A Cost-Based Analysis of Overlay Routing Geometries

Nicolas Christin and John Chuang
University of California, Berkeley
School of Information Management and Systems
{christin,chuang}@sims.berkeley.edu

Problem statement

- Evaluate the amount of resources each peer contributes for being part of an overlay network
- Evaluate the benefits associated with participation
- Study independent of specific overlay protocol
  - Graph-theoretic approach
  - Focus on *geometries*: set of nodes and edges (*topology*) associated with a routing algorithm (shortest path in this talk)

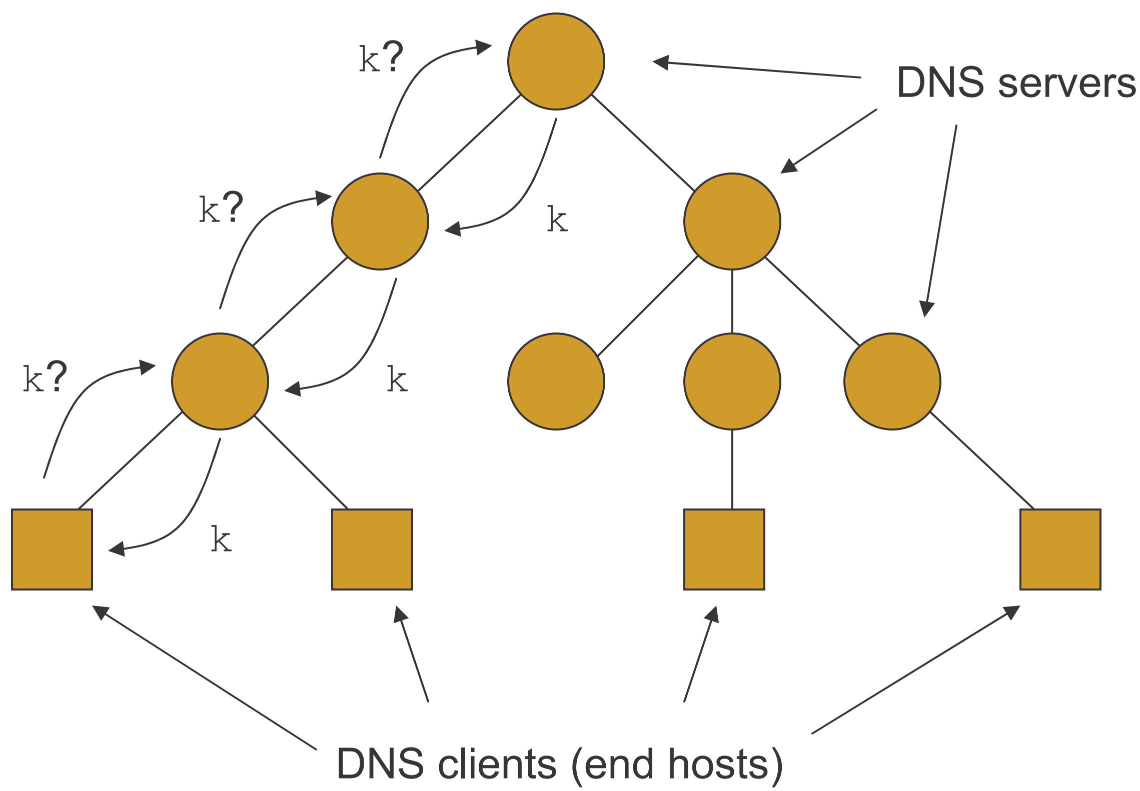
Motivation

- Allows us to predict potential disincentives to collaborate
- Allows us to identify hot spots (e.g., routing)
- Allows us to help design load balancing algorithms
- Benchmark to characterize efficiency of network
  - Can be used to distinguish between proposals for overlays
- Methodology can be applied to other networks
  - e.g., ISP peering relationships

