An overlay network for resource discovery in Grids

Manfred Hauswirth (joint work with Roman Schmidt)

Ecole Polytechnique Fédérale de Lausanne (EPFL)
School of Computer and Communication Sciences

Funded by the European integrated project
DIP – "Data, Information, and Process Integration with Semantic Web Services",

Information Society Technologies

dip
Data, Information and Process Integration with Semantic Web Sevices

Contract no. 507483

Outline of the presentation

- Approaches to resource discovery in Grids
- Using P2P systems for resource discovery in Grids
- Detour: The basics of scalable data access structures and overlay networks
- Our proposal: Using the P-Grid overlay network for resource discovery
- Experimental evaluation
- Conclusions

Approaches to resource discovery in Grids - 1

- Centralized: Condor
  - targets primarily optimal CPU utilization
  - centralized matchmaker to match resource requests with offers
  - efficient for small grids in LANs, but does not scale to larger sizes
- Hierarchical: Monitoring and Discovery Service (MDS)
  - used in Globus
  - based around WSRF (Web Services Resource Framework) standards
  - provides a registry similar to UDDI
  - query and subscription (trigger) interfaces
  - support of global-scale grids: the hierarchical organization and query routing has hot-spots and single-points-of-failure

Approaches to resource discovery in Grids - 2

- Decentralized / P2P: Iamnitchi et al.
  - unstructured P2P network similar to Gnutella combined with Freenet-style query forwarding
  - less traffic than pure Gnutella but no lookup guarantees
- Decentralized / P2P: Gupta et al.
  - based on a range-query-enhanced version of CAN
  - ranges are hashed and indexed ⇒ simple key search operations are not supported or are highly inefficient (both are needed) ⇒ separate indexes
  - CAN does not support efficient updates (update ⇒ new responsible peer)
  - search efficiency is only guaranteed for uniform partitioning of key space

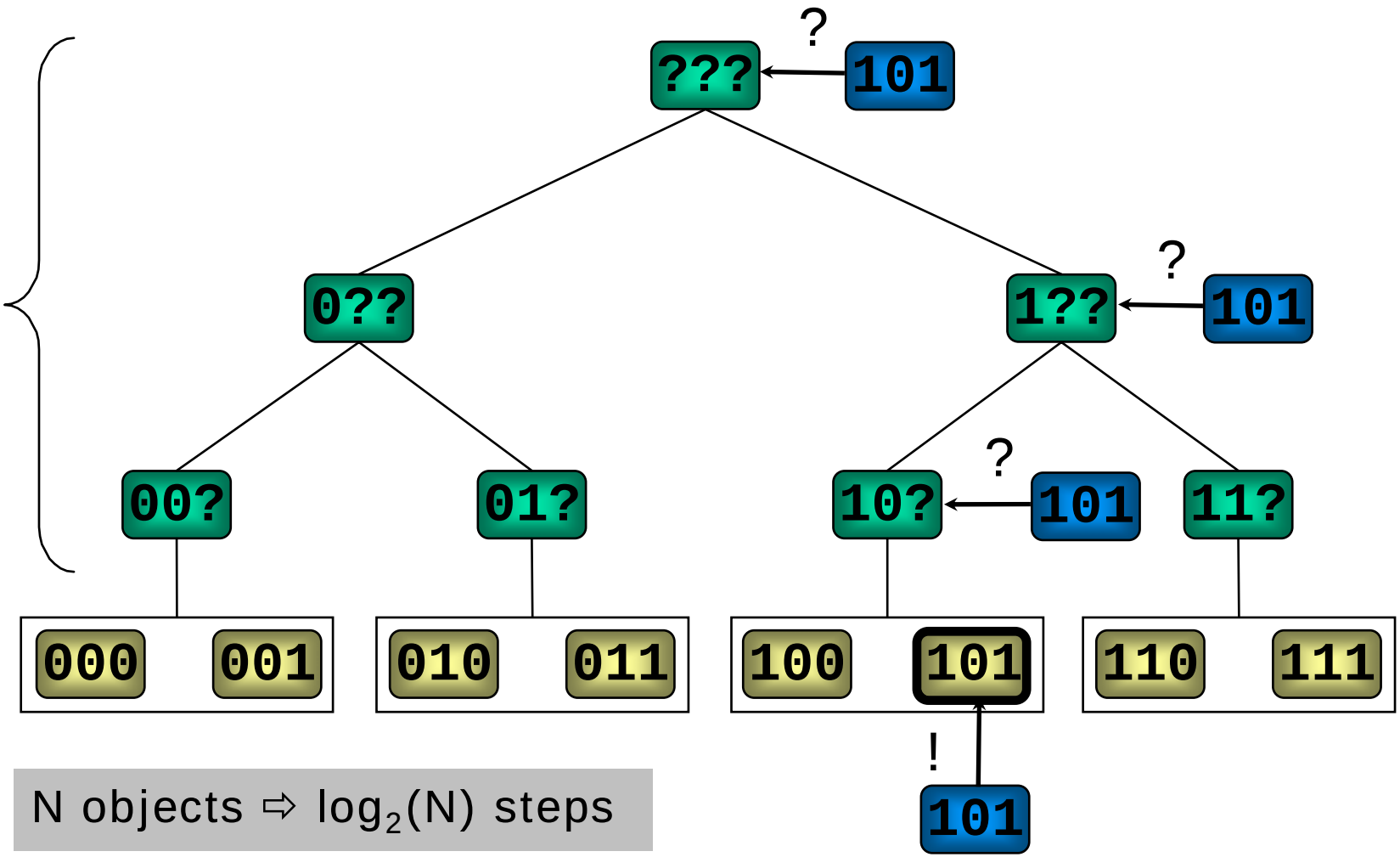
Are overlay networks usable for resource discovery in Grids?

