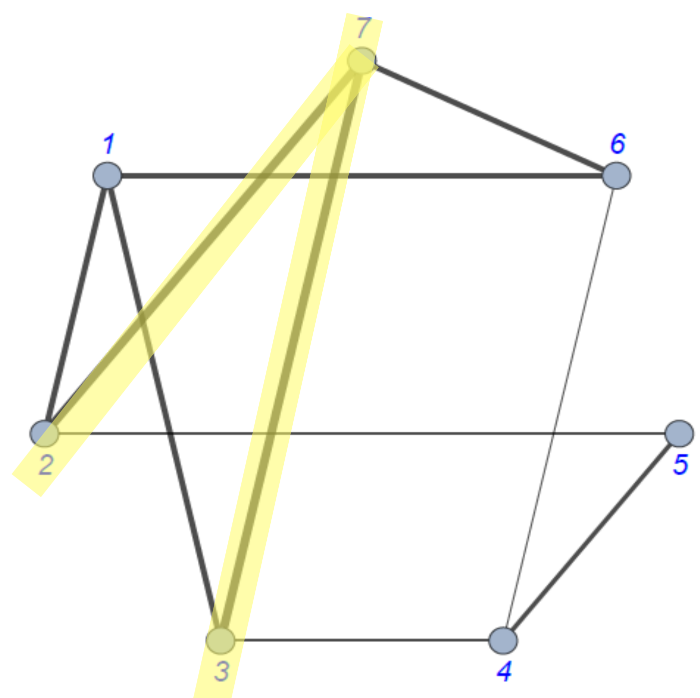
Route length is minimum, not congestion

Challenges: An Example

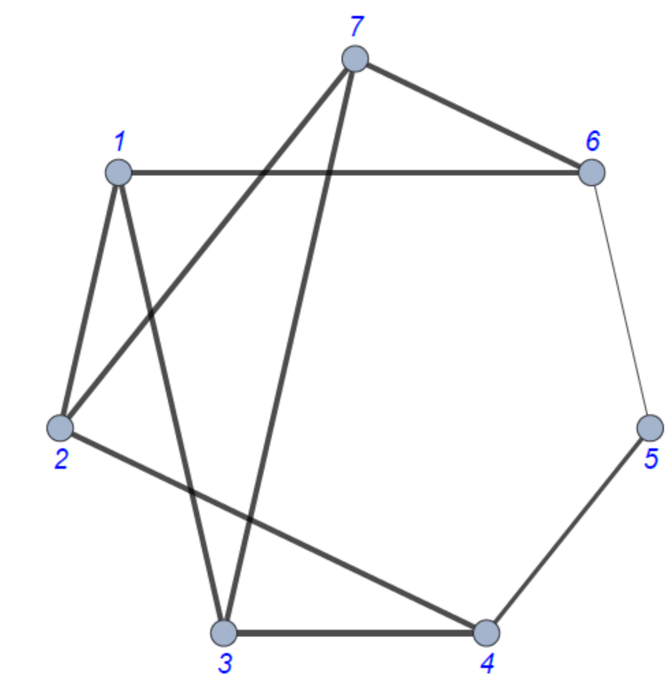
Goal: design an optimal network with bounded degree 3

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

Demand Distribution



Route length is minimum, not congestion



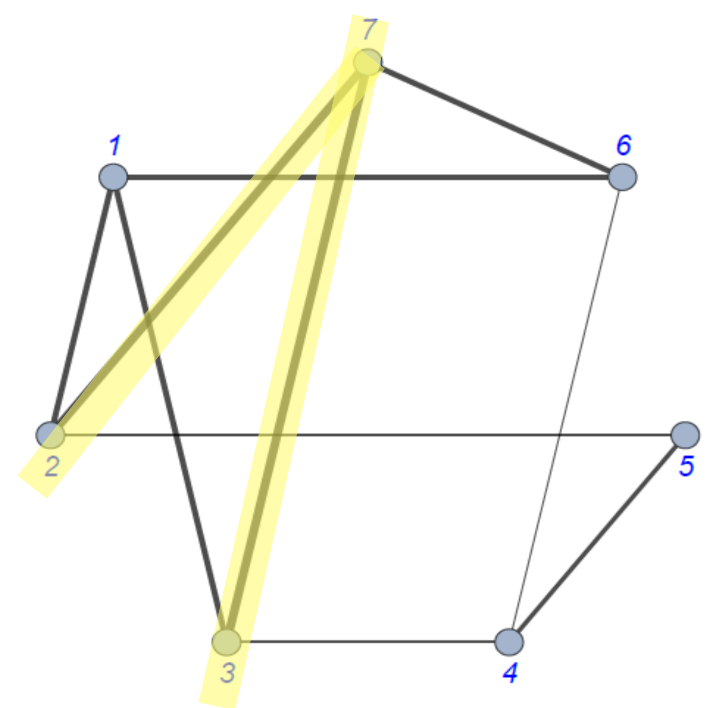
Congestion is minimum, not route length

Challenges: An Example

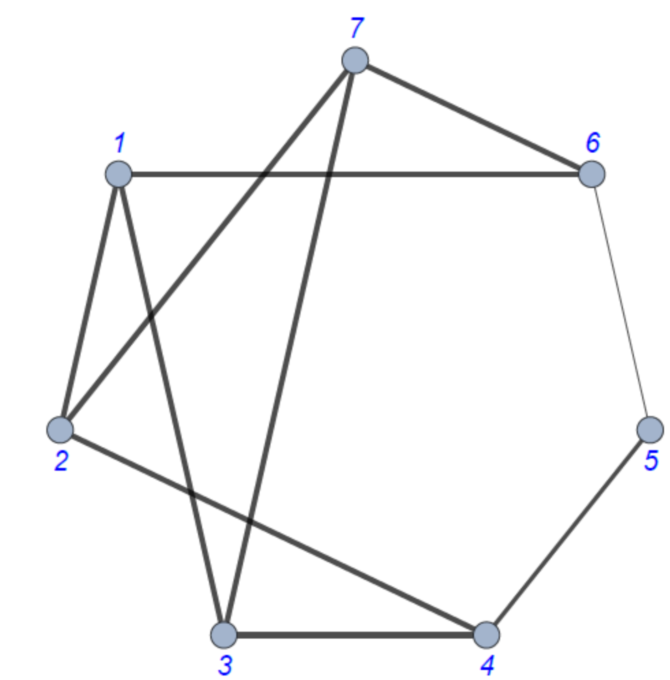
Goal: design an optimal network with bounded degree 3

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

Demand Distribution



Route length is minimum, not congestion



Congestion is minimum, not route length

Can we optimize both simultaneously?

Model and Definitions

Model and Definitions

- Demand joint distribution $\mathcal{D}$

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

Model and Definitions

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

- Demand joint distribution $\mathcal{D}$
- A network N

Model and Definitions

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

- Demand joint distribution $\mathcal{D}$
- A network N
- A routing scheme $\Gamma(N)$

Model and Definitions

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

- Demand joint distribution $\mathcal{D}$
- A network N
- A routing scheme $\Gamma(N)$
- The congestion:

$$\mathrm{C}(\mathcal{D}, \Gamma(N)) = \max_{e \in \Gamma(N)} \sum_{e \in \Gamma_{uv}} \mathrm{p}(u, v)$$

Model and Definitions

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

- Demand joint distribution $\mathcal{D}$
- A network N
- A routing scheme $\Gamma(N)$
- The congestion:

$$\mathrm{C}(\mathcal{D}, \Gamma(N)) = \max_{e \in \Gamma(N)} \sum_{e \in \Gamma_{uv}} \mathrm{p}(u, v)$$

- The weighted path length

$$\mathrm{L}(\mathcal{D}, \Gamma(N)) = \sum_{(u,v) \in \mathcal{D}} \mathrm{p}(u, v) \cdot \mathrm{d}_{\Gamma(N)}(u, v)$$

Goal: (α, β) *cl*-DAN Design

Goal: (α, β) *cl*-DAN Design

Input: matrix, degree

$$\mathcal{D}, \Delta$$

	1	2	3	4	5	6	7
1	0	$\frac{2}{65}$	$\frac{1}{13}$	$\frac{1}{65}$	$\frac{1}{65}$	$\frac{2}{65}$	$\frac{3}{65}$
2	$\frac{2}{65}$	0	$\frac{1}{65}$	0	0	0	$\frac{2}{65}$
3	$\frac{1}{13}$	$\frac{1}{65}$	0	$\frac{2}{65}$	0	0	$\frac{1}{13}$
4	$\frac{1}{65}$	0	$\frac{2}{65}$	0	$\frac{4}{65}$	0	0
5	$\frac{1}{65}$	0	$\frac{3}{65}$	$\frac{4}{65}$	0	0	0
6	$\frac{2}{65}$	0	0	0	0	0	$\frac{3}{65}$
7	$\frac{3}{65}$	$\frac{2}{65}$	$\frac{1}{13}$	0	0	$\frac{3}{65}$	0

Goal: (α, β) *cl*-DAN Design

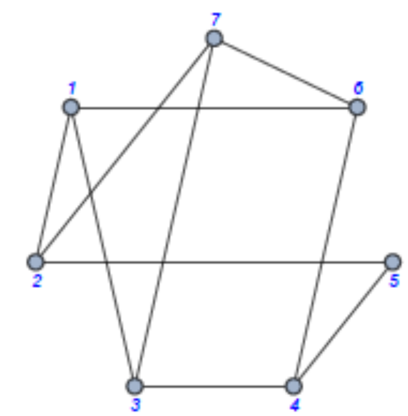
Input: matrix, degree

$\mathcal{D}, \Delta$

	1	2	3	4	5	6	7
1	0	$\frac{2}{65}$	$\frac{1}{13}$	$\frac{1}{65}$	$\frac{1}{65}$	$\frac{2}{65}$	$\frac{3}{65}$
2	$\frac{2}{65}$	0	$\frac{1}{65}$	0	0	0	$\frac{2}{65}$
3	$\frac{1}{13}$	$\frac{1}{65}$	0	$\frac{2}{65}$	0	0	$\frac{1}{13}$
4	$\frac{1}{65}$	0	$\frac{2}{65}$	0	$\frac{4}{65}$	0	0
5	$\frac{1}{65}$	0	$\frac{3}{65}$	$\frac{4}{65}$	0	0	0
6	$\frac{2}{65}$	0	0	0	0	0	$\frac{3}{65}$
7	$\frac{3}{65}$	$\frac{2}{65}$	$\frac{1}{13}$	0	0	$\frac{3}{65}$	0

Output: *cl*-DAN

$N \in \mathcal{N}_\Delta, \Gamma(N)$



Goal: (α, β) *cl*-DAN Design

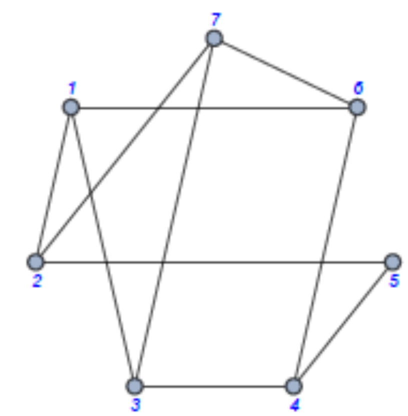
Input: matrix, degree

$\mathcal{D}, \Delta$

	1	2	3	4	5	6	7
1	0	$\frac{2}{65}$	$\frac{1}{13}$	$\frac{1}{65}$	$\frac{1}{65}$	$\frac{2}{65}$	$\frac{3}{65}$
2	$\frac{2}{65}$	0	$\frac{1}{65}$	0	0	0	$\frac{2}{65}$
3	$\frac{1}{13}$	$\frac{1}{65}$	0	$\frac{2}{65}$	0	0	$\frac{1}{13}$
4	$\frac{1}{65}$	0	$\frac{2}{65}$	0	$\frac{4}{65}$	0	0
5	$\frac{1}{65}$	0	$\frac{3}{65}$	$\frac{4}{65}$	0	0	0
6	$\frac{2}{65}$	0	0	0	0	0	$\frac{3}{65}$
7	$\frac{3}{65}$	$\frac{2}{65}$	$\frac{1}{13}$	0	0	$\frac{3}{65}$	0