Related work

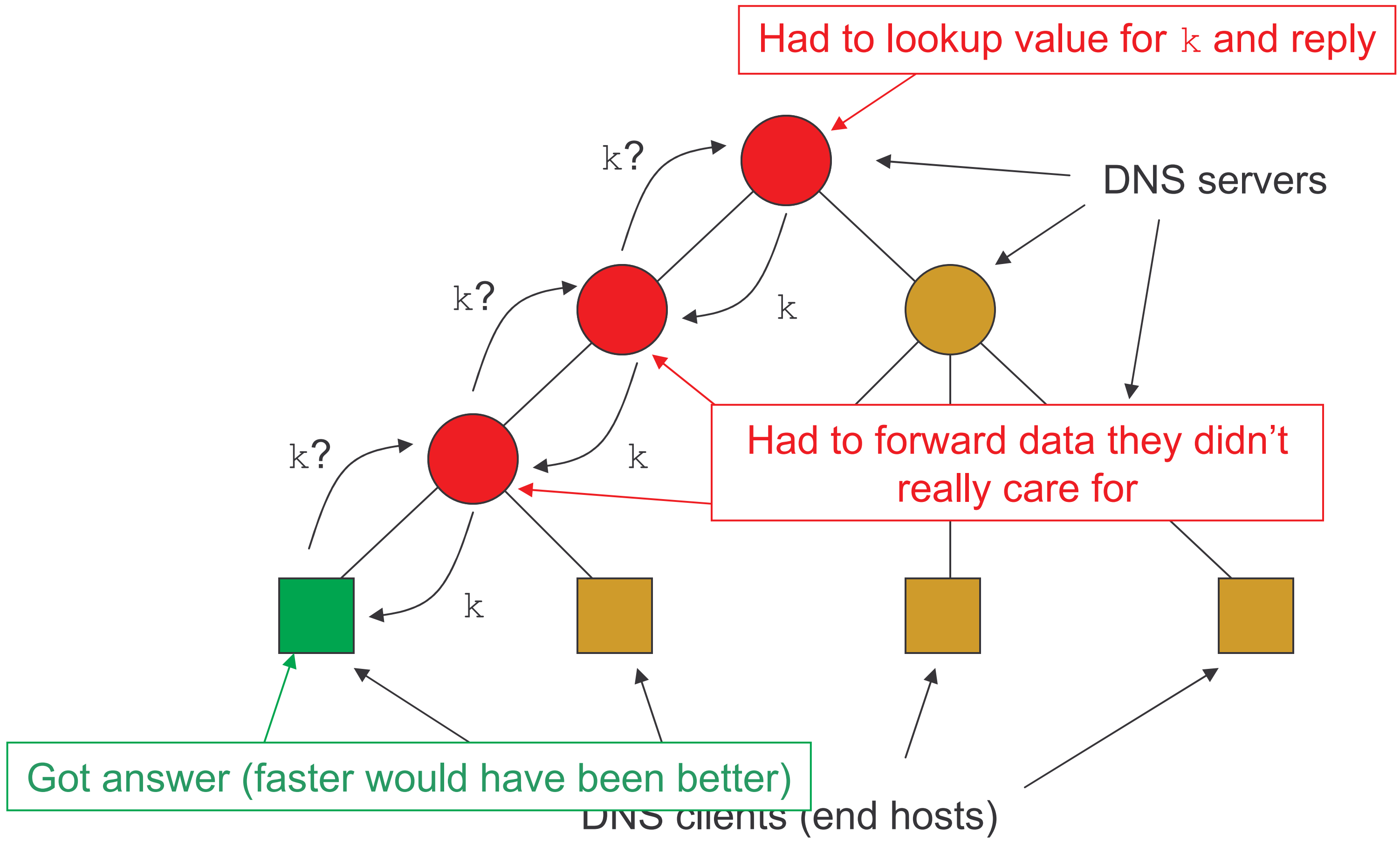
- Cost models
  - Socio-economic networks
    - Jackson and Wolinsky, 1996
  - (Overlay) networks and distributed systems
    - Fabrikant *el al*., 2003
    - Chun *et al*., 2004
  - Generally only consider connectivity (out-degree)
    - Not concerned with service or routing overhead, which may be important factors in a networked system
- Graph-theoretic properties of overlay topologies
  - Loguinov *et al*., 2003
  - Gummadi *et al*., 2003
  - Look at the network as a whole and assume node obedience
    - How about individual nodes?
    - Incentives?

