Freie Universität Berlin

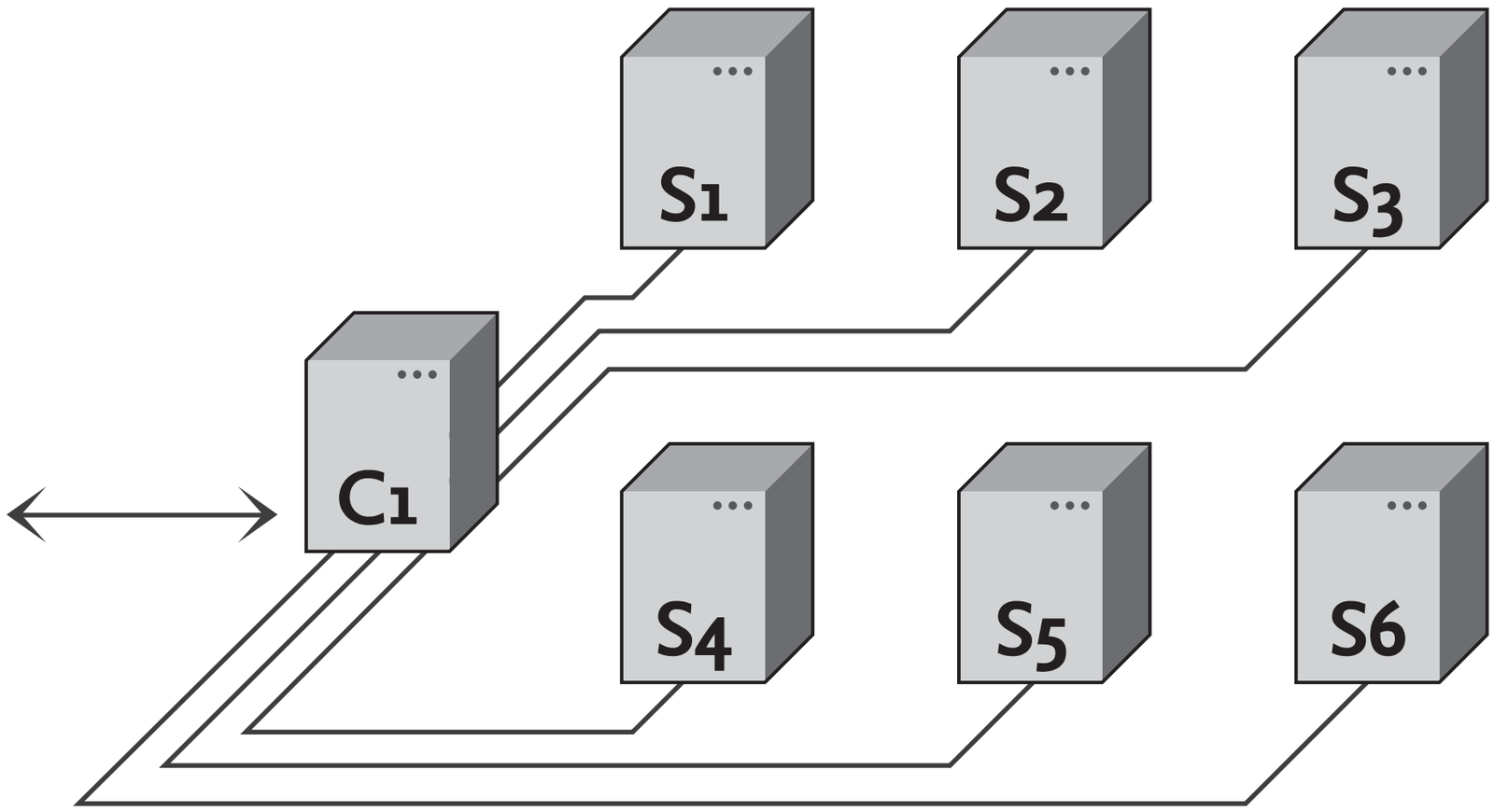
Query Processing in a Self-Organized Storage System