Output: *cl*-DAN

$N \in \mathcal{N}_\Delta, \Gamma(N)$



Where

Optimal congestion

$$\mathrm{L}(\mathcal{D}, \Gamma(N)) \leq \beta \, \mathrm{L}^*(\mathcal{D}, \Delta) + \beta'$$

Goal: (α, β) *cl*-DAN Design

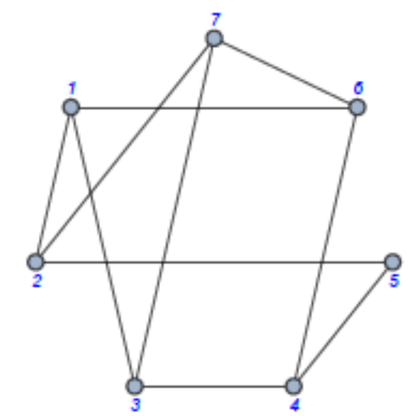
Input: matrix, degree

$\mathcal{D}, \Delta$

	1	2	3	4	5	6	7
1	0	$\frac{2}{65}$	$\frac{1}{13}$	$\frac{1}{65}$	$\frac{1}{65}$	$\frac{2}{65}$	$\frac{3}{65}$
2	$\frac{2}{65}$	0	$\frac{1}{65}$	0	0	0	$\frac{2}{65}$
3	$\frac{1}{13}$	$\frac{1}{65}$	0	$\frac{2}{65}$	0	0	$\frac{1}{13}$
4	$\frac{1}{65}$	0	$\frac{2}{65}$	0	$\frac{4}{65}$	0	0
5	$\frac{1}{65}$	0	$\frac{3}{65}$	$\frac{4}{65}$	0	0	0
6	$\frac{2}{65}$	0	0	0	0	0	$\frac{3}{65}$
7	$\frac{3}{65}$	$\frac{2}{65}$	$\frac{1}{13}$	0	0	$\frac{3}{65}$	0

Output: *cl*-DAN

$N \in \mathcal{N}_\Delta, \Gamma(N)$



Where

Optimal congestion $\quad \mathrm{L}(\mathcal{D}, \Gamma(N)) \leq \beta \cdot \mathrm{L}^*(\mathcal{D}, \Delta) + \beta'$

&

Optimal path length $\quad \mathrm{C}(\mathcal{D}, \Gamma(N)) \leq \alpha \cdot \mathrm{C}^*(\mathcal{D}, \Delta) + \alpha'$

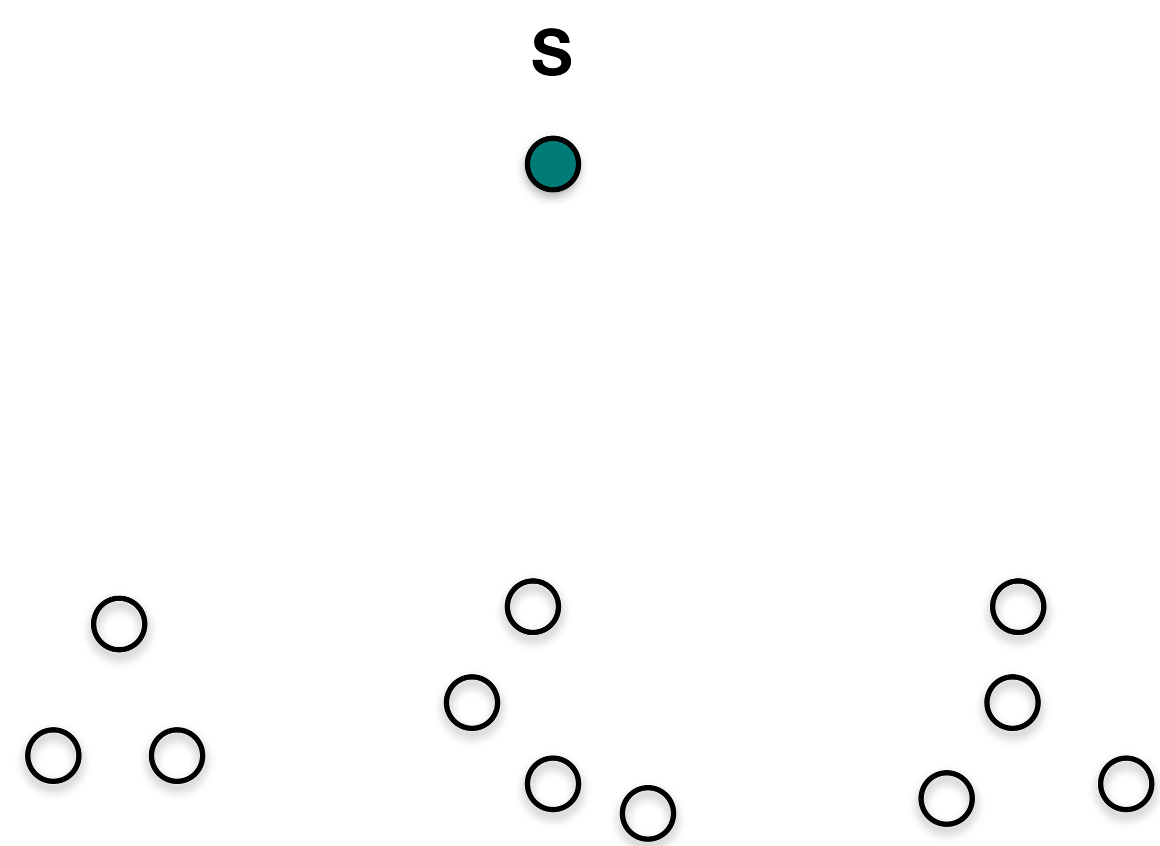
A Building Block: EgoTree

A Building Block: EgoTree

- A single source multiple destination problem

A Building Block: EgoTree

- A single source multiple destination problem

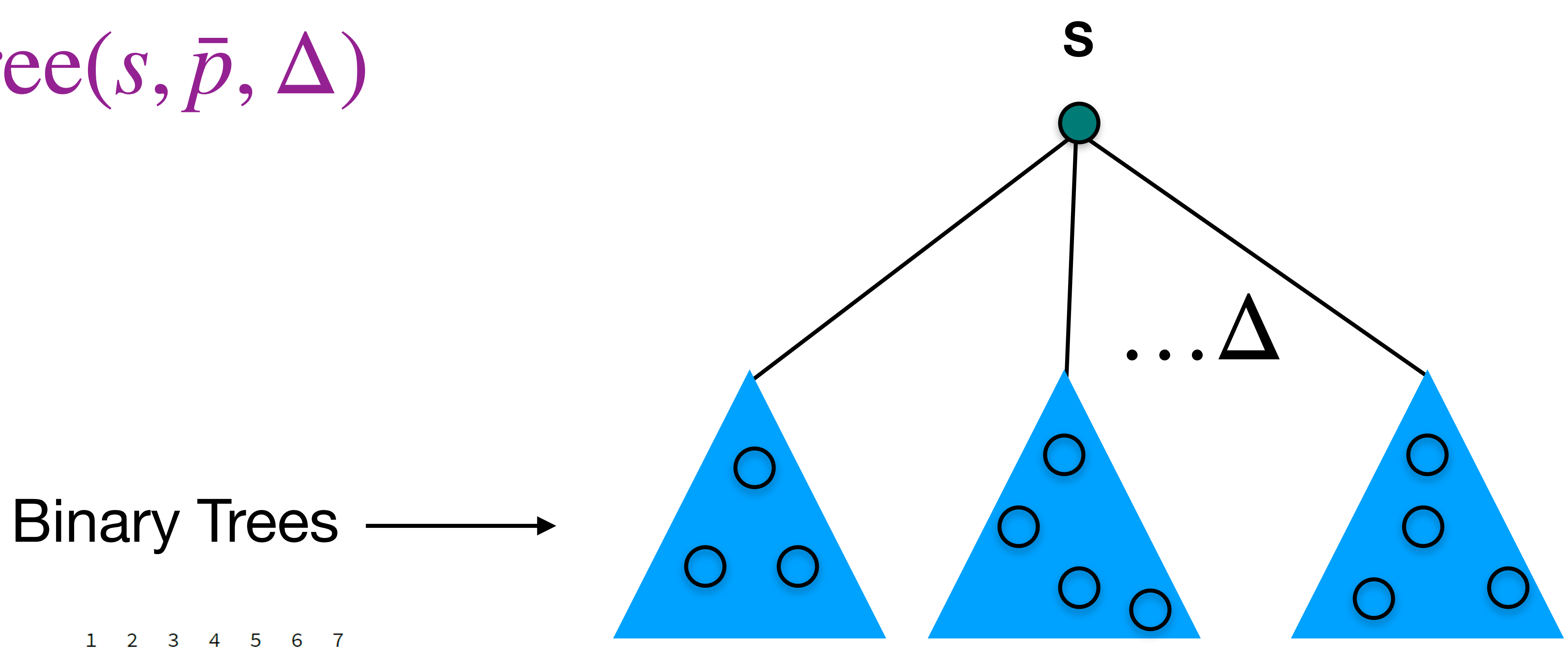


	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

$$\bar{p} = \{p_1, p_2, \ldots, p_k\}$$

A Building Block: EgoTree

- A single source multiple destination problem
- EgoTree$(s, \bar{p}, \Delta)$



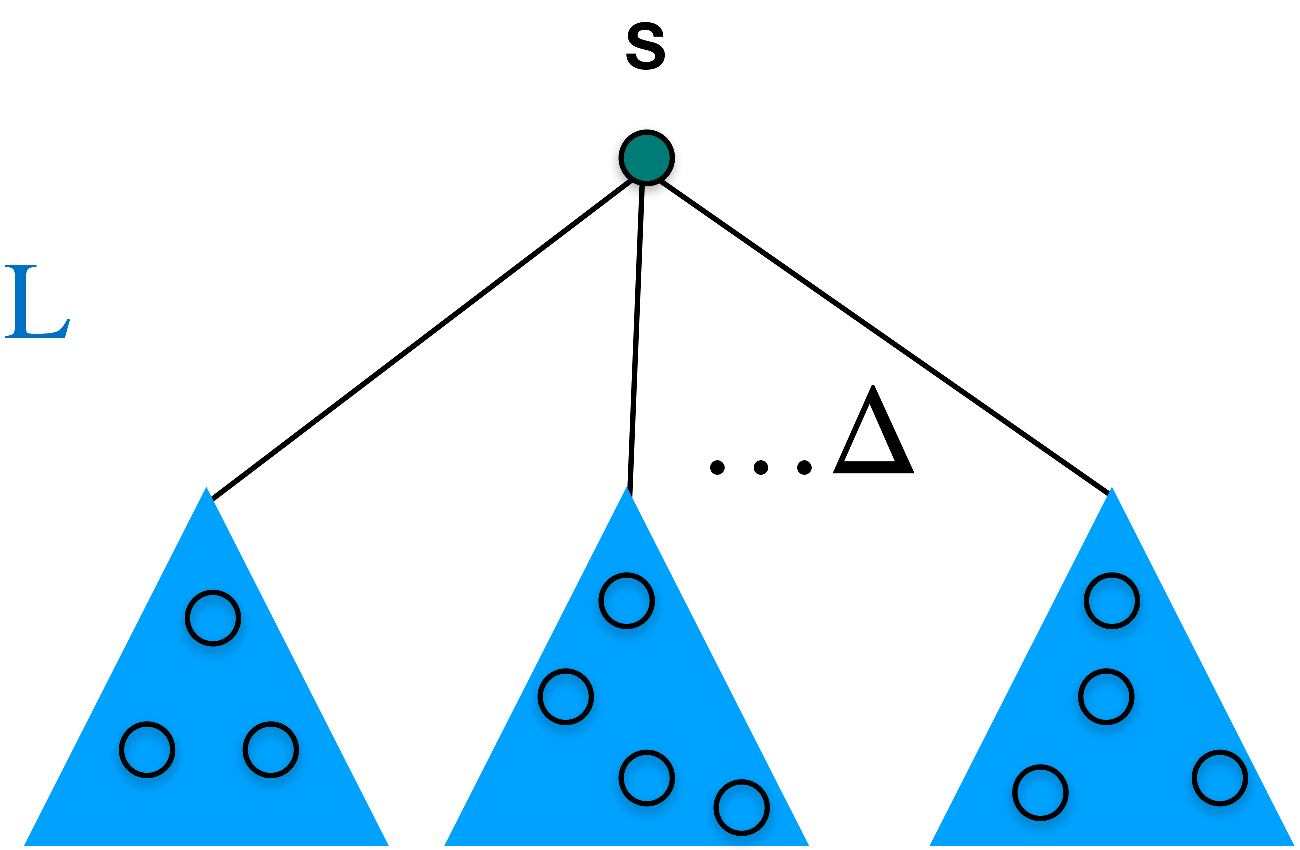
$$\bar{p} = \{p_1, p_2, \ldots, p_k\}$$

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

A Building Block: EgoTree

- A single source multiple destination problem
- $\mathrm{EgoTree}(s, \bar{p}, \Delta)$
- Optimizes both C and L