Example: Costs in a DNS lookup



+ all hosts/servers have to maintain records to know where to find "higher" DNS servers when needed

Example: Costs in a DNS lookup



+ all hosts/servers have to maintain records to know where to find "higher" DNS servers when needed

Cost model

- A given node u requests an item, serves a request, or route requests between other nodes
- Latency cost

$$L_u = \sum_{v \in V} \sum_{k \in K_v} l_{u,k} t_{u,v} \,\mathsf{Pr}[Y = k]$$

- Service cost

$$S_u = \sum_{k \in K_u} s_{u,k} \,\mathsf{Pr}[Y = k]$$

- Routing cost

$$R_u = \sum_{v \in V} \sum_{w \in V} \sum_{k \in K_w} r_{u,k} \,\mathsf{Pr}[X = v] \,\mathsf{Pr}[Y = k] \chi_{v,w}(u)$$

- Maintenance cost

$$M_u = m_u \,\mathsf{deg}(u)$$

Individual and total cost

- Individual cost of node u
  - Sum of latency, routing, service and maintenance costs at node u

$$C_u = L_u + S_u + R_u + M_u$$

- Total cost (of the whole network)
  - Sum of all individual costs

$$C = \sum_u C_u$$

Analysis assumptions

- Homogeneous peers and homogeneous links (i.e., for any u and k, $l_{u,k} = l$, $s_{u,k} = s$, $r_{u,k} = r$ and $m_u = m$.)
- Steady-state regime (i.e., no churn)
- Sources of requests uniformly distributed over the set of nodes (i.e., $\Pr[X = u] = 1/N$)
- Destinations of requests uniformly distributed over the set of nodes (implies $S_u = s/N$)
- Quite idealistic!

Social optimum

- Social optimum: geometry that minimizes total (network) cost C
  - Ideal geometry from the system perspective $m \leq l/N + r/N^2$
  - If number of nodes N is small and/or maintenance operations come cheap (i.e., m is small): **full mesh**
  - Otherwise: **star network**
    - Always local optimum, social optimum if links are bidirectional
  - Proof sketch: start from full mesh, and remove links

