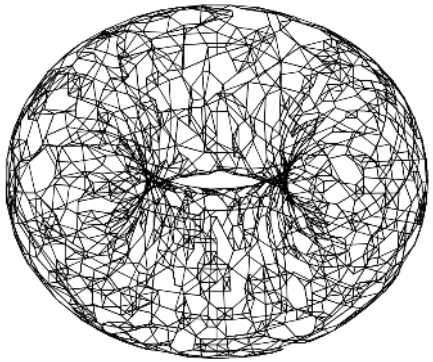
ASAP

Polystyrene: The Decentralized Data Shape that Never Dies



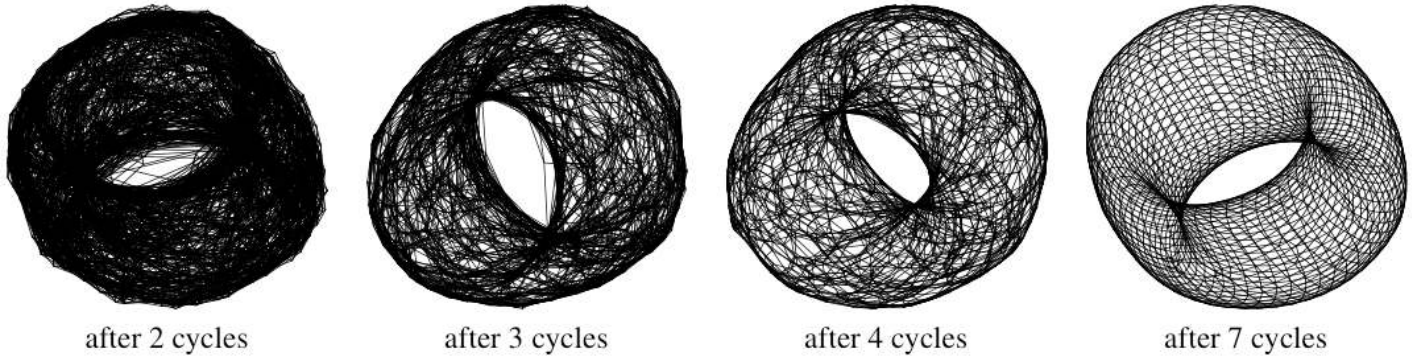
François Taïani

Joint work with:

- Anne Marie Kermarrec (Inria)
- Simon Bouget (ENS Cachan)
- Hoel Kervadec (INSA Rennes)

Focus

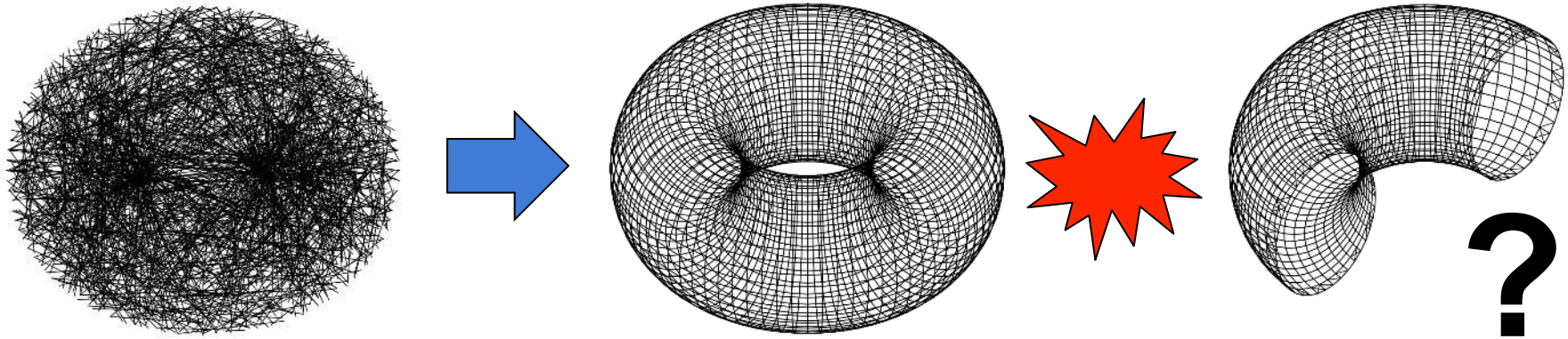
- Epidemic Topology Construction algorithms



Taken from [JMB09]

- ➔ Decentralized, fast, scalable
- ➔ Fundamental building block to higher-level services (DHT, Multicast, Pub-Sub, Recommendations)

Problem: Catastrophic Failure



- The topology **heals**
- But the overall **shape** is **lost**

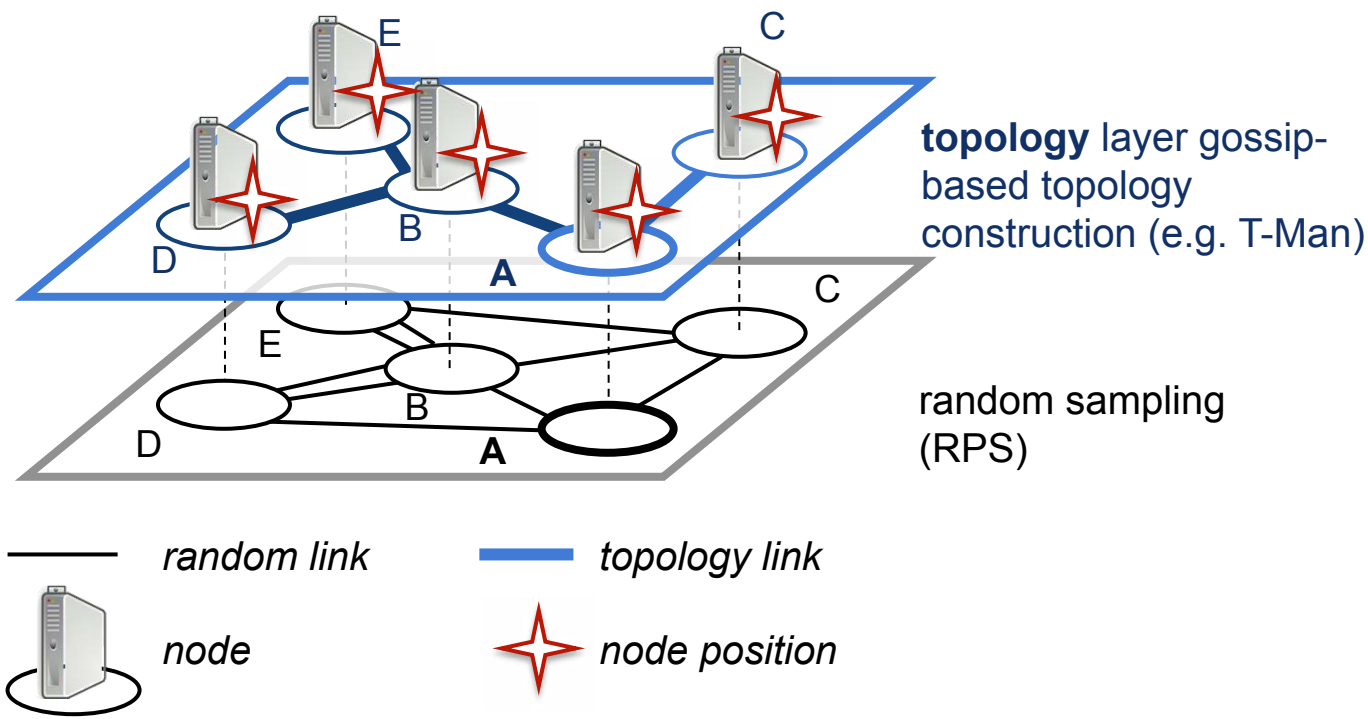
How to recreate whole shape from surviving nodes?