Hannes Mühleisen, supervised by Robert Tolksdorf

Distributed DBs - Goals

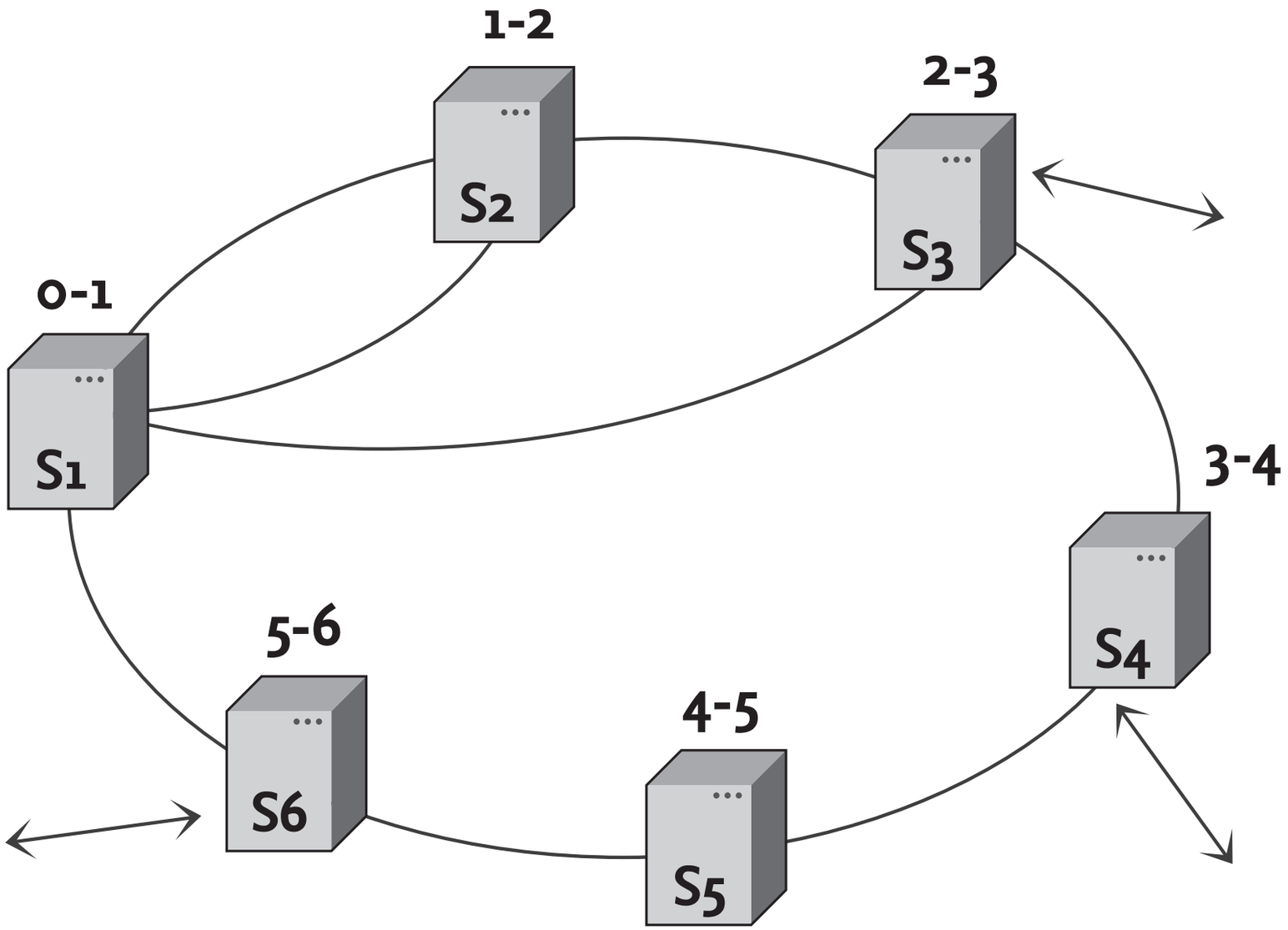
- Scalability
  - Data, Queries, Nodes
- Robustness
  - Node/Network failure
- Adaptiveness
  - “Fair” distribution of load