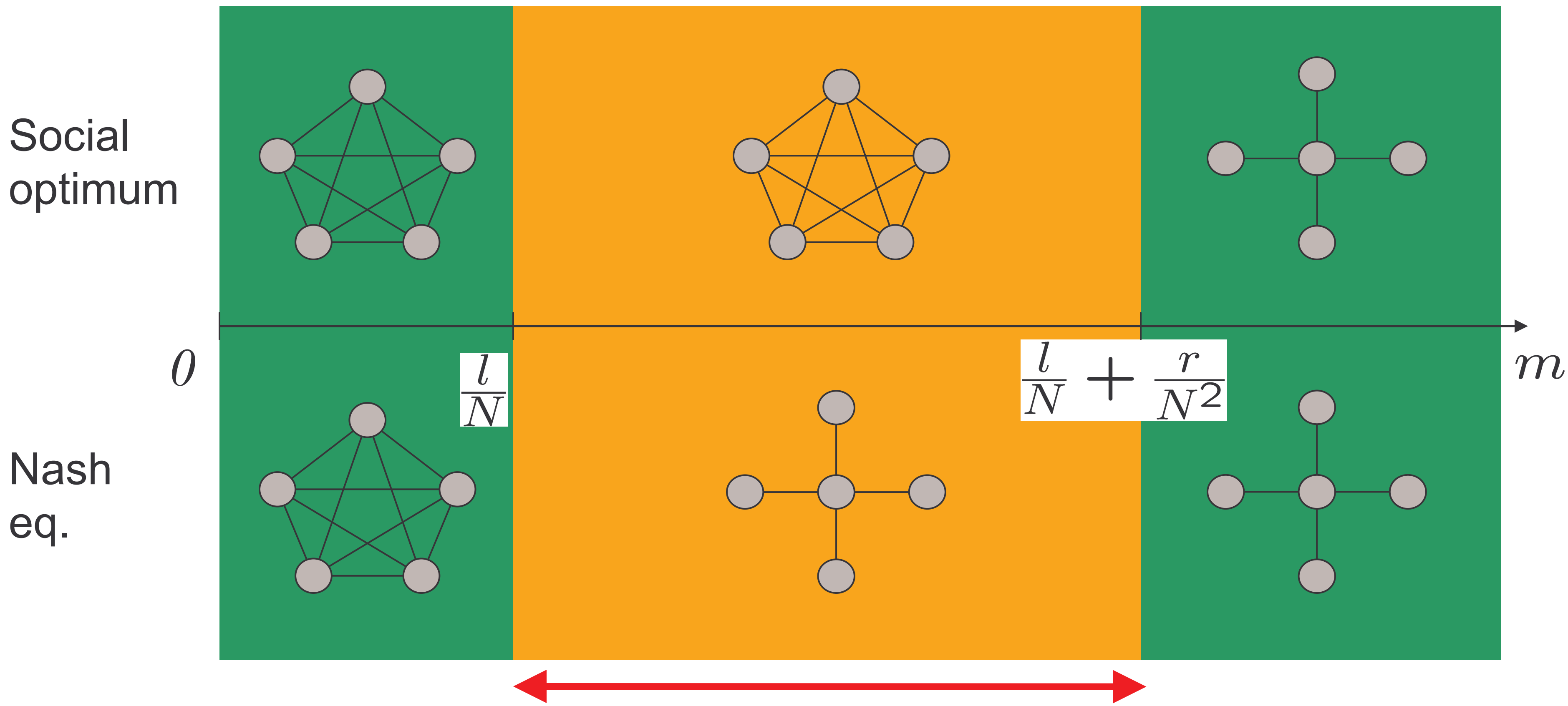
Nash equilibrium

- (Pure) Nash equilibrium: geometry in which no individual node u can decrease its individual cost C_u by (deterministically) creating or removing a link

  $m \leq l/N$

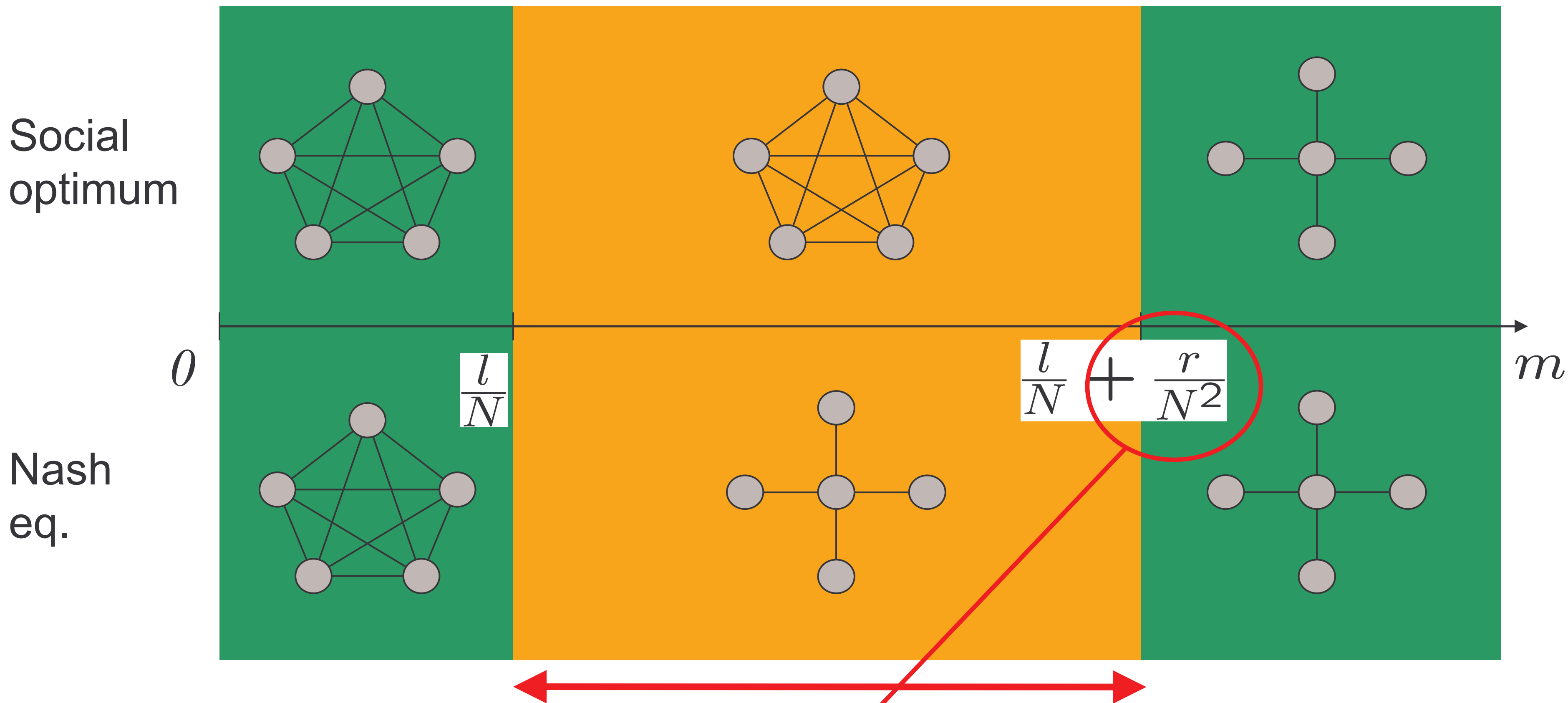
  - Ideal geometry from a selfish node's perspective
  - If number of nodes N is small and/or maintenance operations come cheap (i.e., m is small): **full mesh**
  - Otherwise: **star network** (not necessarily unique)
  - Proof sketch: start from topology, try to add and remove links