Outline

- **Background**:
  Decentralized Topology Construction
- **Polystyrene**:
  Architecture and Protocol
- **Evaluation**
- **Outlook**

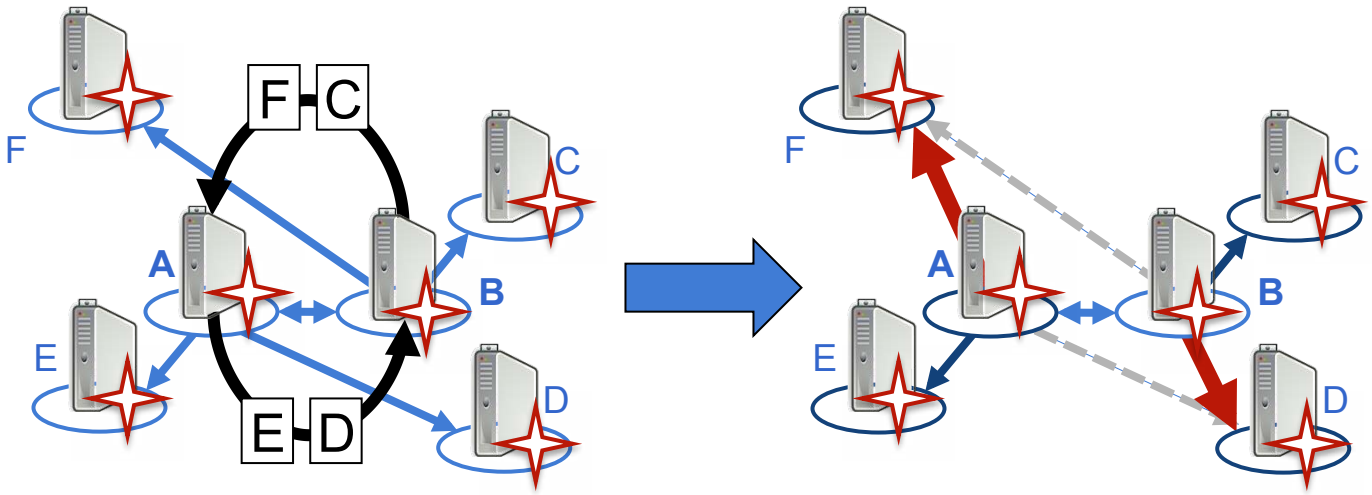
Decentralized Topology Const.

- Each node : some data
- Find k “closest” nodes in system
  - ➔ Decentralized approach, asynchronous rounds



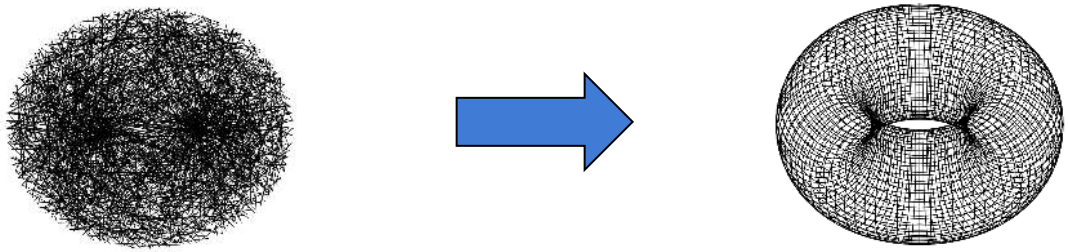
Decentralized Topology Const.