Clustered / Federated



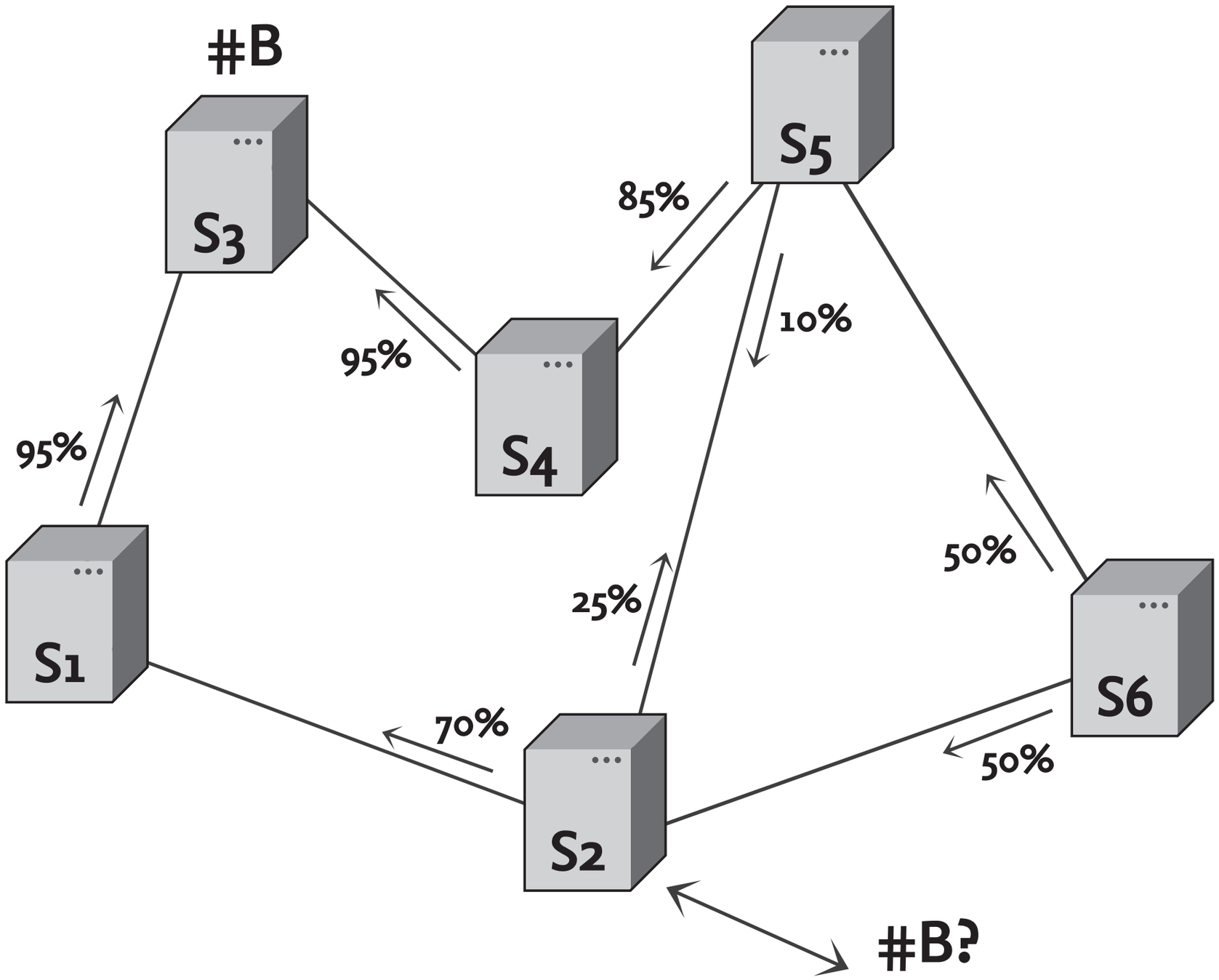
[Bernstein81, Epstein78]