Social optimum vs. Nash equilibrium



Individual incentives not aligned with desirable social outcome

Social optimum vs. Nash equilibrium



Conflicts *only* arise from transporting data for others

The need for rules

- Routing data for others is generally costly (i.e., $r \gg 0$)
  - Social optimum and Nash equilibrium differ
- Optimal topologies (star, full mesh) can be impractical
  - Lack of resiliency or scalability

- Need for rules to realign incentives, ensure resiliency...
  - Monetary compensation
  - Protocol
  - Geometry/Topology
  - …

Topological rules: DHT geometries

- Spread load evenly on all nodes in the network while keeping acceptable overall performance

- How do DHT geometries compare with social optimum/Nash equilibrium?
- Are DHT geometries effective at avoiding blatant individual disincentives?

- Analyzed
  - PRR trees (Pastry, Tapestry, Bamboo, …)
  - *D*-tori (CAN)
  - de Bruijn graphs (Koorde, ODRI, Distance-Halving)

DHT geometries analysis

- Closed form expressions can be derived

- D-tori (CAN)
  - $L_u = l\frac{DN^{1/D}}{4}$
  - $R_u = r\frac{\rho_{u,D}}{N^2}$
  - $M_u = 2mD$

- PRR trees (Pastry, Tapestry, ...)
  - $L_u = l\frac{D\Delta^{D-1}(\Delta-1)}{N}$
  - $R_u = r\frac{\Delta^{D-1}(D(\Delta-1)-\Delta)+1}{N^2}$
  - $M_u = mD(\Delta-1)$

- Same results for Chord rings (with $\Delta=2$)

Asymmetry in de Bruijn graphs

- Different nodes have different latency costs

$$L_{\text{min}} \le L_u \le L_{\text{max}}$$

- Different nodes have different routing costs

$$0 \le R_u \le r\rho_{\text{max}}/N^2$$

- Different nodes have different maintenance costs

$$M_u = m\Delta \quad \text{or} \quad M_u = m(\Delta - 1)$$

Asymmetry in de Bruijn graphs (cnt'd)