Binary Trees →

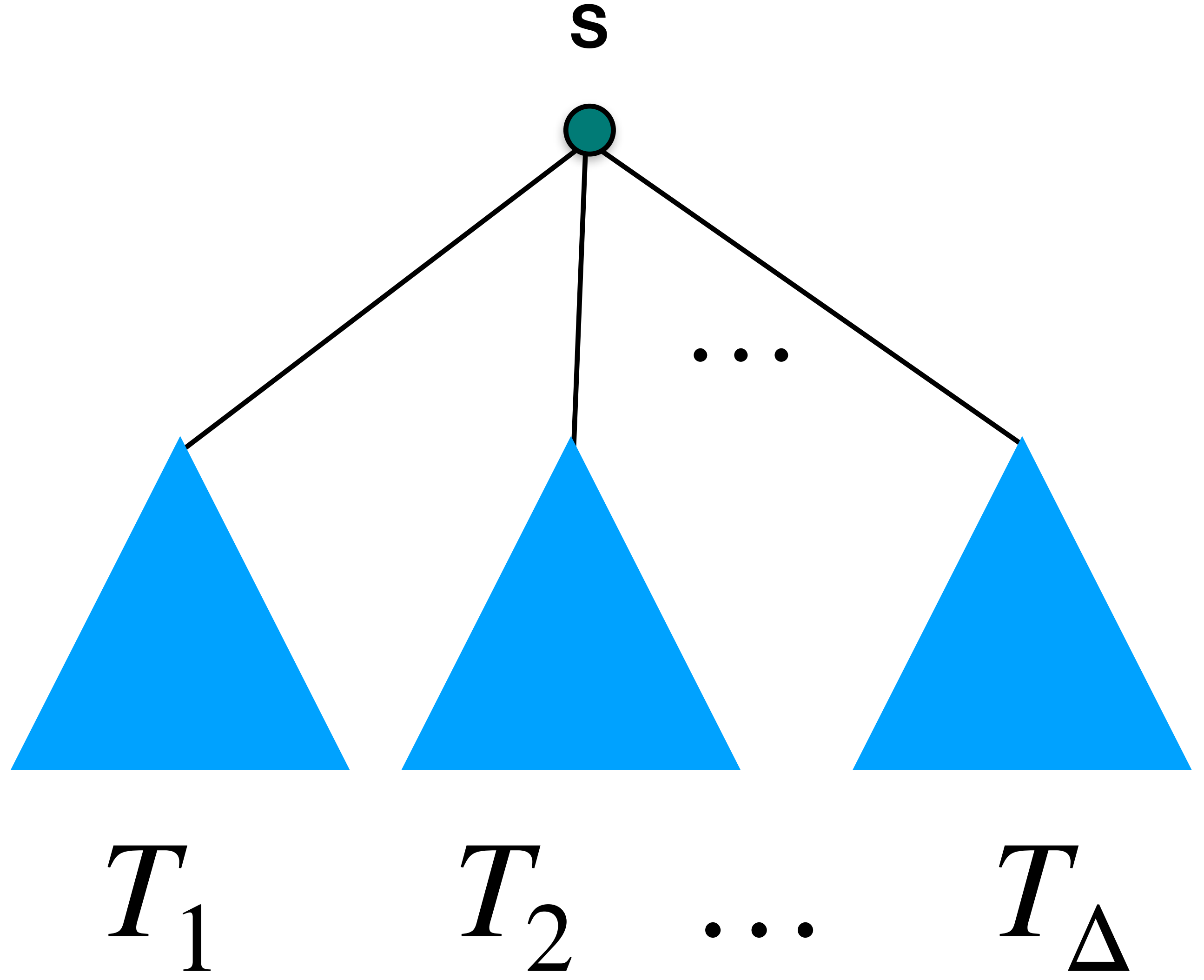


$$\bar{p} = \{p_1, p_2, \ldots, p_k\}$$

	1	2	3	4	5	6	7
1	0	$\frac{3}{60}$	$\frac{4}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$
2	$\frac{3}{60}$	0	$\frac{2}{60}$	0	$\frac{1}{60}$	0	$\frac{4}{60}$
3	$\frac{4}{60}$	$\frac{2}{60}$	0	$\frac{2}{60}$	0	0	$\frac{4}{60}$
4	$\frac{1}{60}$	0	$\frac{2}{60}$	0	$\frac{3}{60}$	0	0
5	$\frac{1}{60}$	$\frac{1}{60}$	0	$\frac{3}{60}$	0	0	0
6	$\frac{1}{60}$	0	0	0	0	0	$\frac{3}{60}$
7	$\frac{1}{60}$	$\frac{4}{60}$	$\frac{4}{60}$	0	0	$\frac{3}{60}$	0

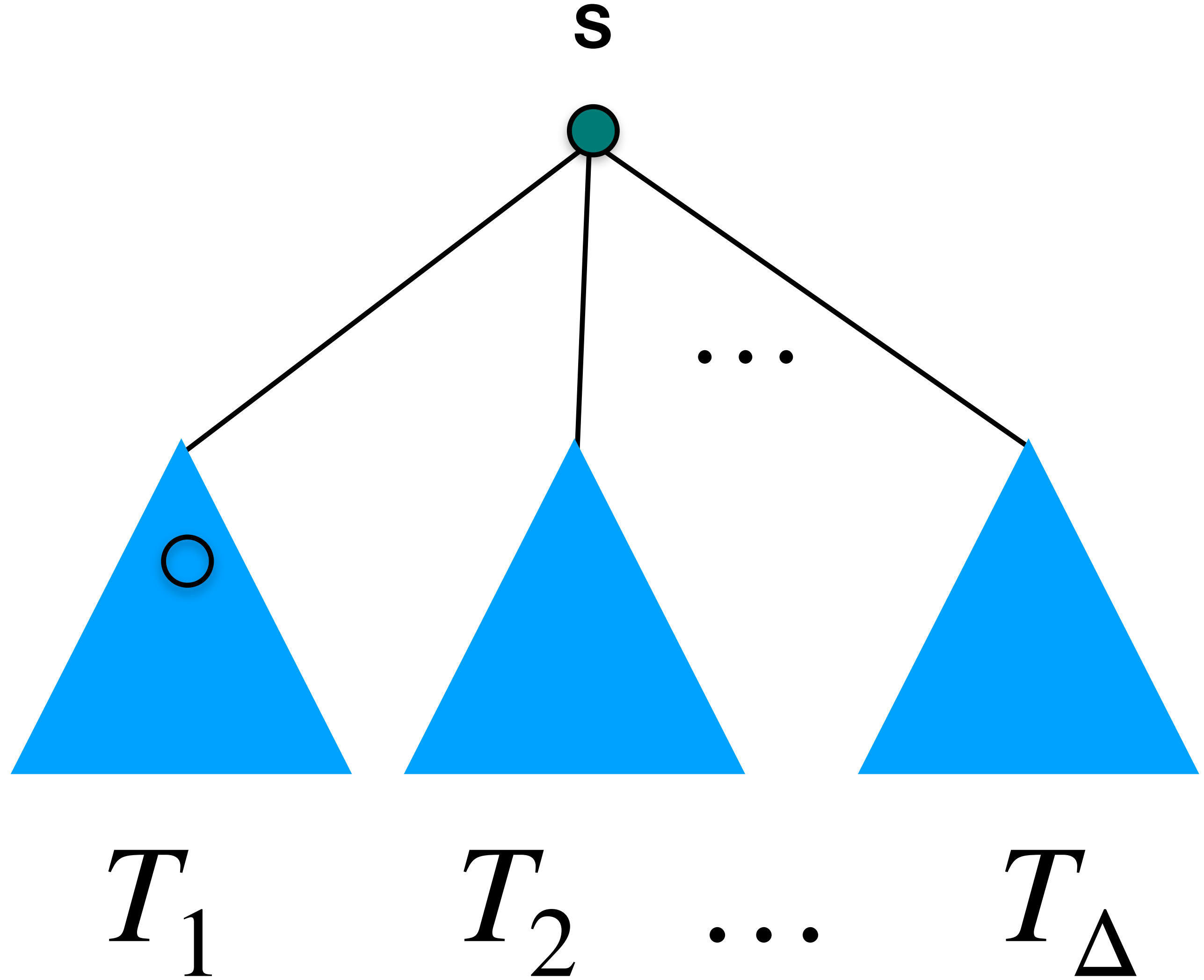
EgoTree: A Greedy Construction

- Sort $= \{p_1, p_2, \ldots, p_k\}$ from large to small
- Greedily Place destination nodes (one at a time)



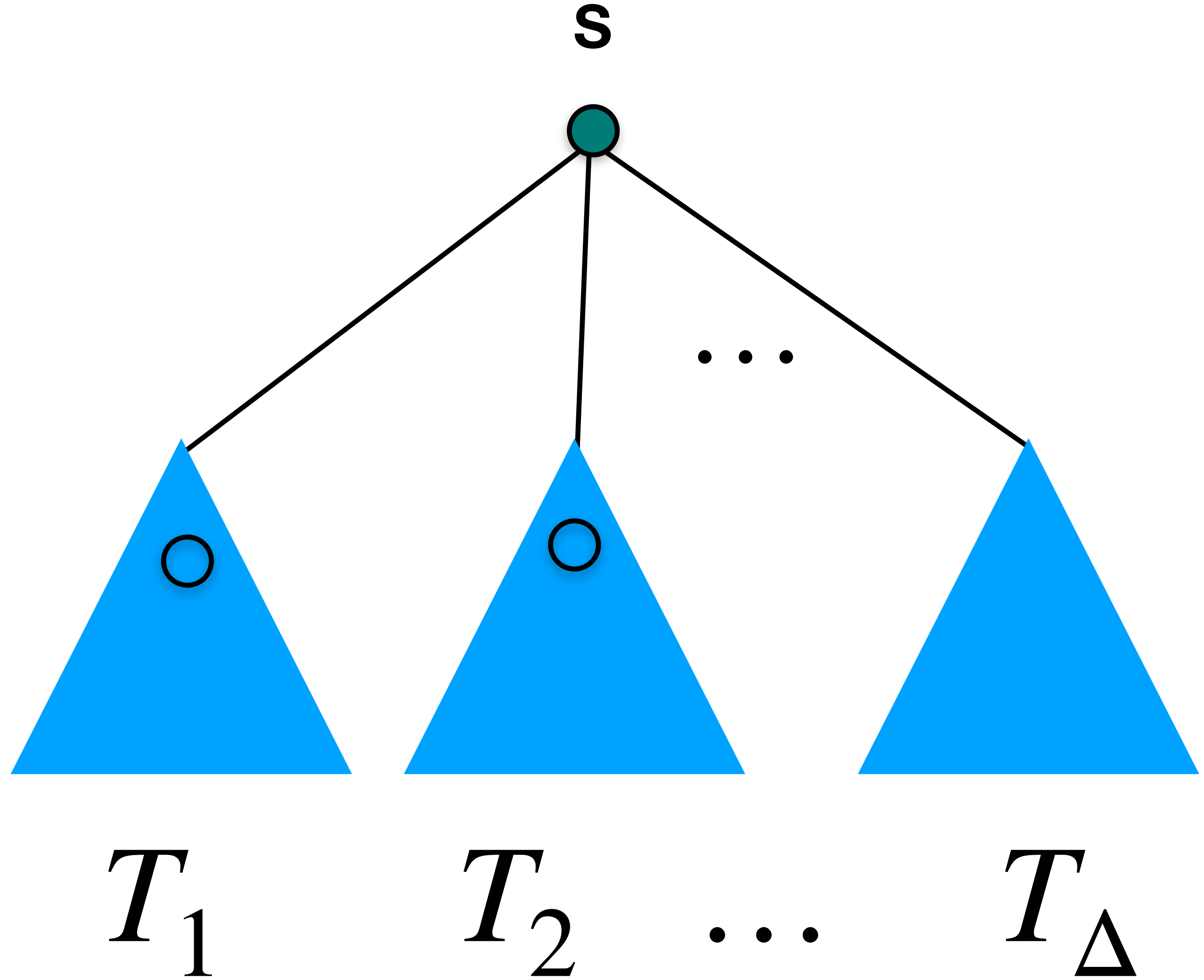
EgoTree: A Greedy Construction

- Sort $= \{p_1, p_2, \ldots, p_k\}$ from large to small
- Greedily Place destination nodes (one at a time)