Global Laws

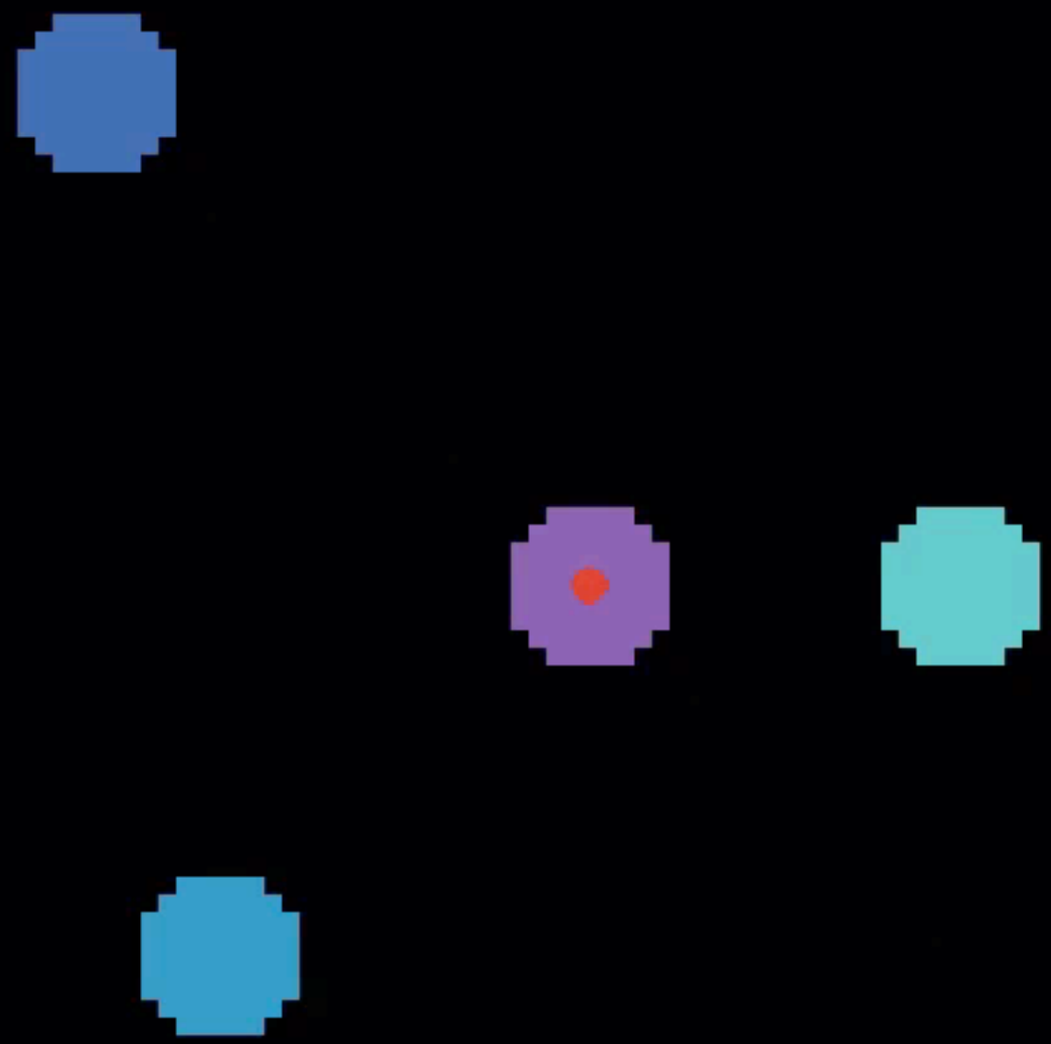


[Harreno2,Karnstedto4,Röscho5]

Probabilistic Request Routing



[Lindgreno3]



[Wilensky97, NetLogo Ants model]