- Main idea: **greedy neighbourhood optimization**

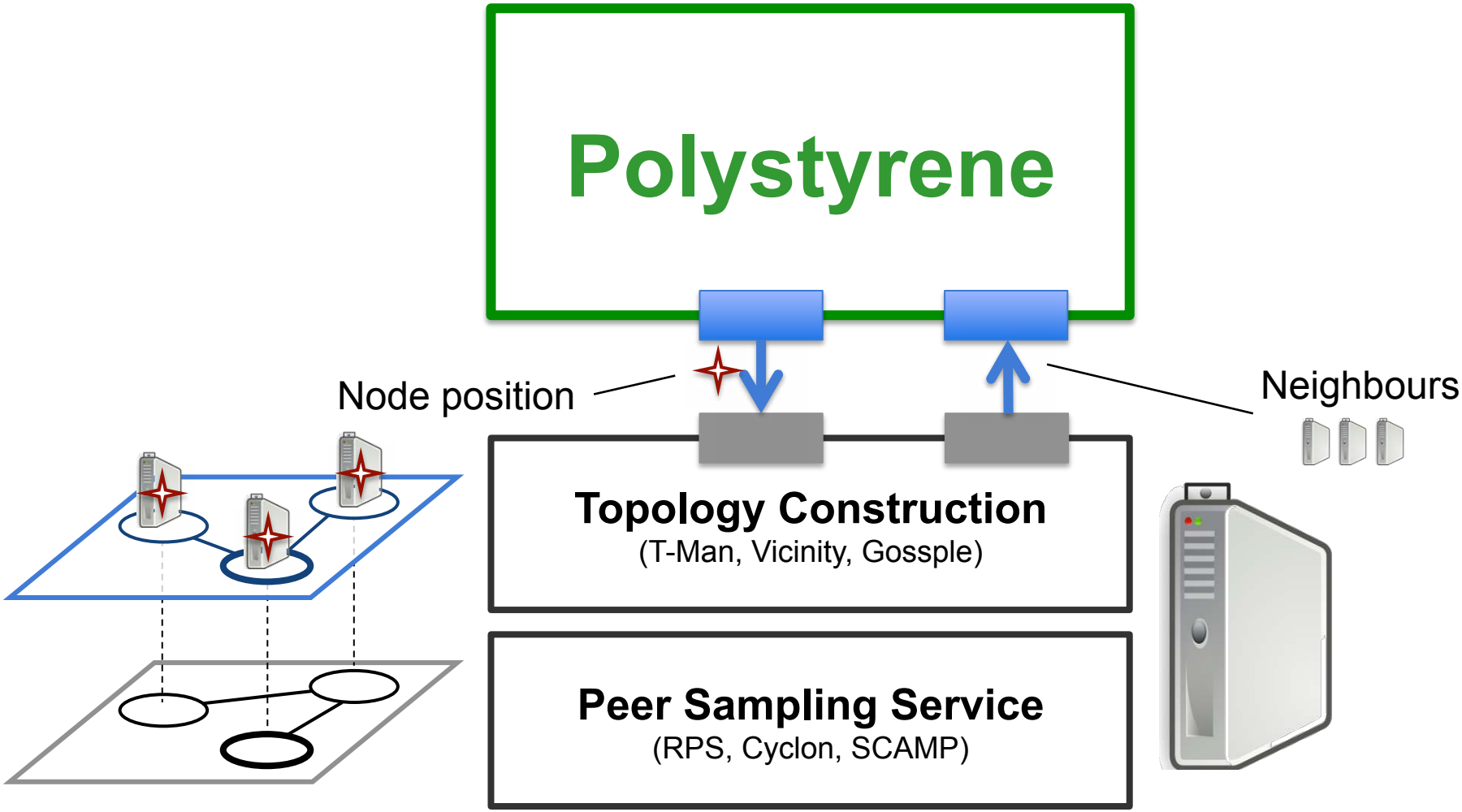


❶ exchange of neighbors lists

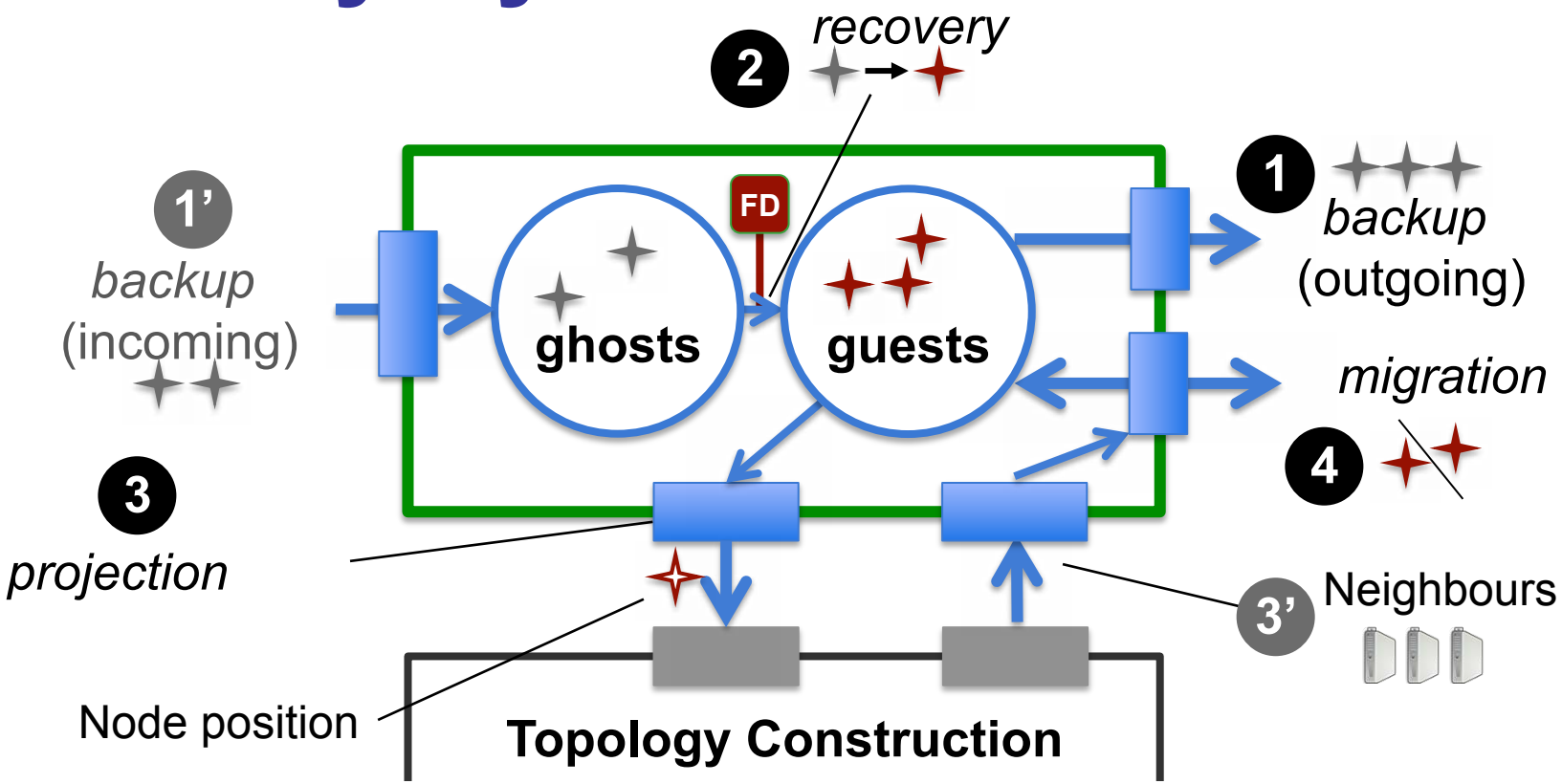
❷ neighborhood optimization



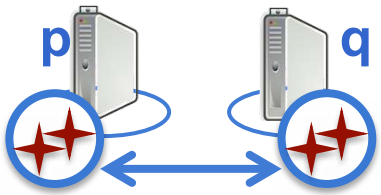
Polystyrene's Architecture

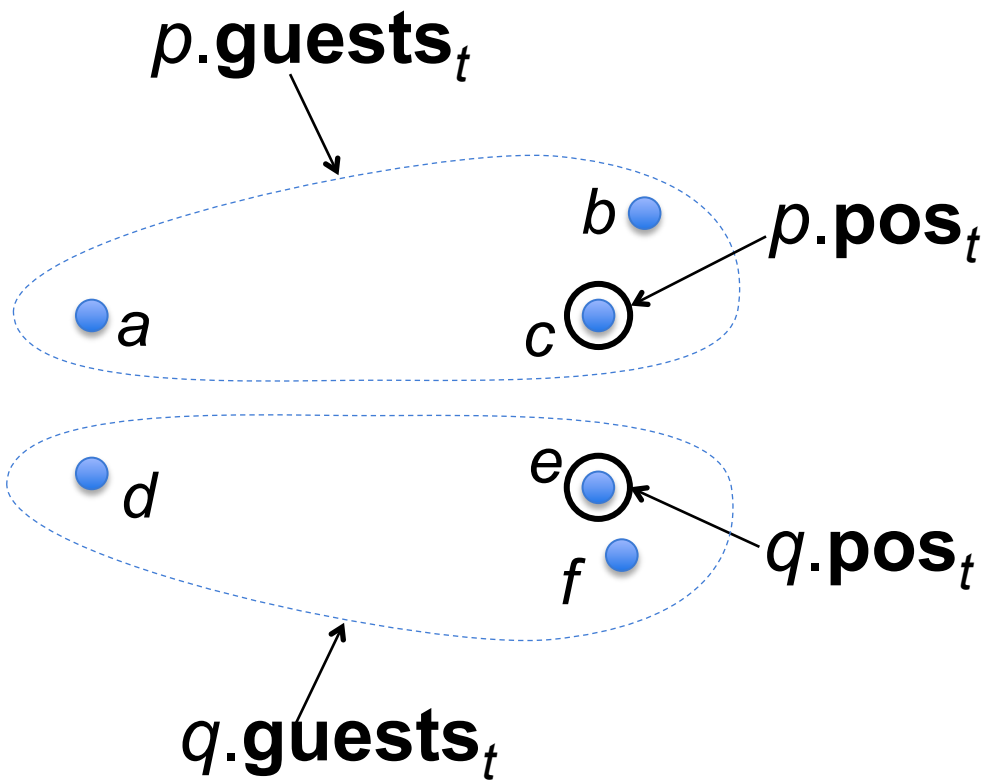


Polystyrene Protocol

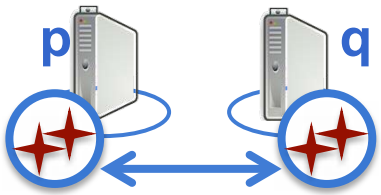