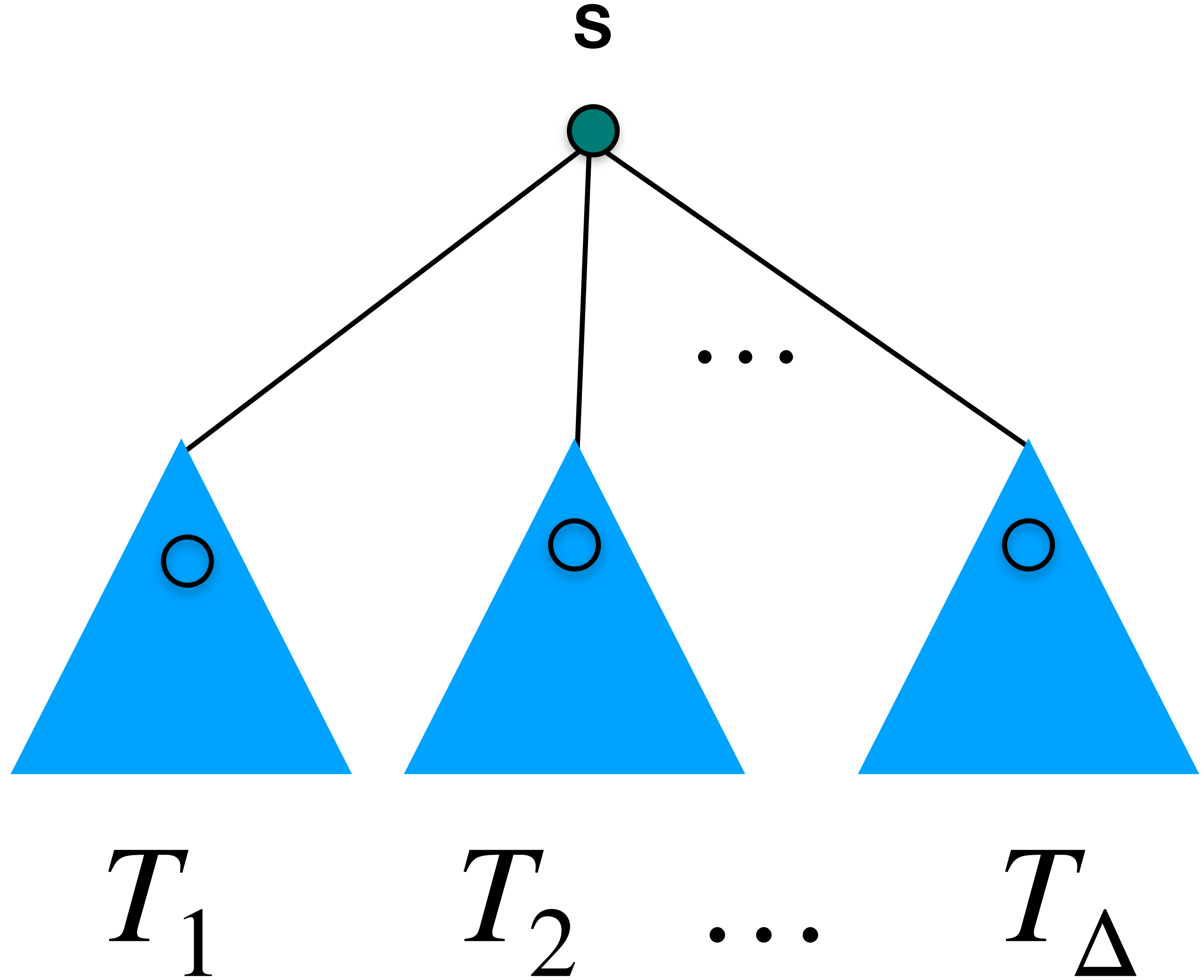
EgoTree: A Greedy Construction

- Sort $= \{p_1, p_2, \ldots, p_k\}$ from large to small
- Greedily Place destination nodes (one at a time)



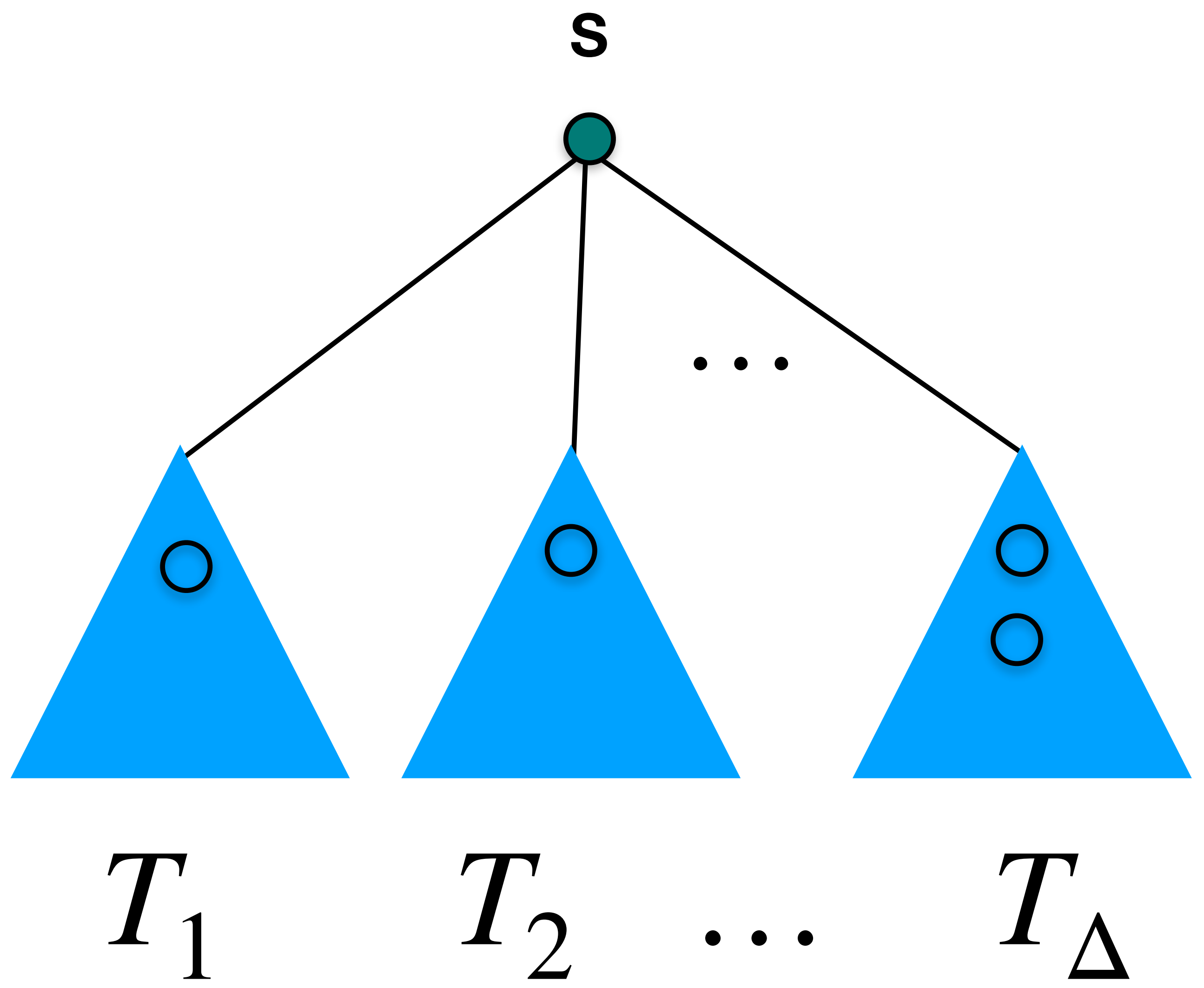
EgoTree: A Greedy Construction

- Sort $= \{p_1, p_2, \ldots, p_k\}$ from large to small
- Greedily Place destination nodes (one at a time)



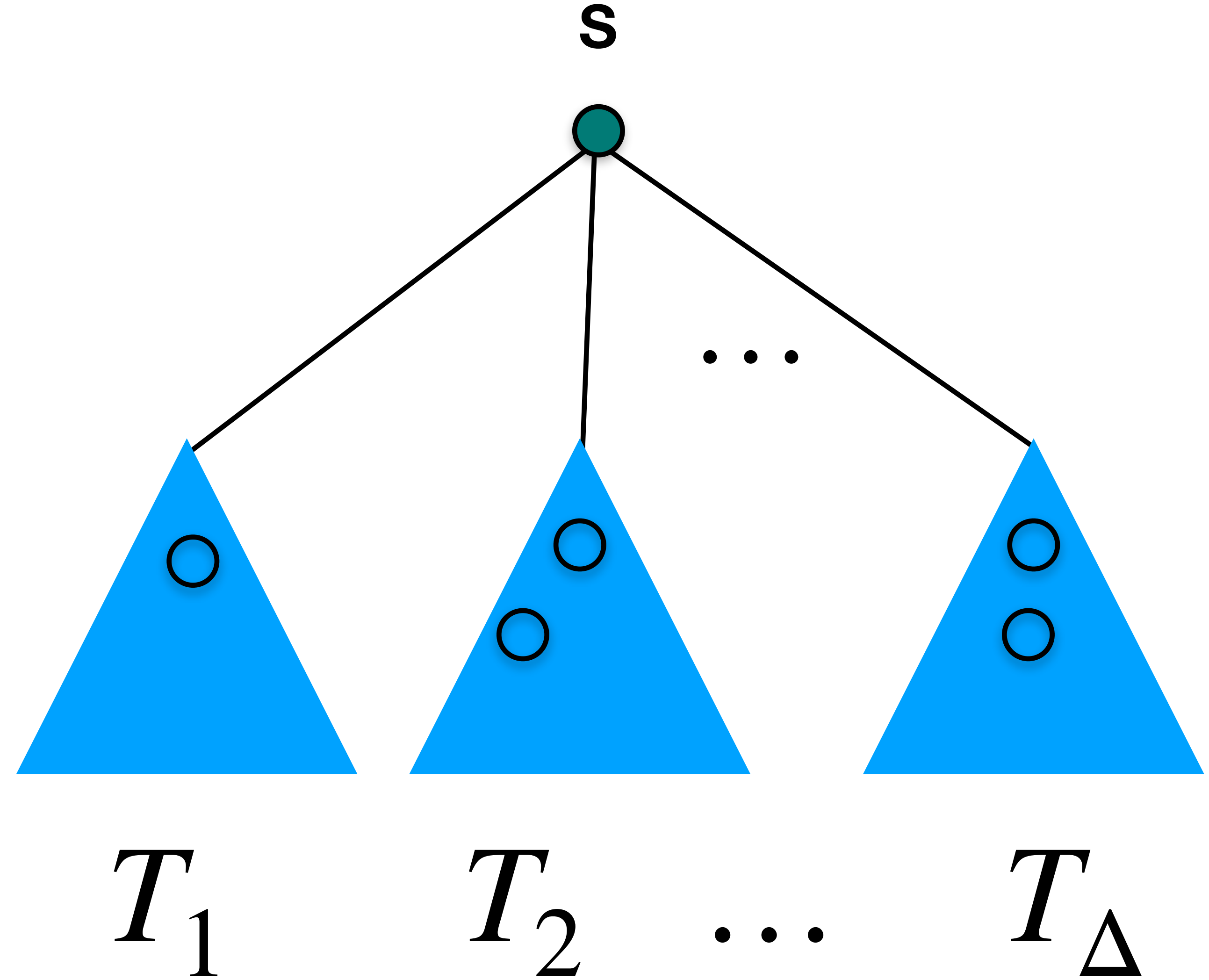
EgoTree: A Greedy Construction

- Sort $= \{p_1, p_2, \ldots, p_k\}$ from large to small
- Greedily Place destination nodes (one at a time)