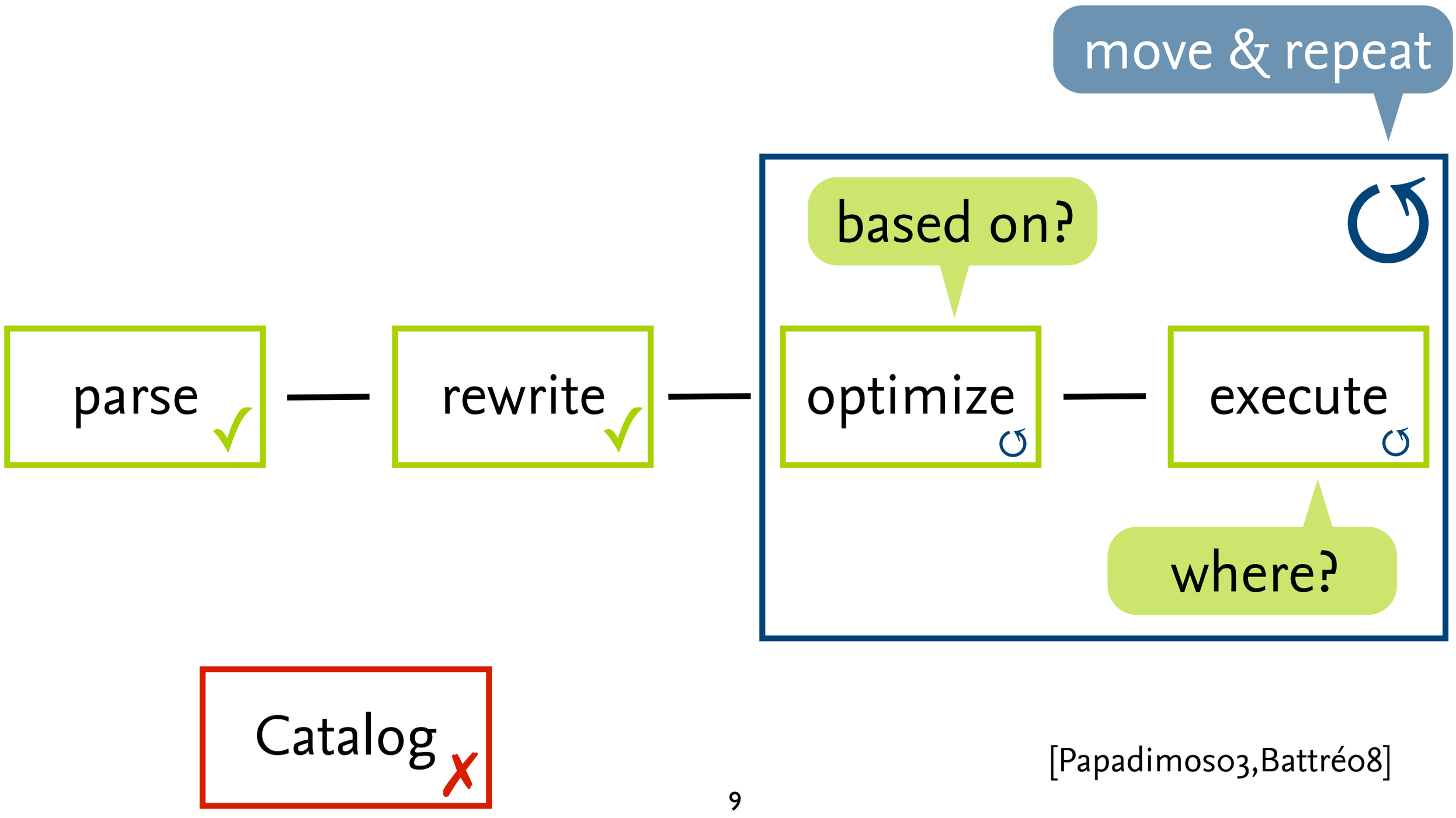
Distribution Paradigms

	Scalability	Adaptability	Robustness	Completeness	Complex Queries
Stand-Alone	low	**high**	low	**high**	✓
Federated	**high**	**high**	fair	**high**	✓
Global-Law	**high**	fair	**high**	**high**	✓
Probabilistic e.g. Swarms	**high**	**high**	**high**	fair	?

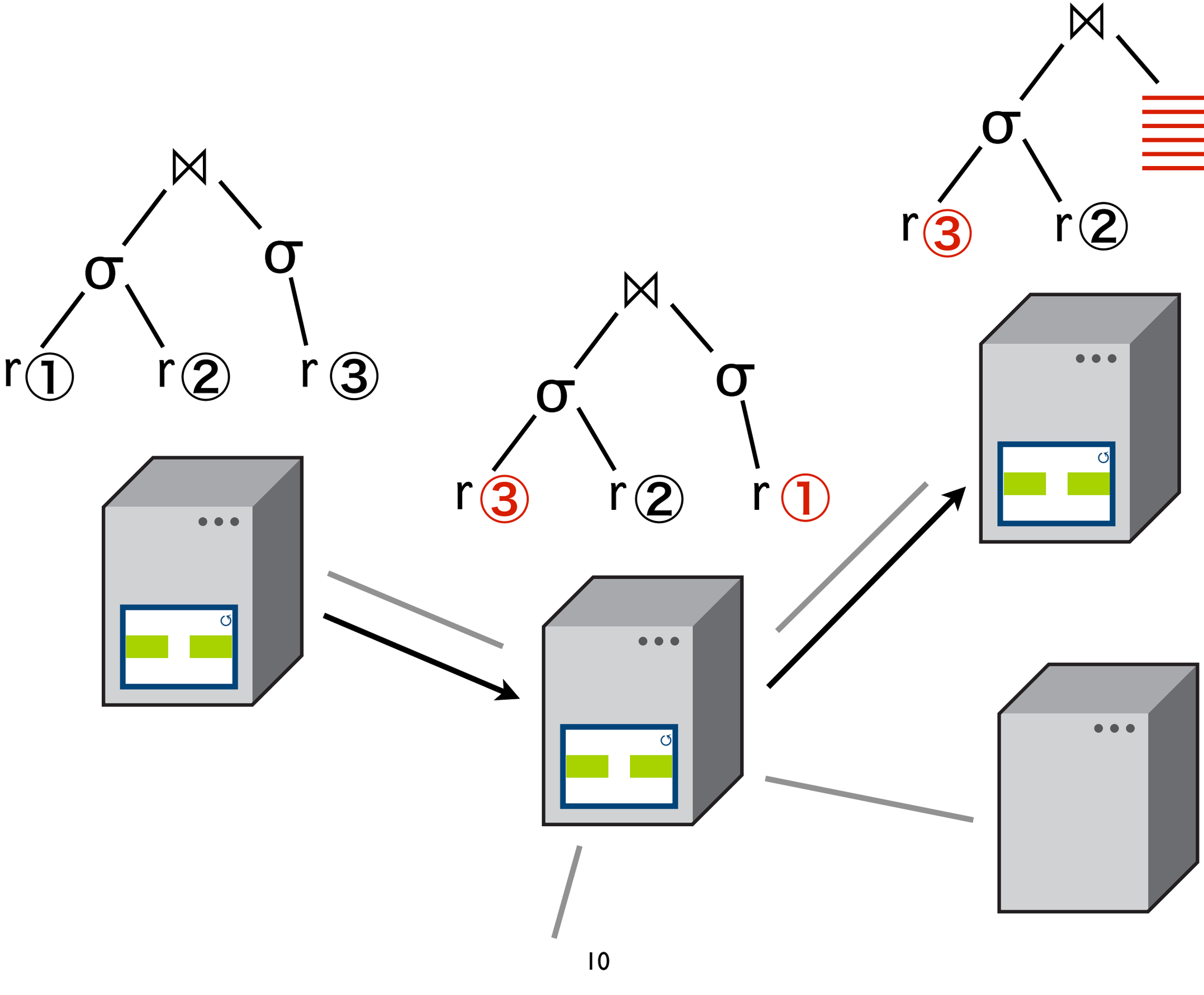
Research Question

Can complex queries be evaluated efficiently in a swarm-based distributed storage system?