Δ: alphabet size
D: network diameter

$$L_{\text{max}} = \max_u L_u$$

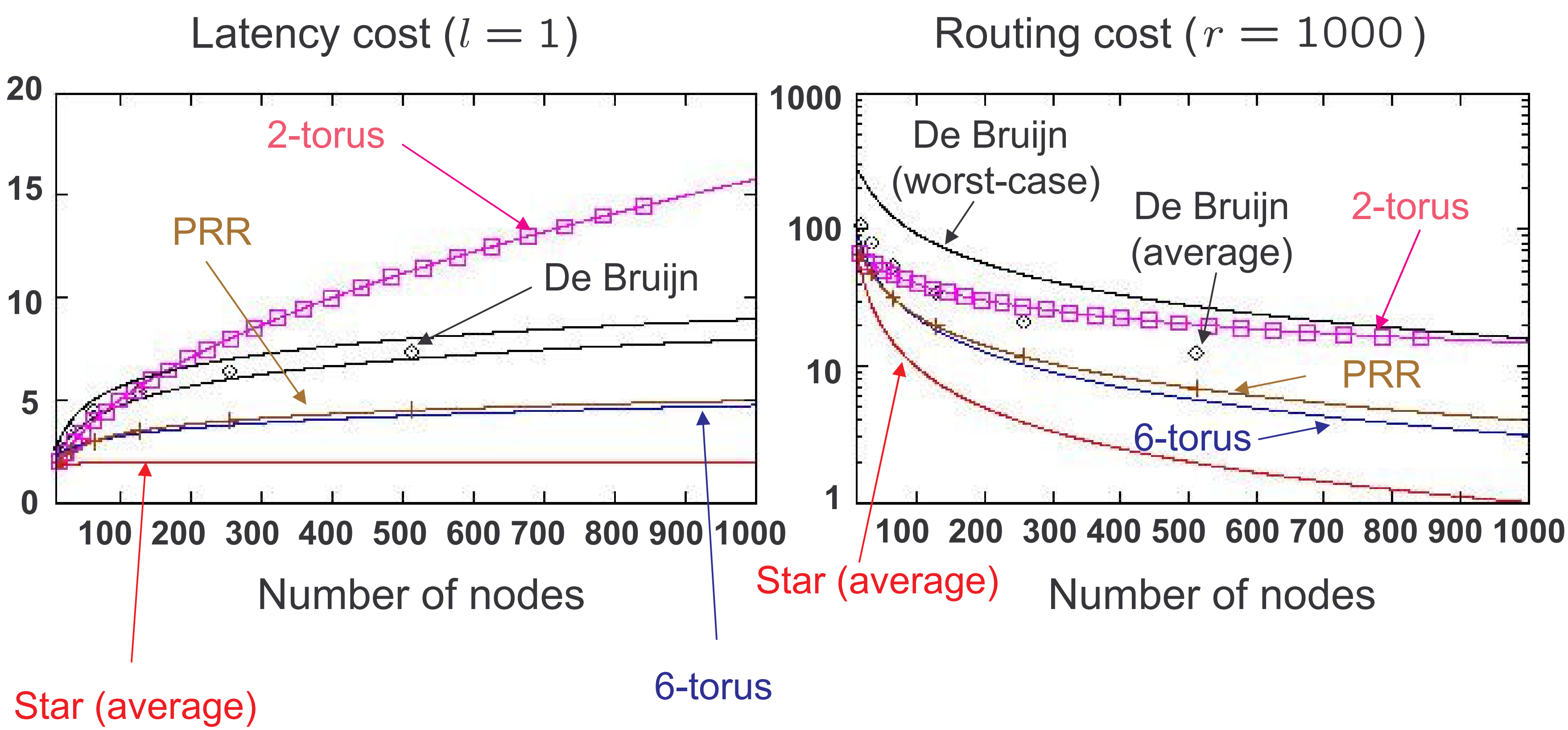
$$L_{\text{min}} = \min_u L_u$$

$$R_{\text{max}} = \max_u R_u$$

$$R'_{\text{min}} = \min_u \{R_u : R_u > 0\}$$

(Δ, D)	$\frac{L_{\text{max}}}{L_{\text{min}}}$	$\frac{R_{\text{max}}}{R'_{\text{min}}}$
(2,9)	**1.11**	**4.51**
(3,6)	**1.04**	**4.41**
(4,4)	**1.03**	**2.71**
(5,4)	**1.02**	**2.78**
(6,3)	**1.01**	**1.86**

Routing and latency costs



Numerical results

- Analysis relies on very stringent set of assumptions
- Use simulations to evaluate impact of
  - Asymmetry in item popularity on individual costs
  - Sparse population of the identifier space
    - e.g., Pastry has 2^{128} available identifiers, so that the number of nodes in the system at any time is $N \ll 2^{128}$
- 1,024 experimental runs
- 100,000 requests per run

Asymmetry in item popularity

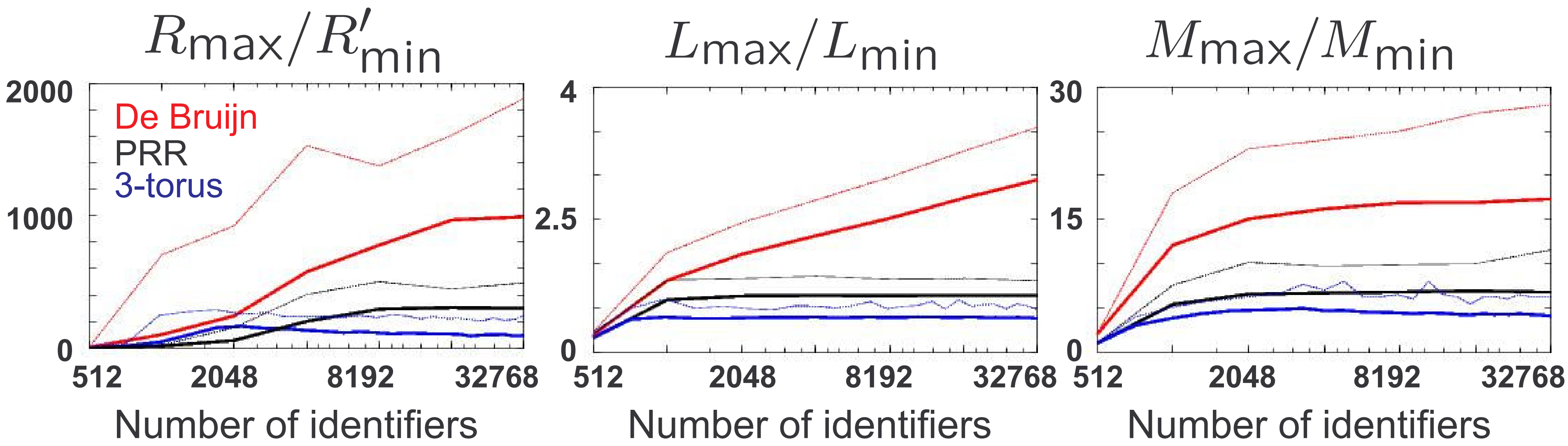
- Item popularity follows a Zipf distribution with $\alpha = 0.75$