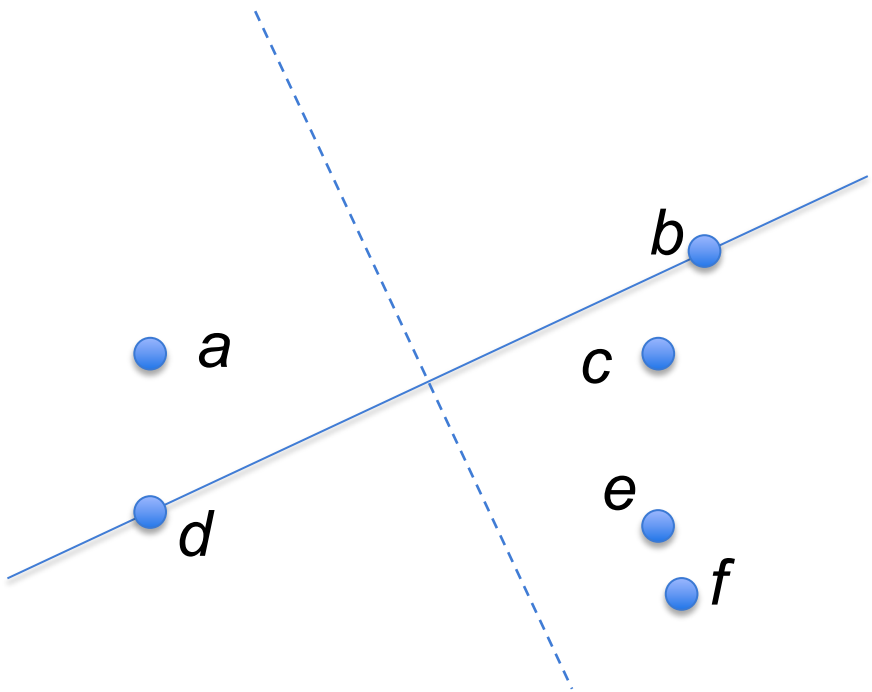
The Migration Process





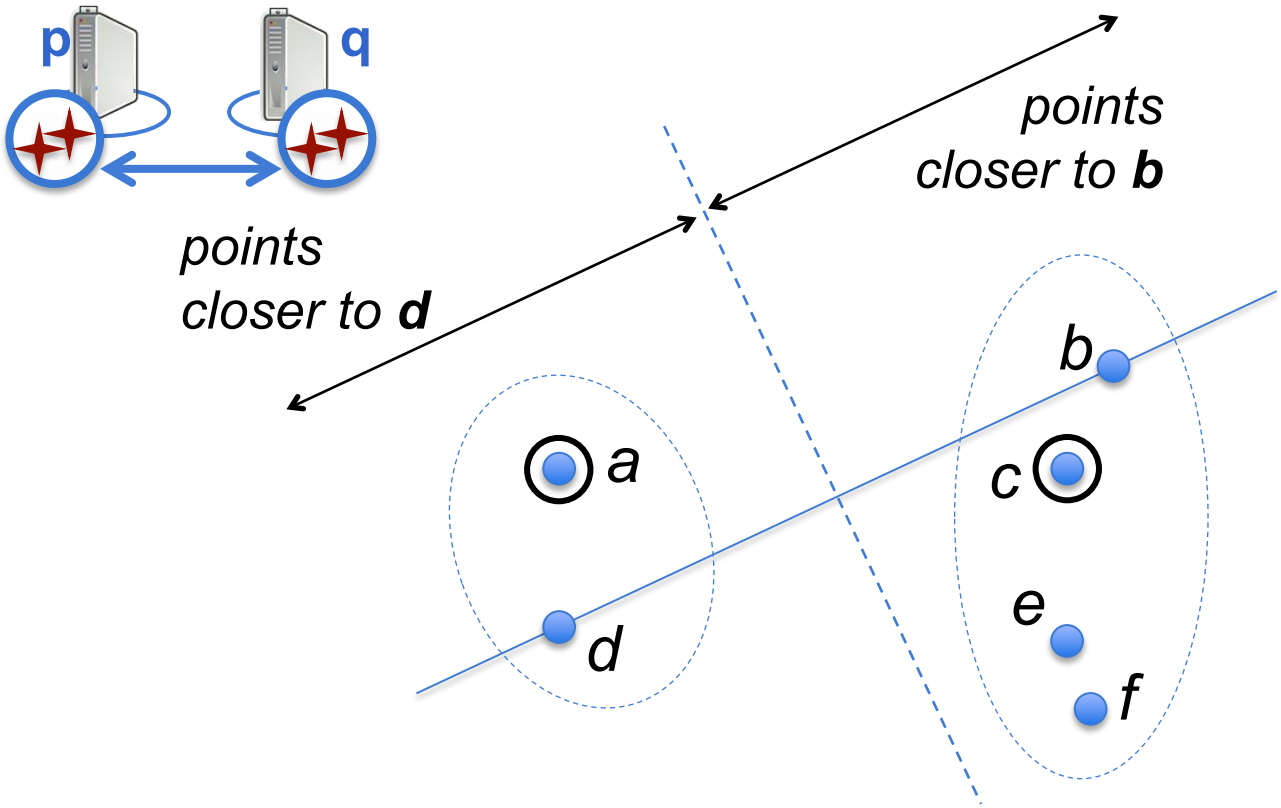
The Migration Process





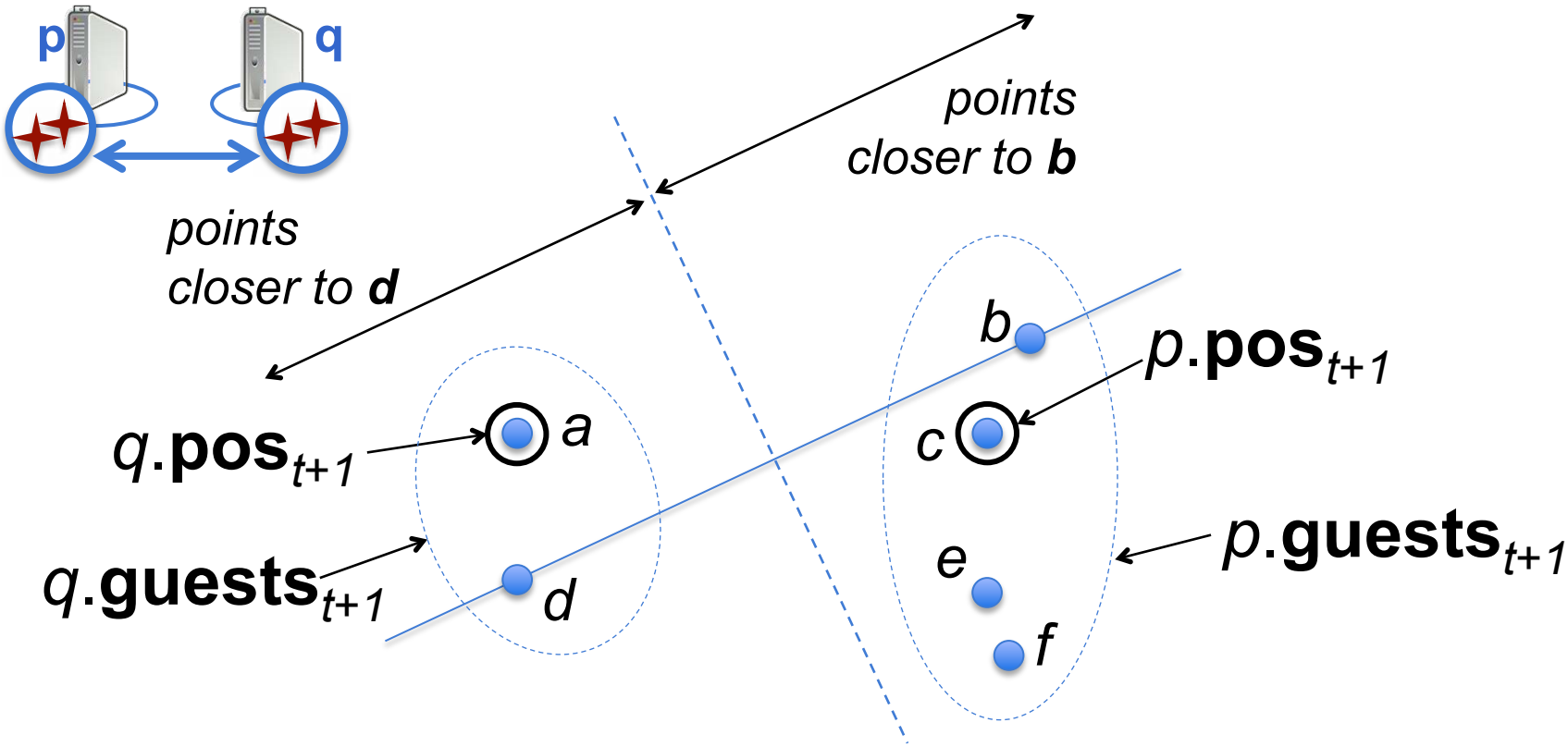
- Bi-clustering of guest points
  - ➔ Heuristics : diameter

The Migration Process



- Bi-clustering of guest points
  - ➔ Heuristics : diameter

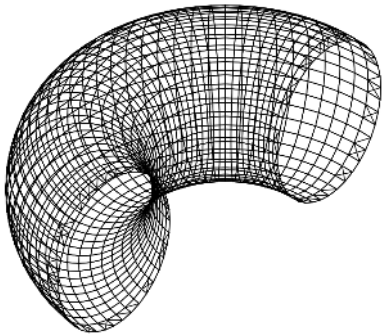