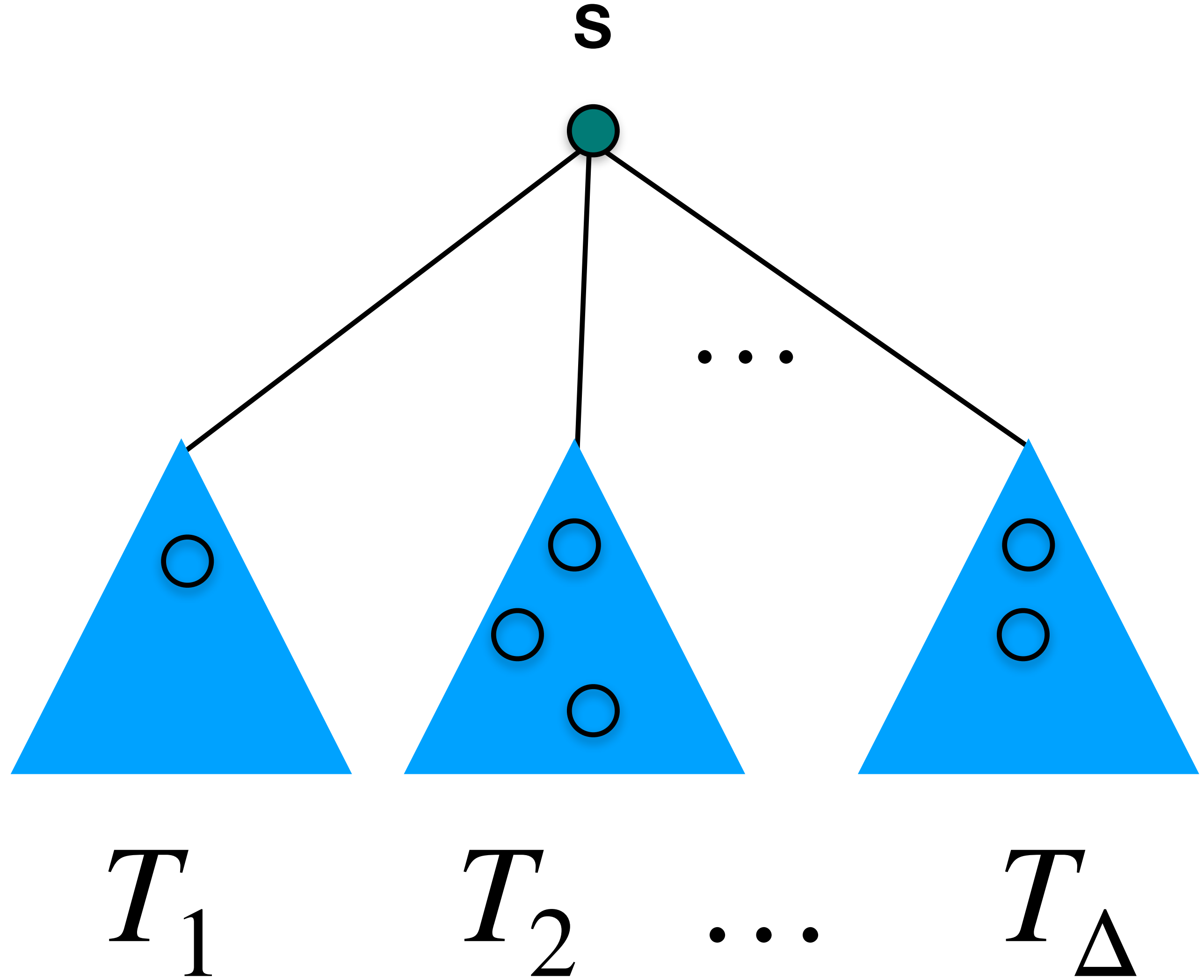
EgoTree: A Greedy Construction

- Sort $= \{p_1, p_2, \ldots, p_k\}$ from large to small
- Greedily Place destination nodes (one at a time)



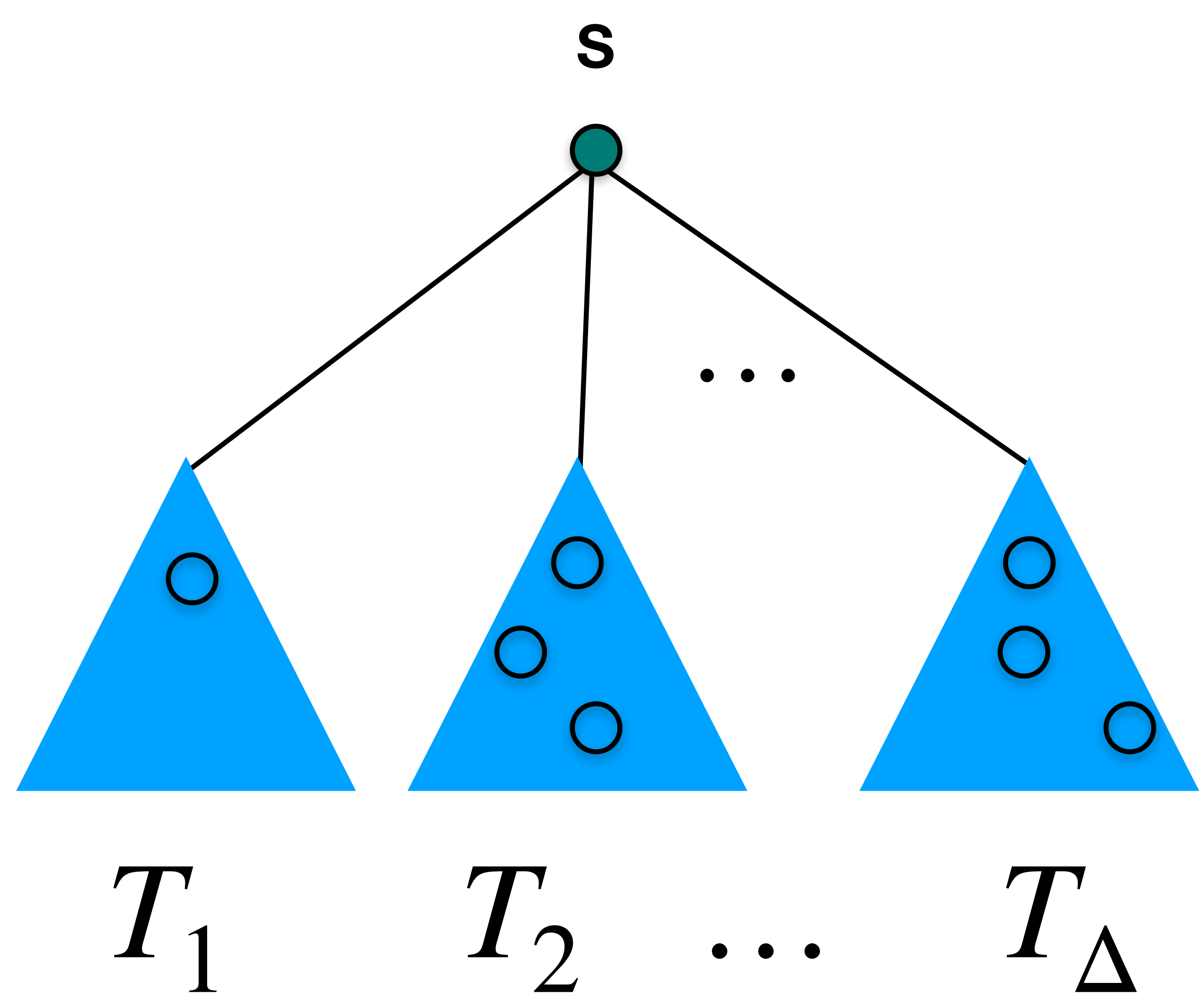
EgoTree: A Greedy Construction

- Sort $= \{p_1, p_2, \ldots, p_k\}$ from large to small
- Greedily Place destination nodes (one at a time)



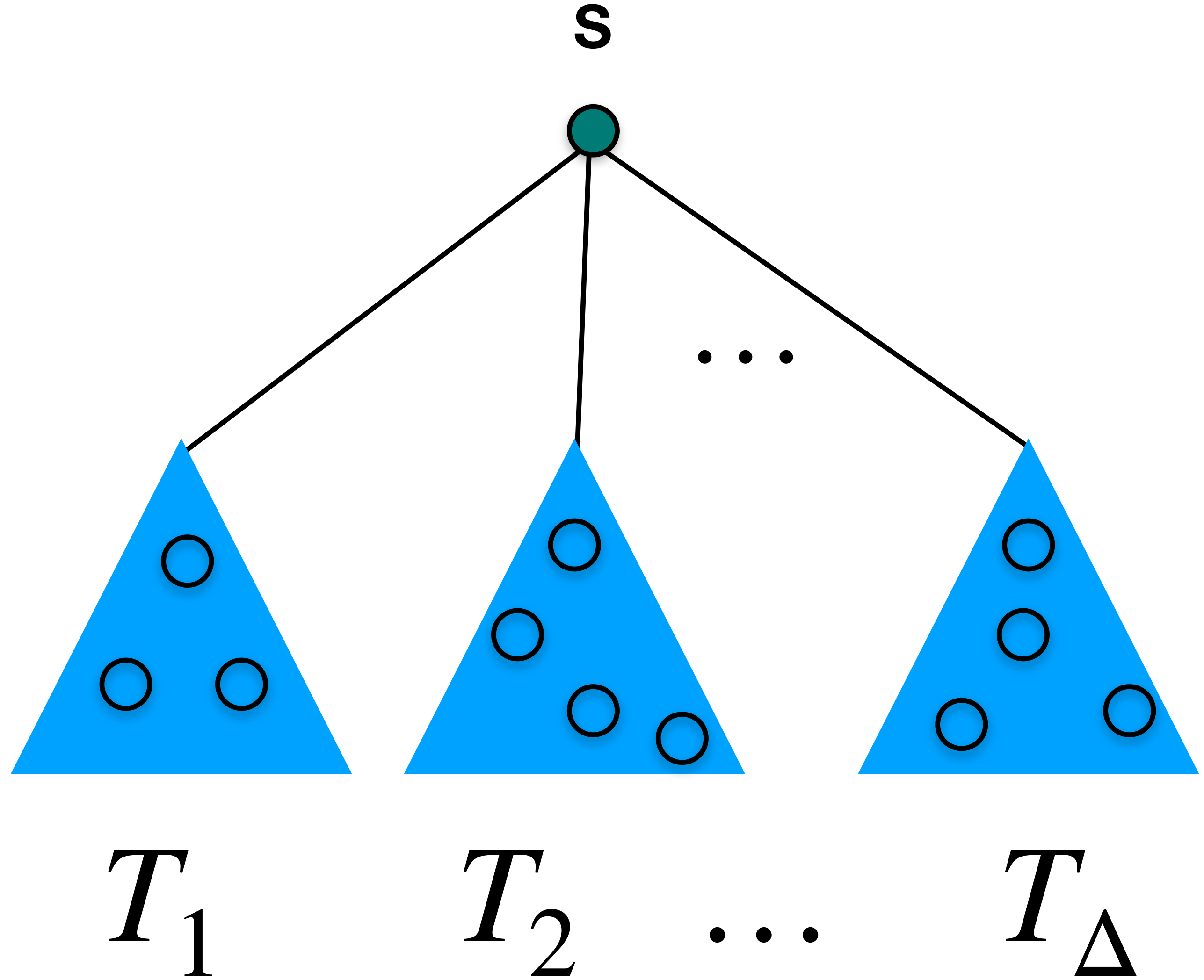
EgoTree: A Greedy Construction

- Sort $= \{p_1, p_2, \ldots, p_k\}$ from large to small
- Greedily Place destination nodes (one at a time)