Mutable Moving Query Plans

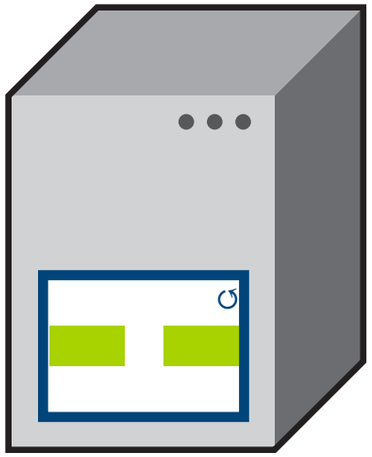


[Papadimoso3,Battréo8]


⋈
σ
σ
r ①
r ②
r ③
⋈
σ
σ
r ③
r ②
r ①
⋈
σ
r ③
r ②

① r(#)

② r(*)

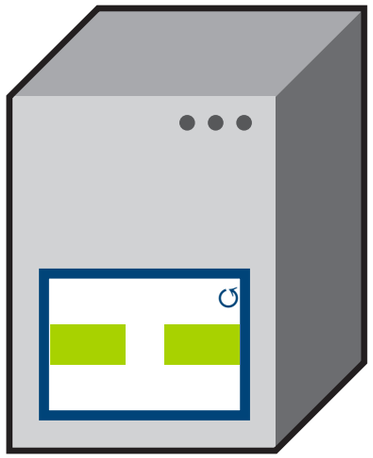


p(#)=53%

p(*)= 3%

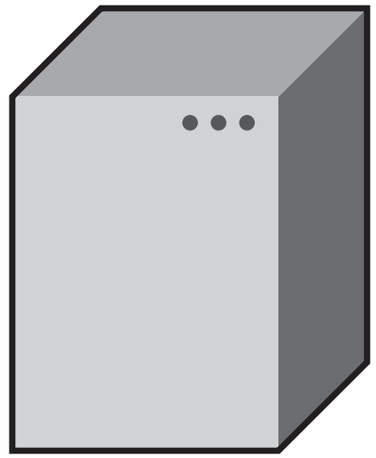
① r(*)

② r(#)