The Migration Process



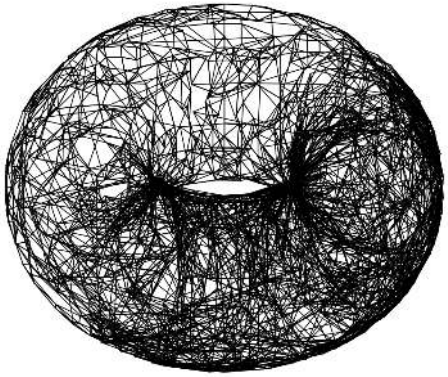
- Bi-clustering of guest points
  - ➔ Heuristics : diameter + minimum move

Evaluation

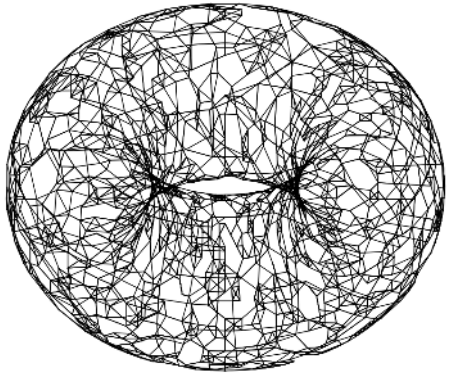
- Shape : 3D torus (replication K = 4)
- Round 20 : 50% correlated node crashes
- Round 100 : 50% node reinjection (repair)



After failure (r=20)