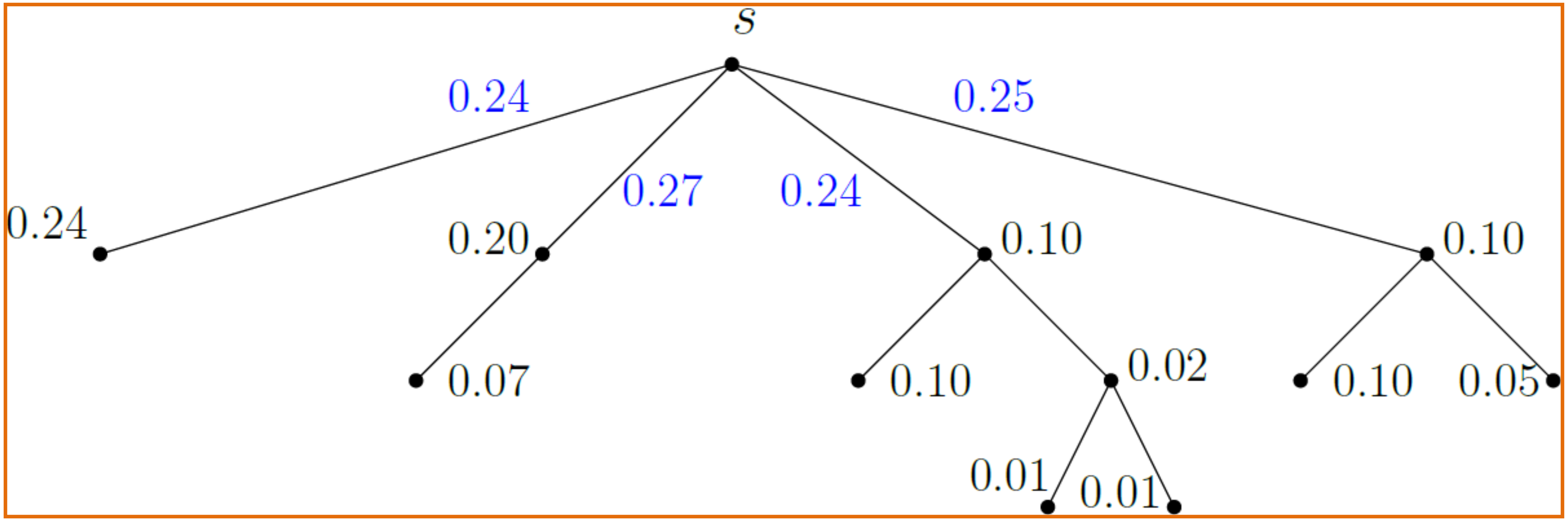
EgoTree: A Greedy Construction

- Sort $= \{p_1, p_2, \ldots, p_k\}$ from large to small
- Greedily Place destination nodes (one at a time)



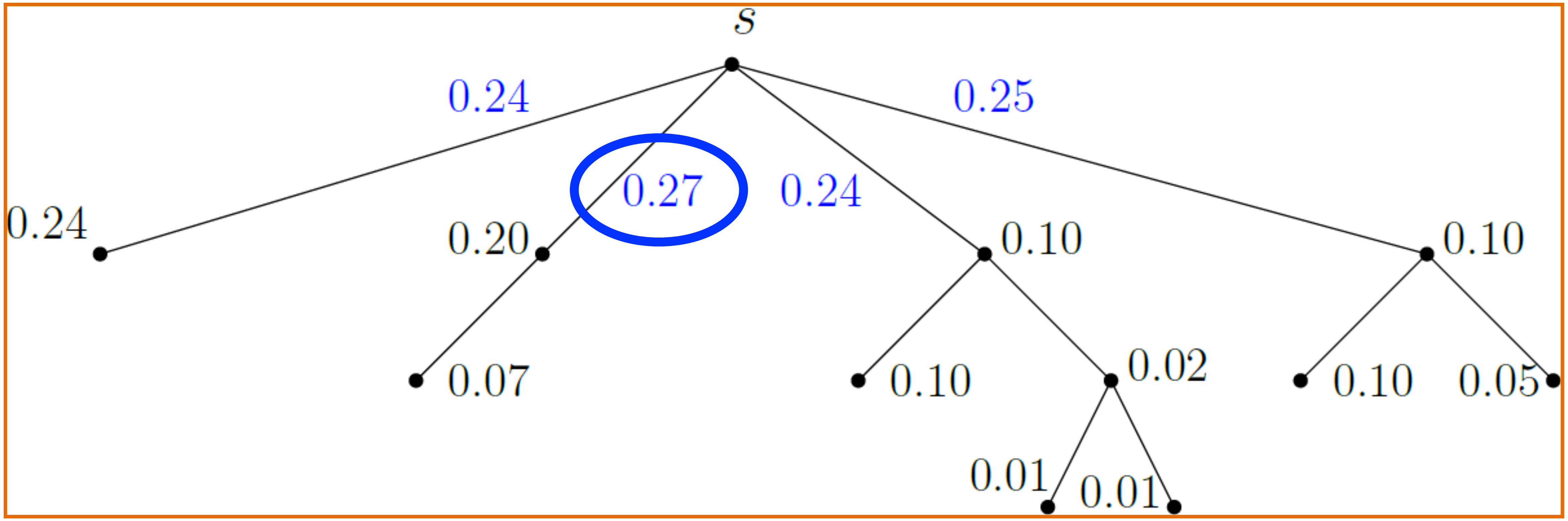
EgoTree: An Example

- Ego-tree is not balanced w.r.t. sizes of the subtrees
- Almost balanced w.r.t. the probability mass that defines congestion



EgoTree: An Example

- Ego-tree is not balanced w.r.t. sizes of the subtrees
- Almost balanced w.r.t. the probability mass that defines congestion



Analysis on Congestion

Analysis on Congestion

- Minimizing congestion is a NP-Hard problem

- $\mathrm{EgoTree}(s, \bar{p}, \Delta)$ provides 4/3 approximation to the minimum congestion (i.e., Longest Processing Time [Graham69])

Analysis on Route-Length

Analysis on Route-Length

- Considering all the binary subtrees and using entropy grouping properties, we have an Entropy bound:

$$\mathrm{L}(\bar{p}, T_s) = 1 + \sum_{i=1}^{\Delta} \mathrm{L}(\Pi_i, T_i) \leq 1 + \sum_{i=1}^{\Delta} S_i H(\Pi_i')$$

$$= 1 + H(\bar{\mathrm{p}}) - H(S_1, S_2, \ldots, S_\Delta) \leq H(\bar{\mathrm{p}})$$

Analysis on Route-Length

- On any optimal Δ-ary tree:

$$\mathrm{L}(\bar{\mathrm{p}}, T_\Delta^*) \geq \frac{1}{\log(\Delta + 1)} H_\Delta(\bar{\mathrm{p}})$$

- Combining all, we now have optimality on L

$$\frac{H(\bar{p})}{\log^2(\Delta + 1)} \leq \mathrm{L}(\bar{p}, T_\Delta^*) \leq \mathrm{L}(\bar{p}, T_s^*) \leq H(\bar{p})$$

Analysis on Route-Length

- On any optimal Δ-ary tree:

$$\mathrm{L}(\bar{\mathrm{p}}, T_\Delta^*) \geq \frac{1}{\log(\Delta + 1)} H_\Delta(\bar{\mathrm{p}})$$

- Combining all, we now have optimality on L

$$\frac{H(\bar{p})}{\log^2(\Delta + 1)} \leq \mathrm{L}(\bar{p}, T_\Delta^*) \leq \mathrm{L}(\bar{p}, T_s^*) \leq H(\bar{p})$$

From Trees to Networks

From Trees to Networks

From Trees to Networks

From Trees to Networks

- Real distributions are **sparse**: datacentre's traffic shows demand distributions are sparse

From Trees to Networks

- Real distributions are **sparse**: datacentre's traffic shows demand distributions are sparse

- We assume a distribution $\mathcal{D}$ with average degree ρ

From Trees to Networks

- Real distributions are **sparse**: datacentre's traffic shows demand distributions are sparse

- We assume a distribution $\mathcal{D}$ with average degree ρ

From Trees to Networks

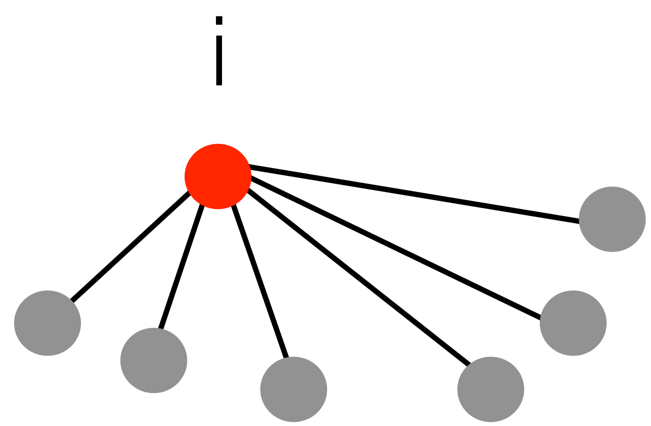
- Real distributions are **sparse**: datacentre's traffic shows demand distributions are sparse
- We assume a distribution $\mathcal{D}$ with average degree ρ
- ρ is a constant

From Trees to Networks

- Real distributions are **sparse**: datacentre's traffic shows demand distributions are sparse
- We assume a distribution $\mathcal{D}$ with average degree ρ
- ρ is a constant
- Half of the nodes of lowest degree are defined as low degree nodes; others are high degree nodes