	$L_{\text{max}}/L_{\text{min}}$	$R_{\text{max}}/R'_{\text{min}}$
3-torus	**1.27** (± 0.04)	**5.28** (± 0.35)
De Bruijn	**1.25** (± 0.02)	**30.73** (± 9.6)
PRR	**1.26** (± 0.04)	**9.22** (± 0.66)

- Little or no correlation between the different costs (see paper)
  - Some nodes just get a "rotten deal"

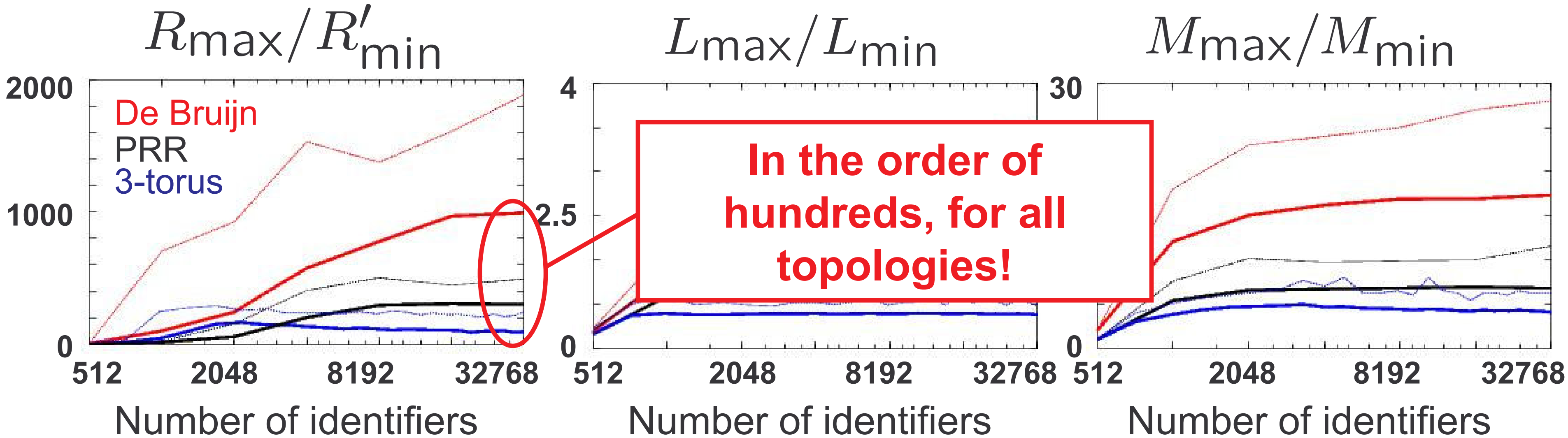
Sparse population of the ID space

- 512 nodes
- Vary number of identifiers
- Assign each unused identifier to closest node



Sparse population of the ID space

- 512 nodes
- Vary number of identifiers
- Assign each unused identifier to closest node



Summary

- General cost model for participating in overlay
  - Takes into account routing, latency, maintenance and service costs
  - Probably applicable beyond overlays (ISP-ISP peering?)
- Notion of routing cost is important
  - Explains why individual incentives are not necessarily aligned with overall welfare
- May need rules to realign incentives
  - DHT geometries: Implementing rules can be tough
    - Very balanced geometries in theory
    - Potentially large imbalances (esp. routing) in practice

Questions