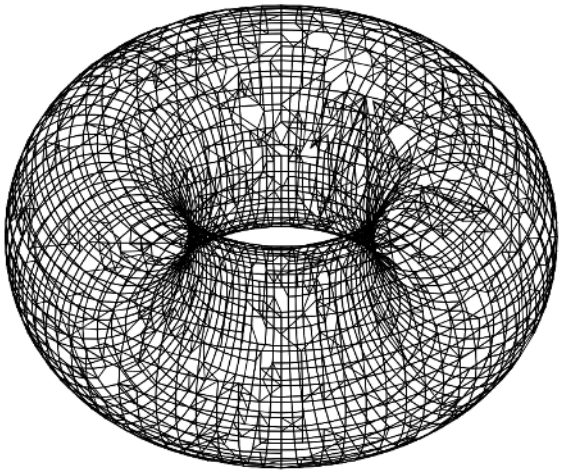
(r=22)



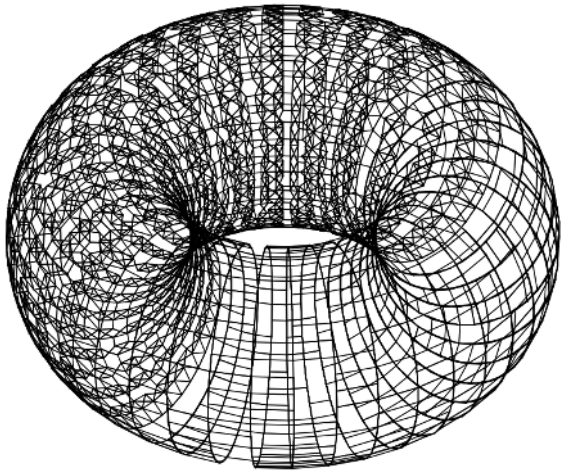
(r=28)

Polystyrene recreates shape with surviving nodes

Evaluation : Node Reinjection



Polystyrene, r = 125

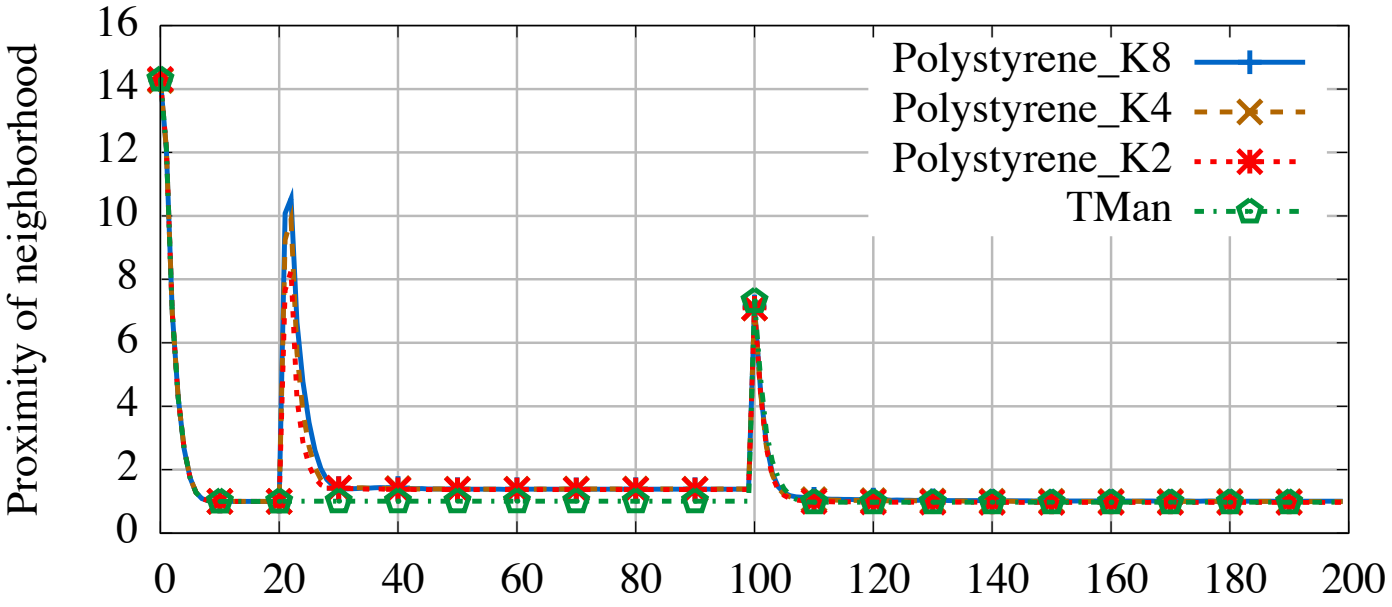


T-Man, r = 125

Polystyrene recreates uniform shape after repair

Eval: Quality of Neigborhoods

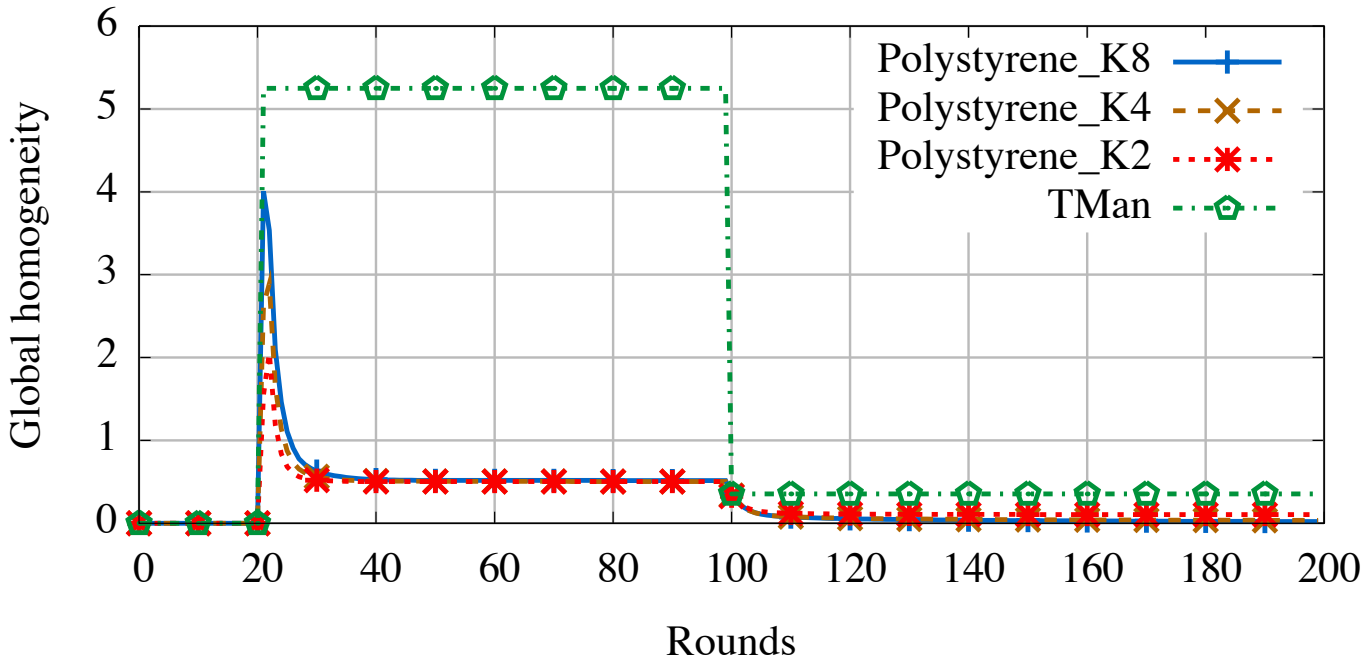
$$proximity = \mathop{\mathbb{E}}_{n_i \in \text{nodes}} \left(\mathop{\mathbb{E}}_{n_j \in n_i.\text{neighbors}} d\big(n_i.\mathbf{pos}, n_j.\mathbf{pos}\big) \right)$$