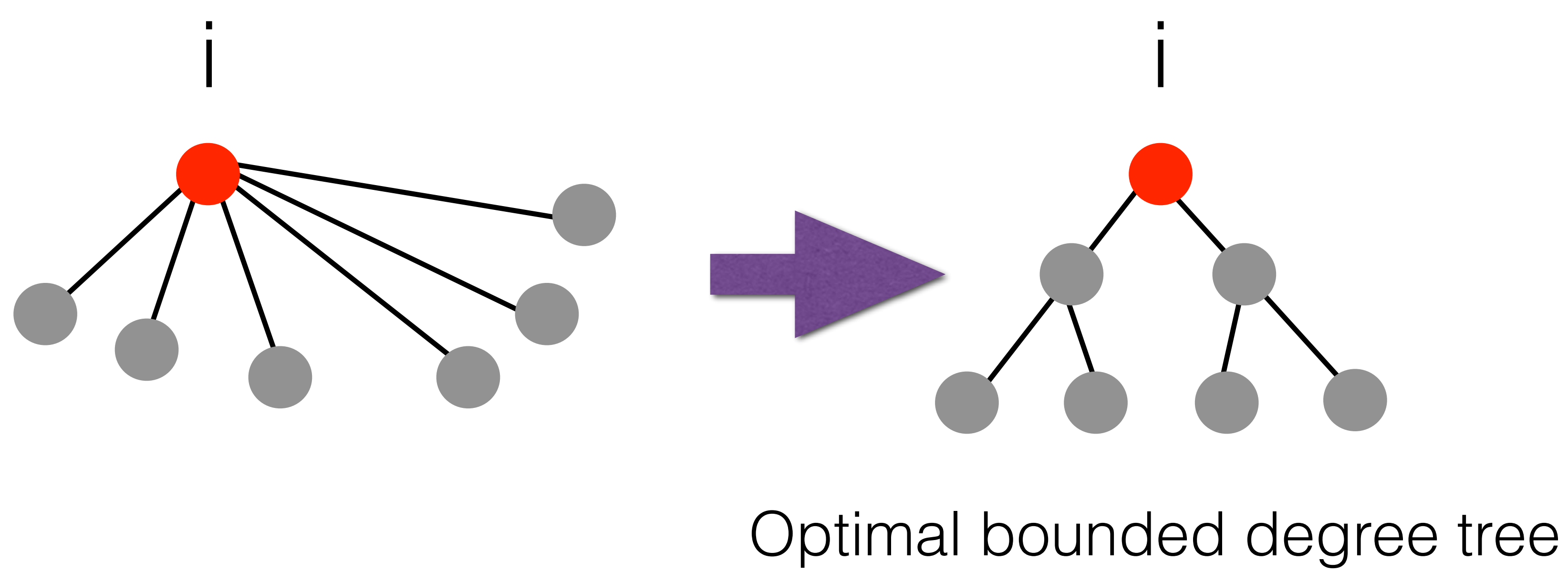
Sparse Distributions

- Proof idea



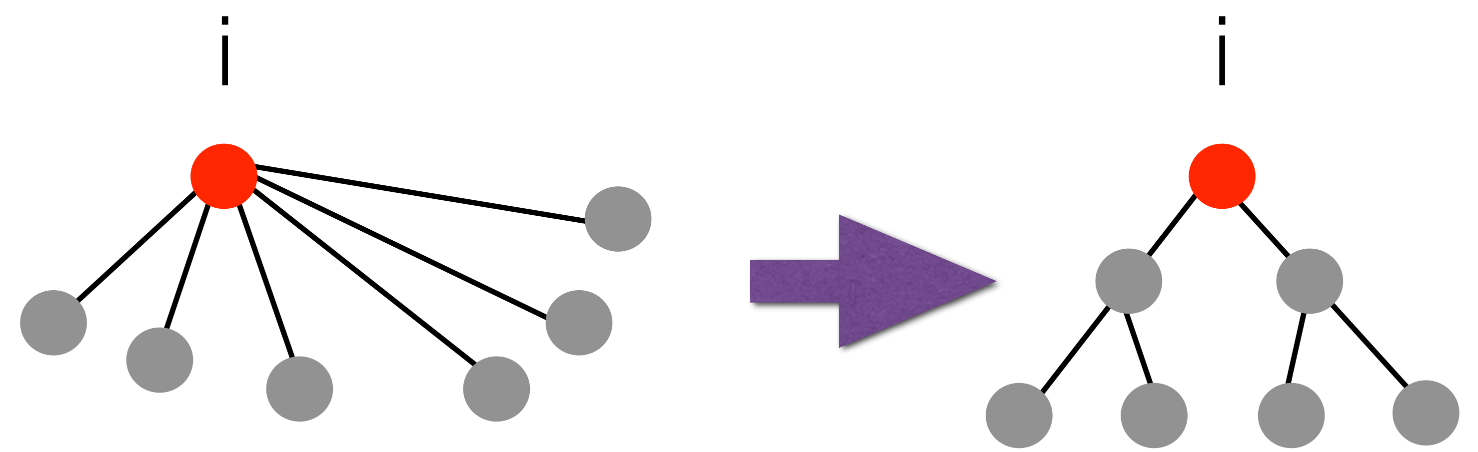
Sparse Distributions

- Proof idea



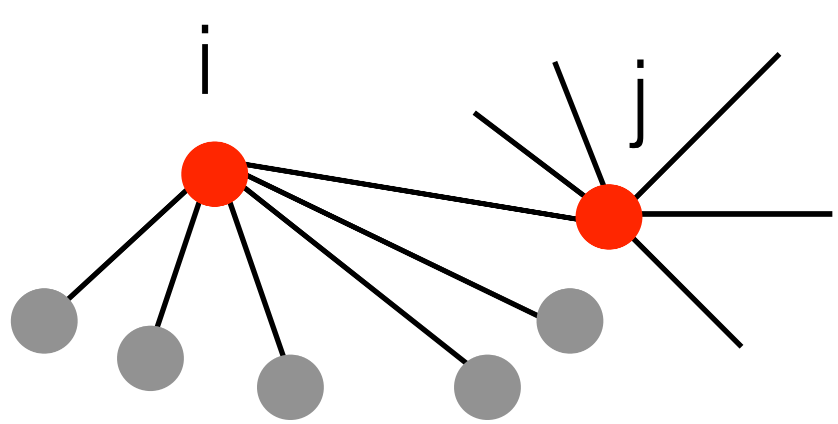
Sparse Distributions

- Proof idea



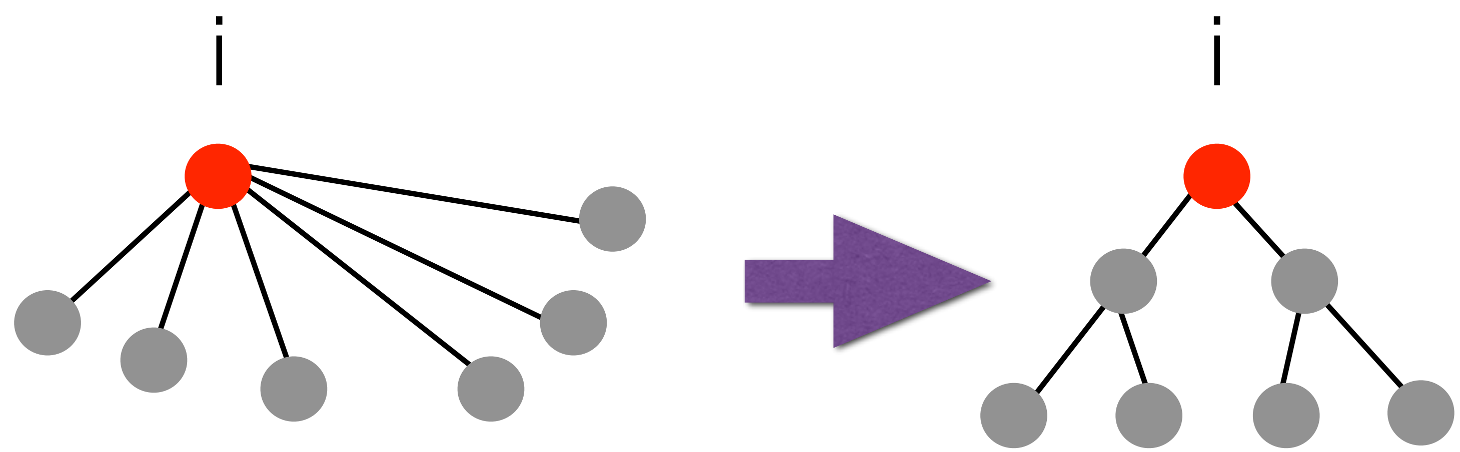
Optimal bounded degree tree

Problem



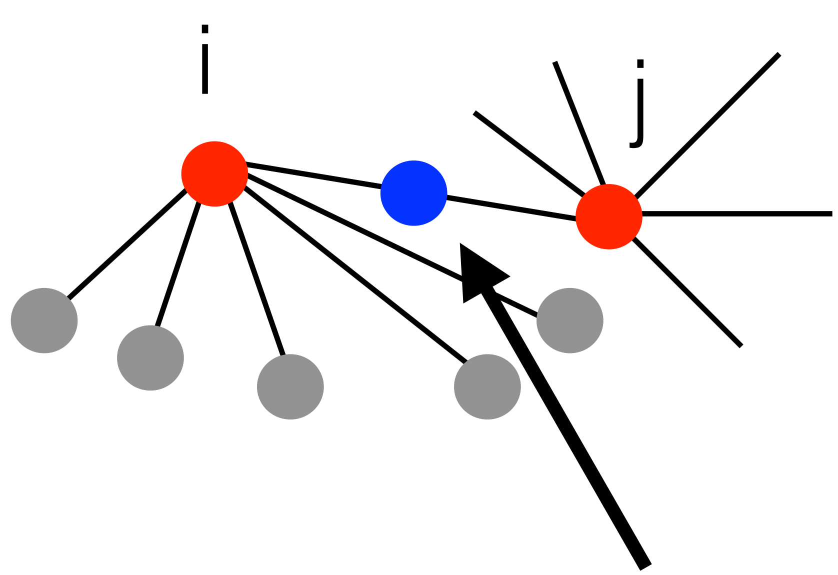
Sparse Distributions

- Proof idea



Problem

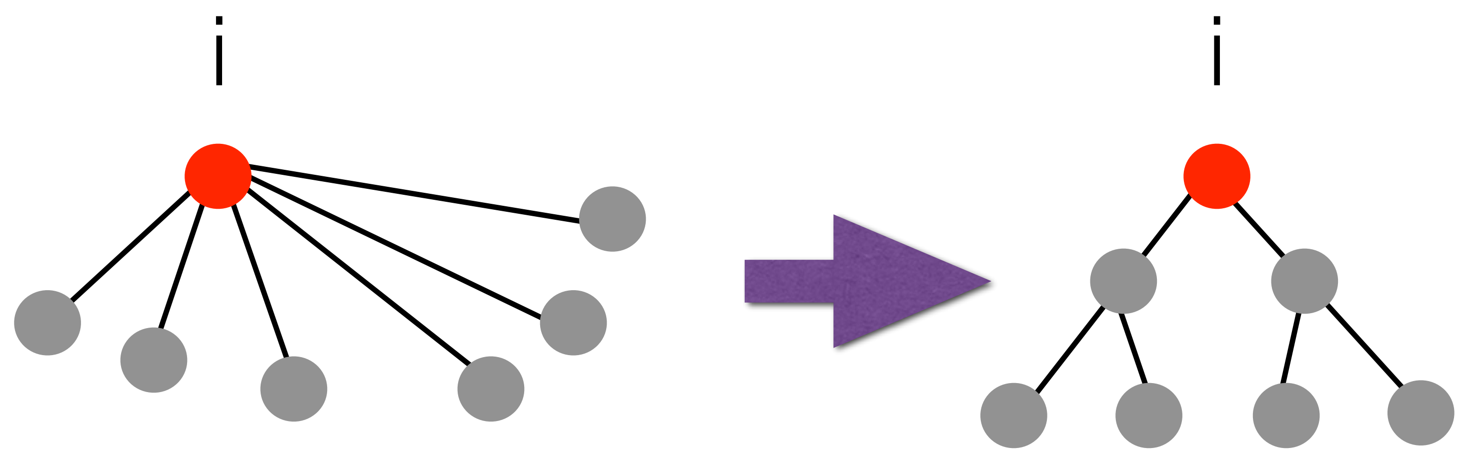
Optimal bounded degree tree



Solution - “helper” nodes

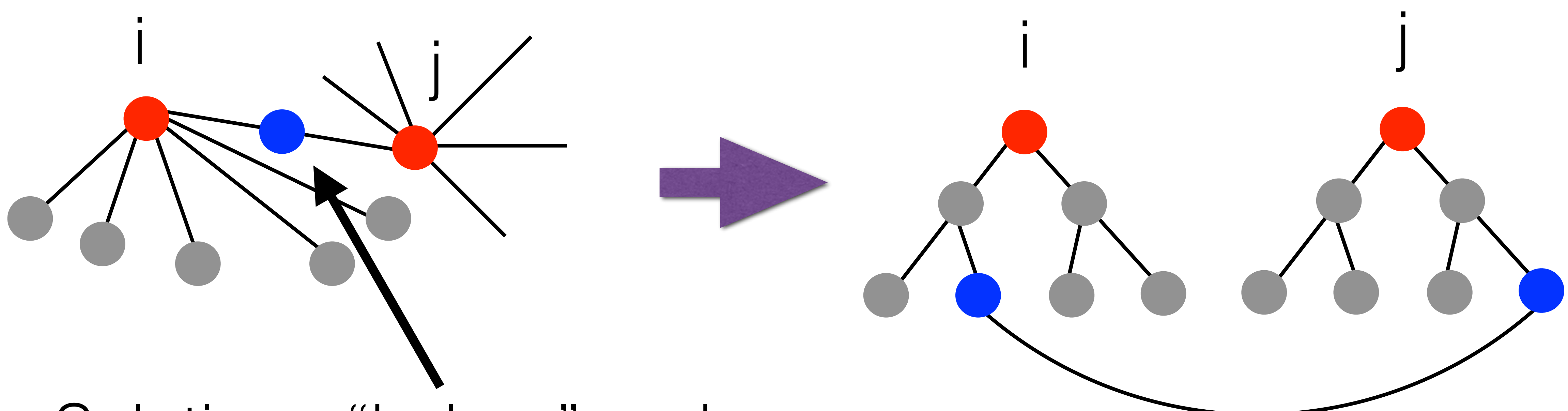
Sparse Distributions

- Proof idea



Optimal bounded degree tree

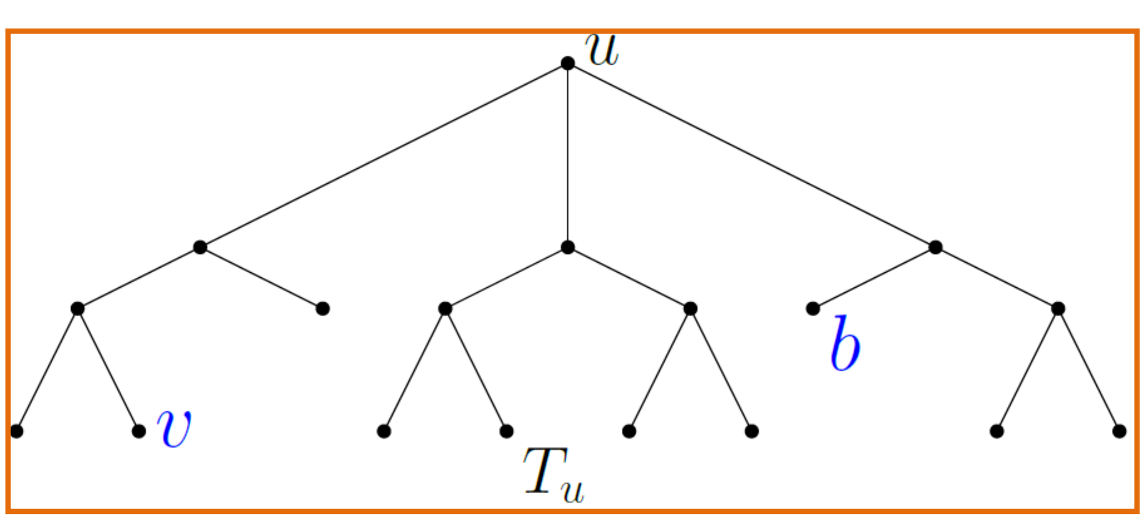
Problem

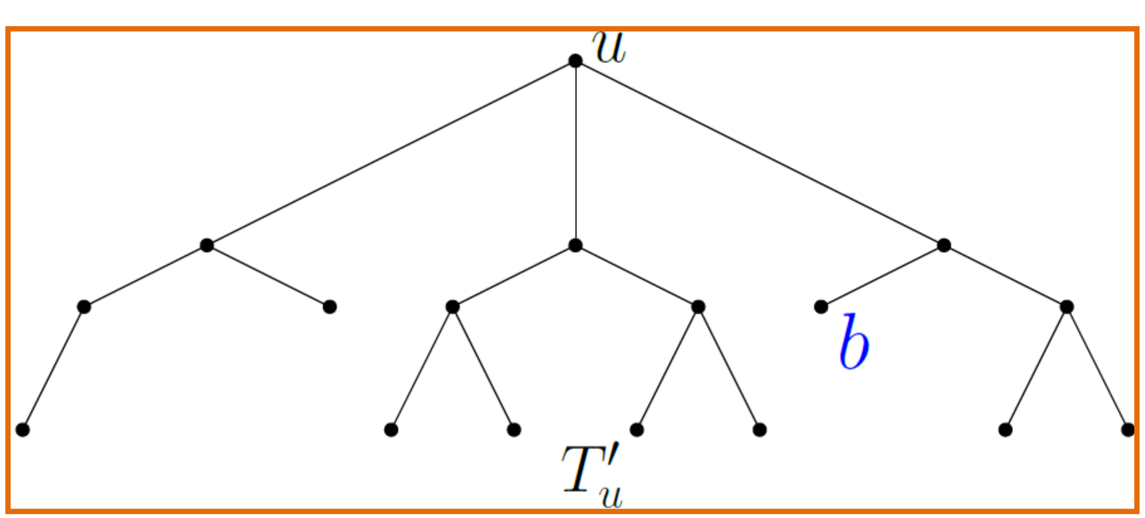


Solution - “helper” nodes

Ego-tree Modification

- Modify ego-tree T_u of a high degree node u to T'_u
  - let v be a high degree neighbor of u: b be the helper node
    - we remove v from ego-tree of u if $p(u,b)>p(u,v)$
    - else we put b in place of v





cl-DANs : Sparse Distributions

cl-DANs : Sparse Distributions

- For Δ=12 ρ

cl-DANs : Sparse Distributions

- For Δ=12 ρ
- Degree of nodes are at most Δ

cl-DANs : Sparse Distributions

- For Δ=12 ρ
- Degree of nodes are at most Δ
- Congestion is at most:

$$C(\mathcal{D}, \Gamma(N)) \leq 1 + 8/9 \Delta C^*(\mathcal{D}, \Delta)$$

c/-DANs : Sparse Distributions

- For Δ=12 ρ
- Degree of nodes are at most Δ
- Congestion is at most:

$$C(\mathcal{D}, \Gamma(N)) \leq 1 + 8/9\Delta C^*(\mathcal{D}, \Delta)$$

cl-DANs : Sparse Distributions

- For Δ=12 ρ
- Degree of nodes are at most Δ
- Congestion is at most:

$$C(\mathcal{D}, \Gamma(N)) \leq 1 + 8/9\Delta C^*(\mathcal{D}, \Delta)$$

- Average Path length is at most

$$\mathrm{L}(\mathcal{D}, \Gamma(N)) \leq 1 + 4\log^2(\Delta+1)\mathrm{L}^*(\mathcal{D}, \Delta)$$

cl-DANs : Sparse Distributions

- For Δ=12 ρ
- Degree of nodes are at most Δ
- Congestion is at most:

$$C(\mathcal{D}, \Gamma(N)) \leq 1 + 8/9\Delta C^*(\mathcal{D}, \Delta)$$

- Average Path length is at most

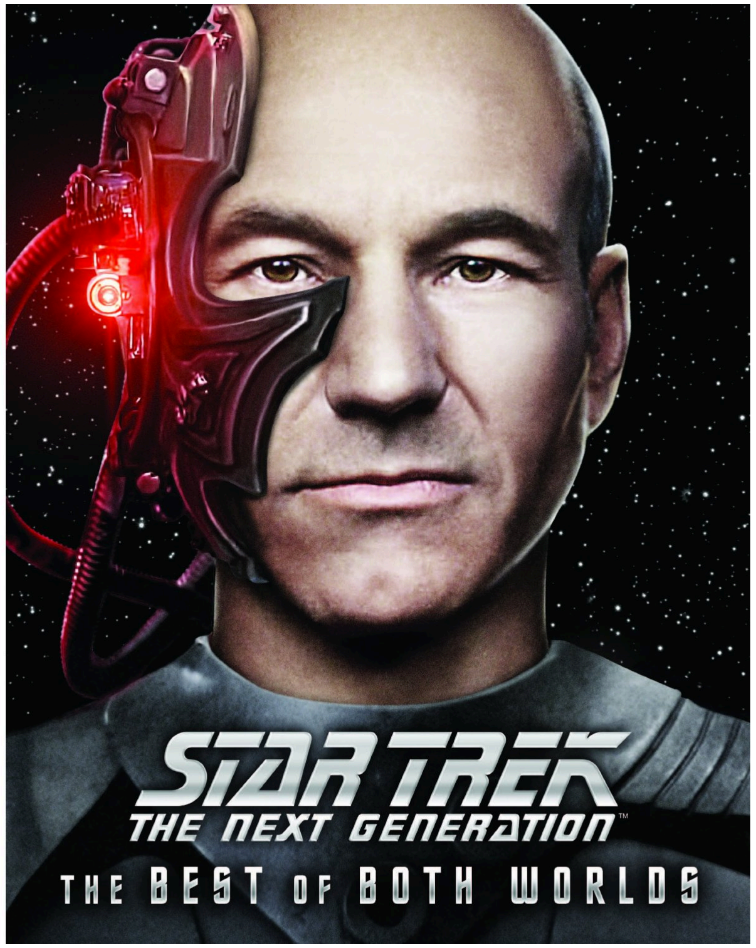
$$\mathrm{L}(\mathcal{D}, \Gamma(N)) \leq 1 + 4\log^2(\Delta + 1)\mathrm{L}^*(\mathcal{D}, \Delta)$$

cl-DANs : Sparse Distributions

- What have we shown?
    - For sparse distributions we can have

c/-DANs : Sparse Distributions