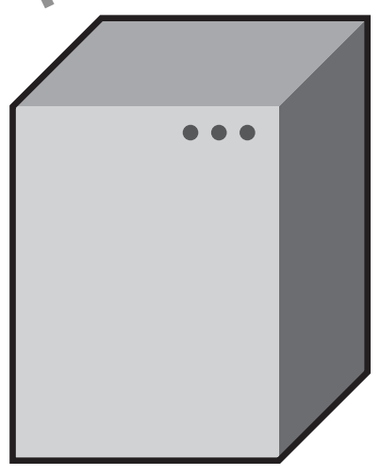
p(#)= 2%

p(*)=78%

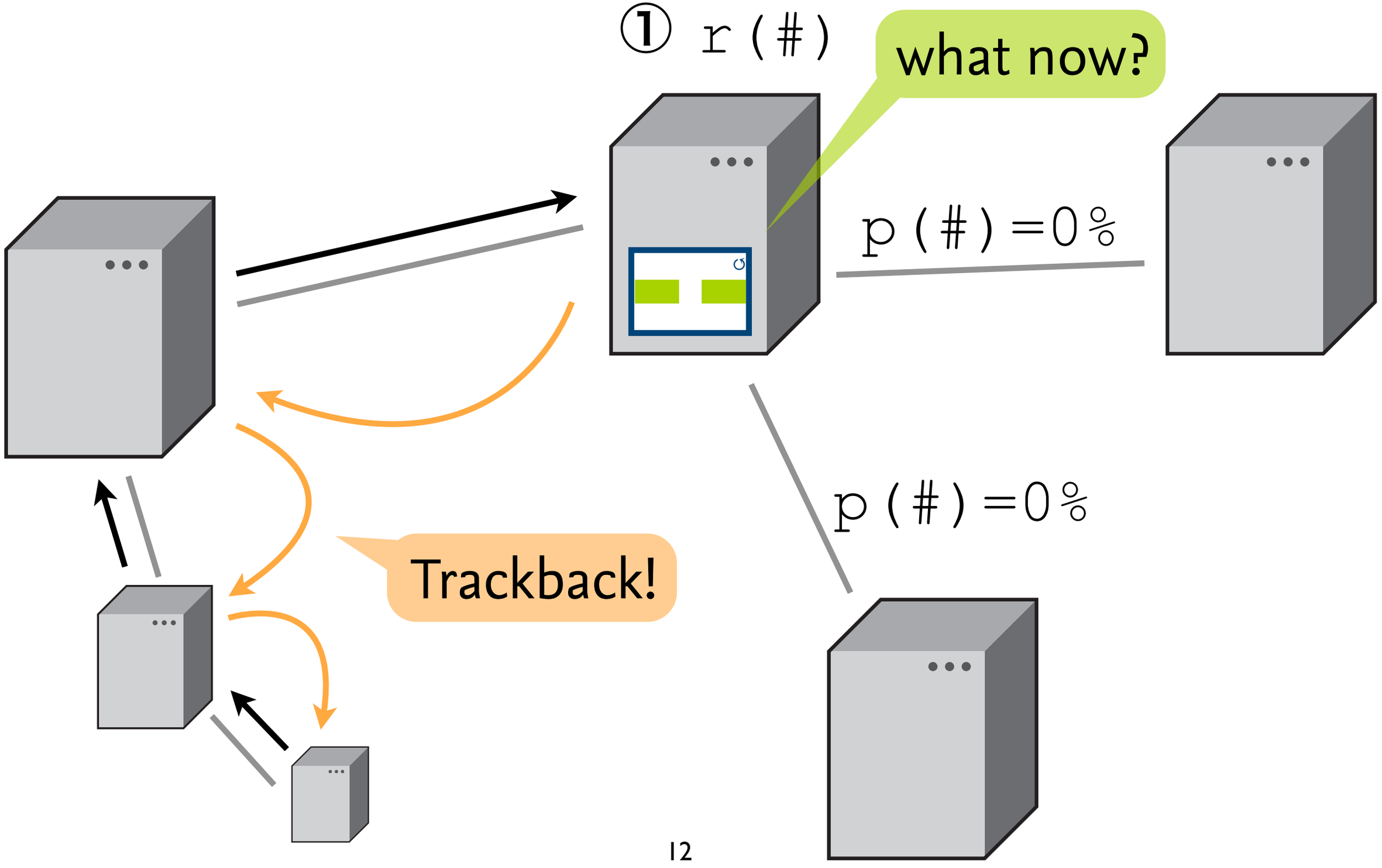


p(#)= 2%

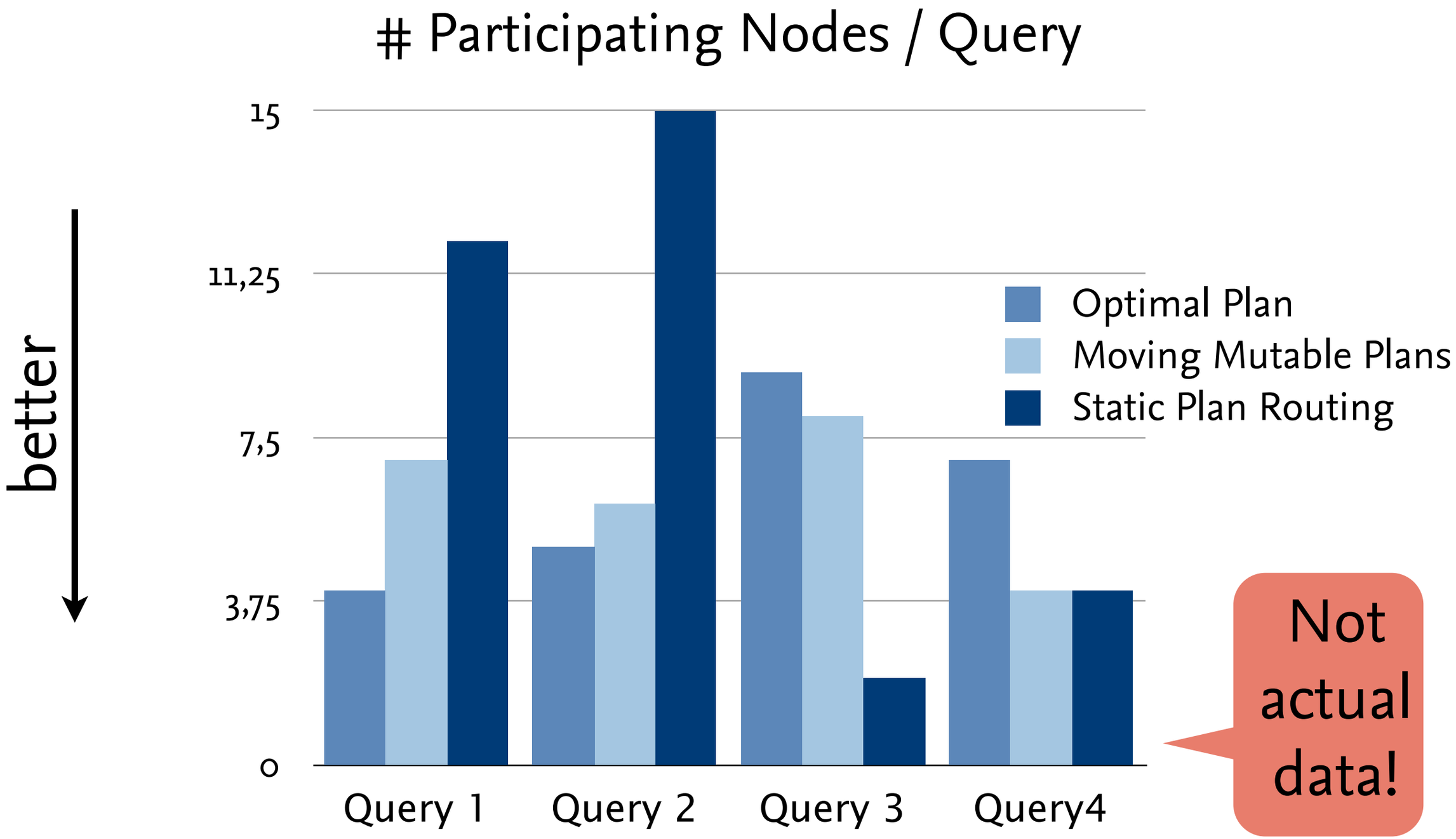
p(*)=10%