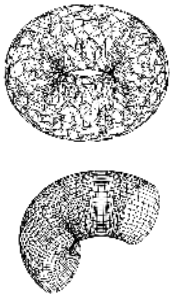


Polystyrene maintains good neighborhoods

Eval: Quality of Shape

$$homogeneity = \mathop{\mathbb{E}}_{\mathbf{x} \in \text{data-points}} \left(\min_{n \in \hat{\text{guests}}^{-1}(\mathbf{x})} \{ d(\mathbf{x}, n.\mathbf{pos}) \} \right)$$

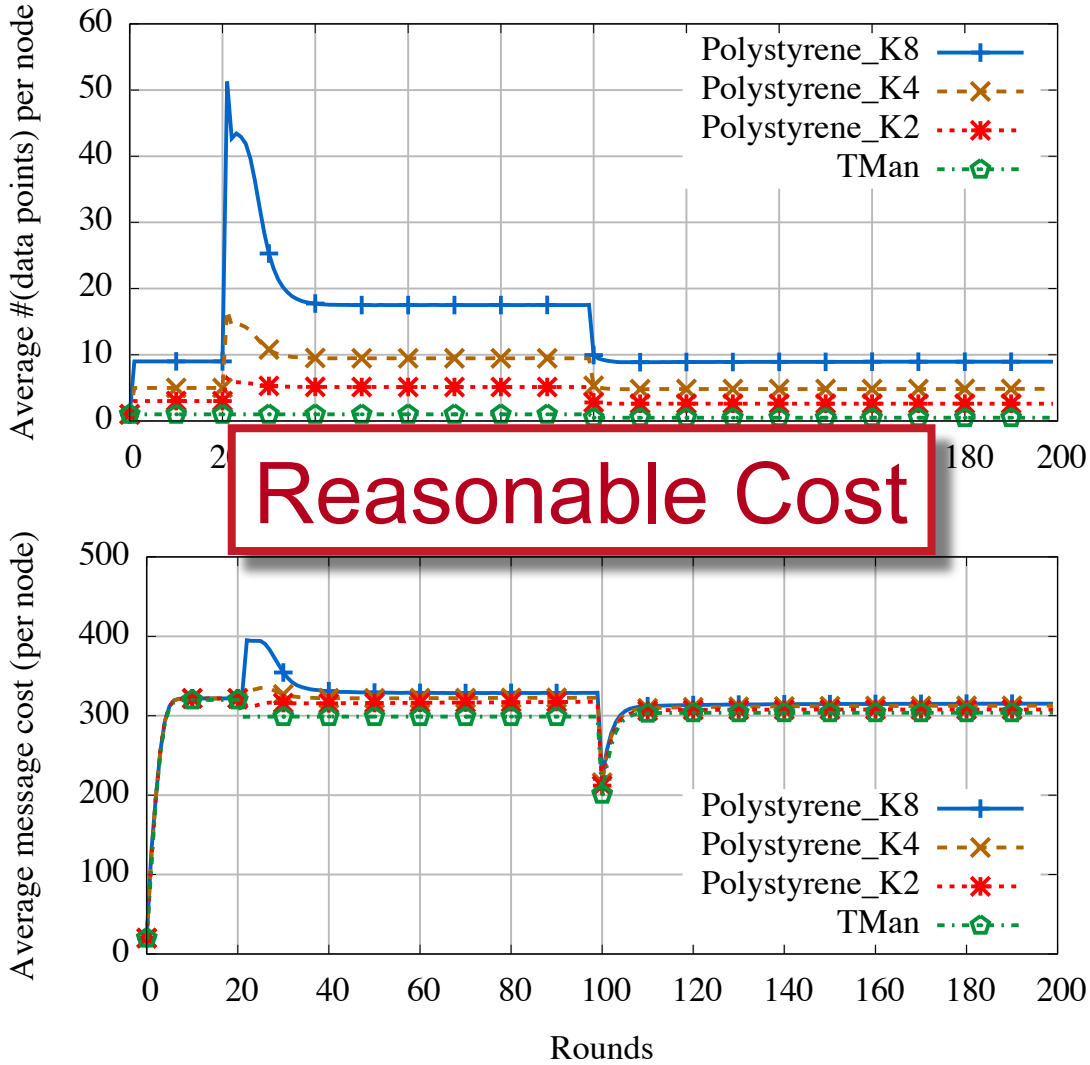




And the torus gets restored!