- What have we shown?
    - For sparse distributions we can have

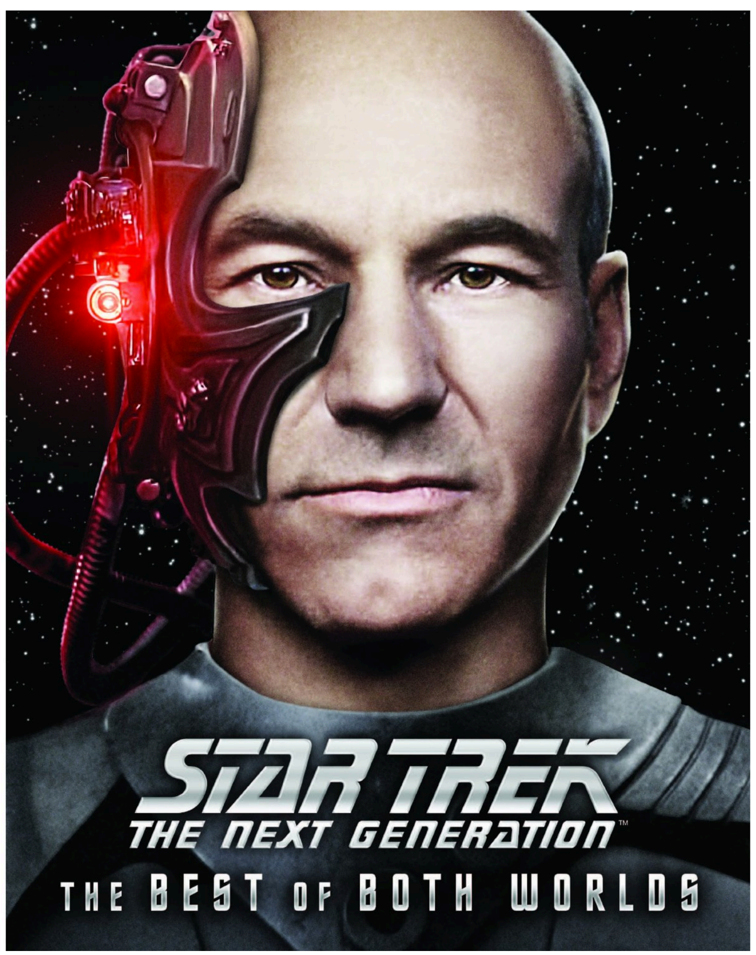


Near optimal congestion

Near optimal path length

cl-DANs : Sparse Distributions

- What have we shown?
  - For sparse distributions we can have



Near optimal congestion

Near optimal path length

Take home

Take home

- Demand Aware Networks are possible

Take home

- Demand Aware Networks are possible
- The “challenge” is **Self-Adjusting Networks**
  - without knowing the future

Take home

- Demand Aware Networks are possible
- The "challenge" is **Self-Adjusting Networks**
  - without knowing the future
- What about:
  - Δ is forced.
  - routing protocols?
  - cost metrics?
  - Sync with upper layers

Take home

- Demand Aware Networks are possible
- The “challenge” is **Self-Adjusting Networks**
  - without knowing the future
- What about:
  - Δ is forced.
  - routing protocols?
  - cost metrics?
  - Sync with upper layers
- Much work and Interesting…

Thank you!

Further Reading

Toward Demand-Aware Networking: A Theory for Self-Adjusting Networks
Chen Avin and Stefan Schmid.
SIGCOMM CCR, October 2018.

Demand-Aware Network Designs of Bounded Degree
Chen Avin, Kaushik Mondal, and Stefan Schmid.
31st International Symposium on Distributed Computing (**DISC**), Vienna, Austria, October 2017.

Online Balanced Repartitioning
Chen Avin, Andreas Loukas, Maciej Pacut, and Stefan Schmid.
30th International Symposium on Distributed Computing (**DISC**), Paris, France, September 2016.

rDAN: Toward Robust Demand-Aware Network Designs
Chen Avin, Alexandr Hercules, Andreas Loukas, and Stefan Schmid.
Information Processing Letters (**IPL**), Elsevier, 2018.

SplayNet: Towards Locally Self-Adjusting Networks
Stefan Schmid, Chen Avin, Christian Scheideler, Michael Borokhovich, Bernhard Haeupler, and Zvi Lotker.
IEEE/ACM Transactions on Networking (**TON**), Volume 24, Issue 3, 2016. Early version: IEEE **IPDPS** 2013.

Demand-aware network design with minimal congestion and route lengths
C. Avin, K. Mondal, and S. Schmid,
IEEE INFOCOM, 2019.

Distributed self-adjusting tree networks
B. Peres, O. Souza, O. Goussevskaia, S. Schmid, and C. Avin,
Proc. IEEE INFOCOM, 2019.

GIF
German-Israeli Foundation for Scientific Research and Development