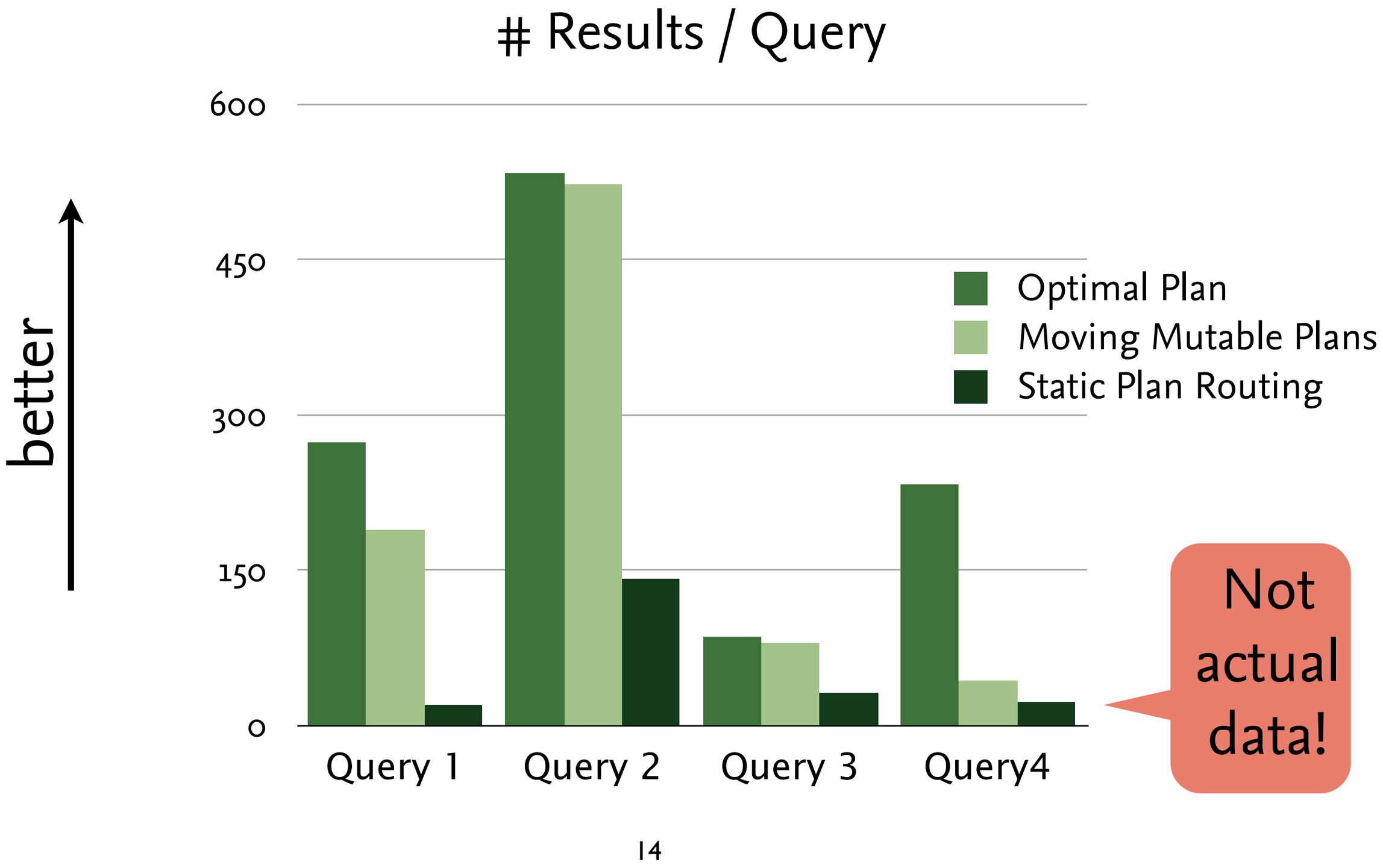
Handling Routing #Failures



Evaluation Methodology



Evaluation Methodology



Thank You!

Questions?

Web Page: `http://hannes.muehleisen.org`