Eval: Overheads

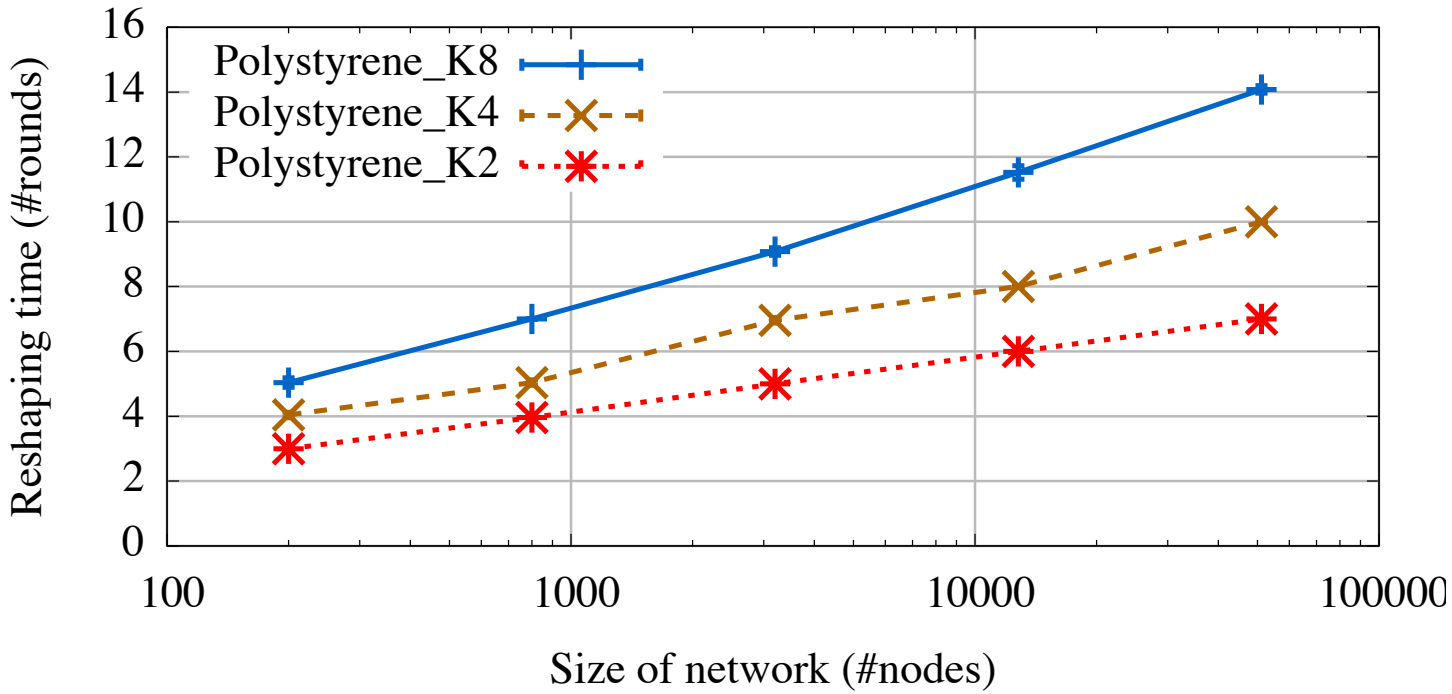
- Message + Memory



Eval: Scalability (1)

- Time (rounds) until homeogeneity less than

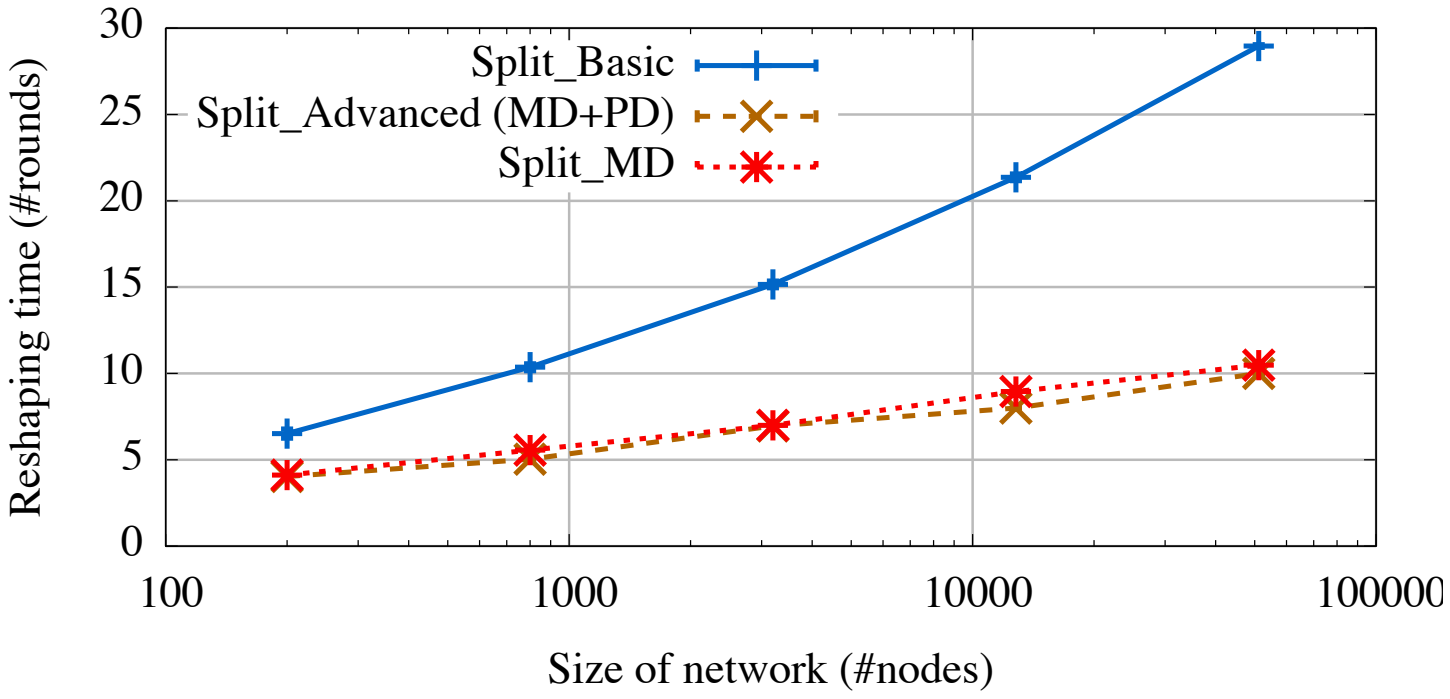
$$\mathcal{H}_{\mathcal{A}}^{|N|} = \frac{1}{2}\sqrt{\frac{\mathcal{A}}{|N|}}$$



Logarithmic convergence!

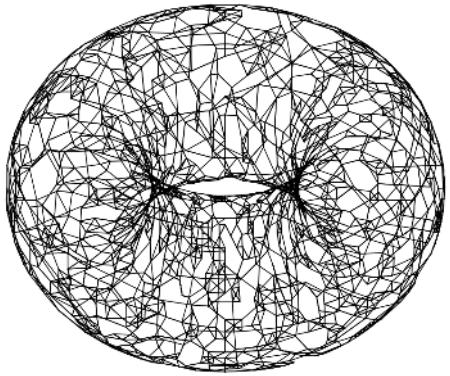
Eval: Scalability (2)

- Strong influence of splitting mechanism



Outlook

- Behaviour in **different shapes**
  - ➔ Ring, grid, logical torus



- **Proof** of logarithmic convergence
- **Concrete application** on top of polystyrene
  - ➔ DHT, recommendation
- Larger picture
  - ➔ **Self-organising data** primitives for **geo-systems**

References

- [JMB09] Mark Jelasity, Alberto Montresor, and Ozalp Babaoglu. 2009. **T-Man: Gossip-based fast overlay topology construction**. *Comput. Netw.* 53, 13 (August 2009), 2321-2339.
- *Simon Bouget, Hoel Kervadec, Anne-Marie Kermarrec, François Taiani*, **Polystyrene: The Decentralized Data Shape that Never Dies**, 2014 IEEE 34th International Conference on Distributed Computing Systems (ICDCS 2014), 30 June - 3 July 2014 in Madrid, Spain (10p)