- Grid community
  - often uses inefficient versions of existing P2P approaches
- P2P community
  - does not address the specific needs of Grid computing
    - exact search is fast, other search predicates do not exist or are inefficient
    - frequent update of resource state required but updates are either not supported or are inefficient
  - some assumptions are inadequate
    - Grids normally do not have very large numbers of nodes and data
    - node population is rather stable
- Advantages of P2P approach
  - no dedicated nodes required
  - no “single point of failure” (node, network)
  - implicit load distribution and balancing
  - no dedicated infrastructure needed - “the system is the directory”

Resource discovery based on overlay networks seems an interesting approach for global-scale / large-scale Grids, otherwise other approaches may be more applicable

Detour: Data access structures

Search tree (prefix tree)

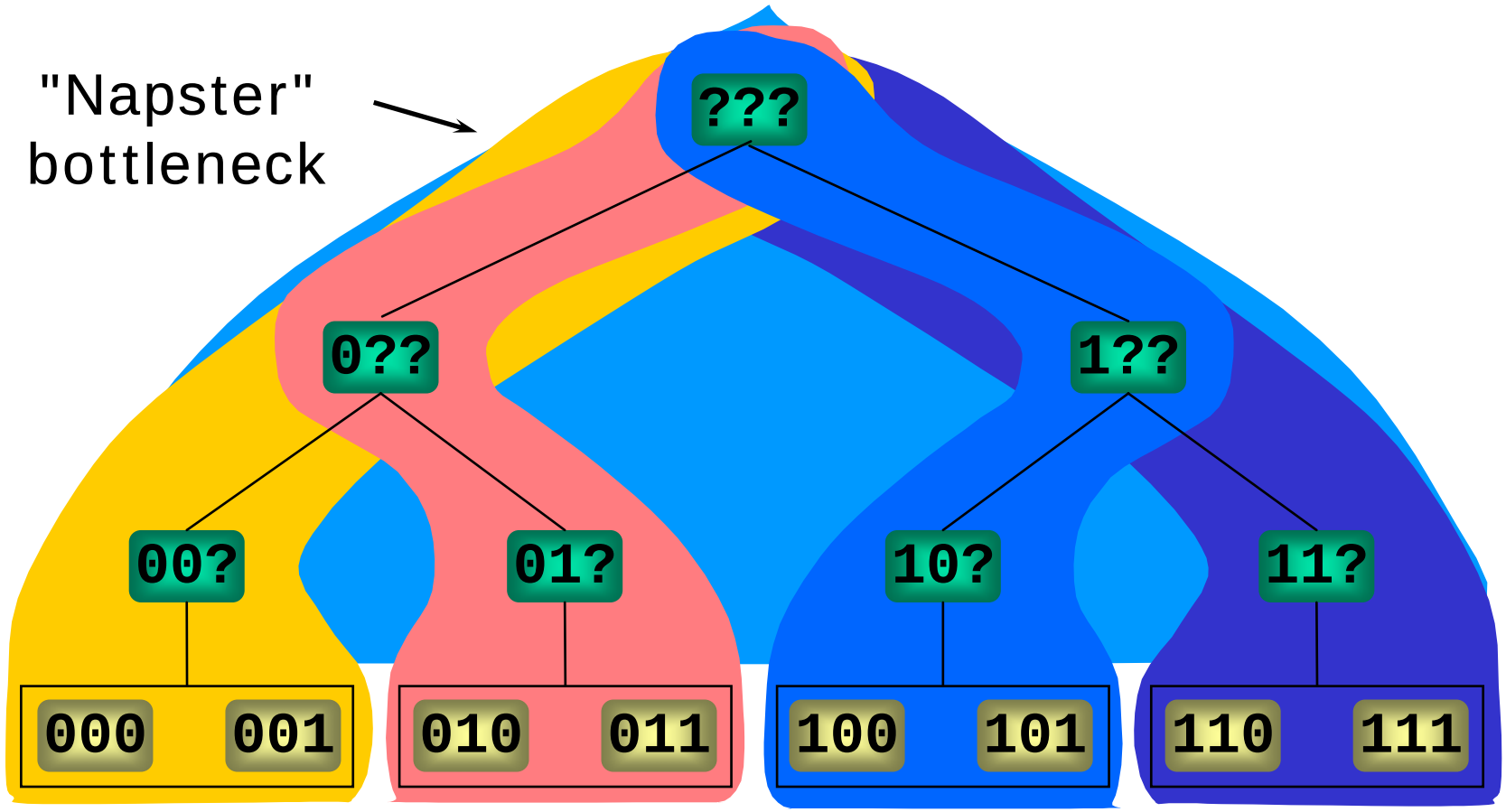


N objects ⇨ $\log_2(N)$ steps

Detour: Scalable data access structures - 1

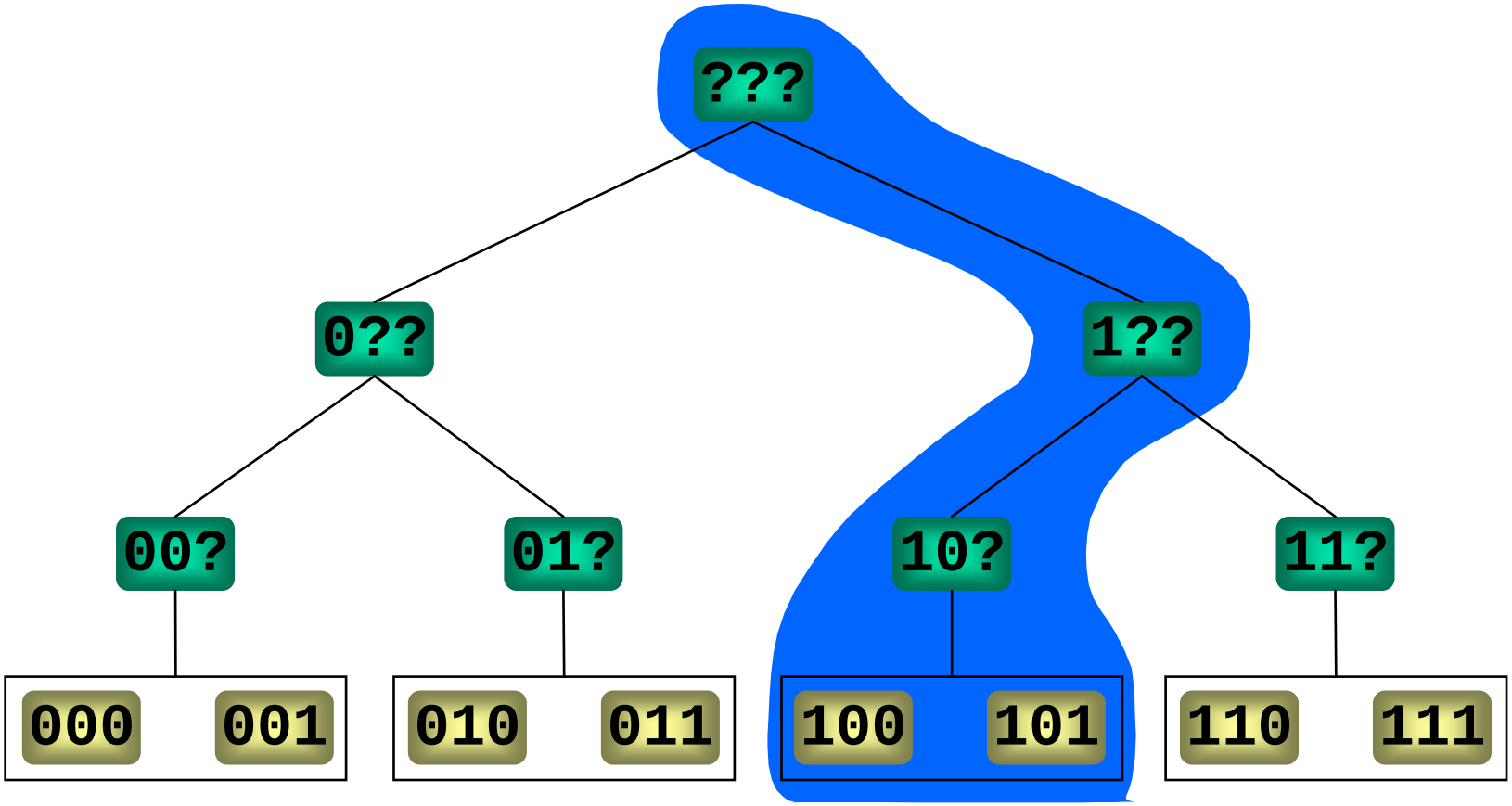
- Assume number of data objects >> storage of one node
  - Distributed storage
- Given a data access structure
  - Size of data access structure = number of data objects
  - Size of data access structure >> storage of one node
- Problem: where to store?

Detour: Scalable data access structures - 2

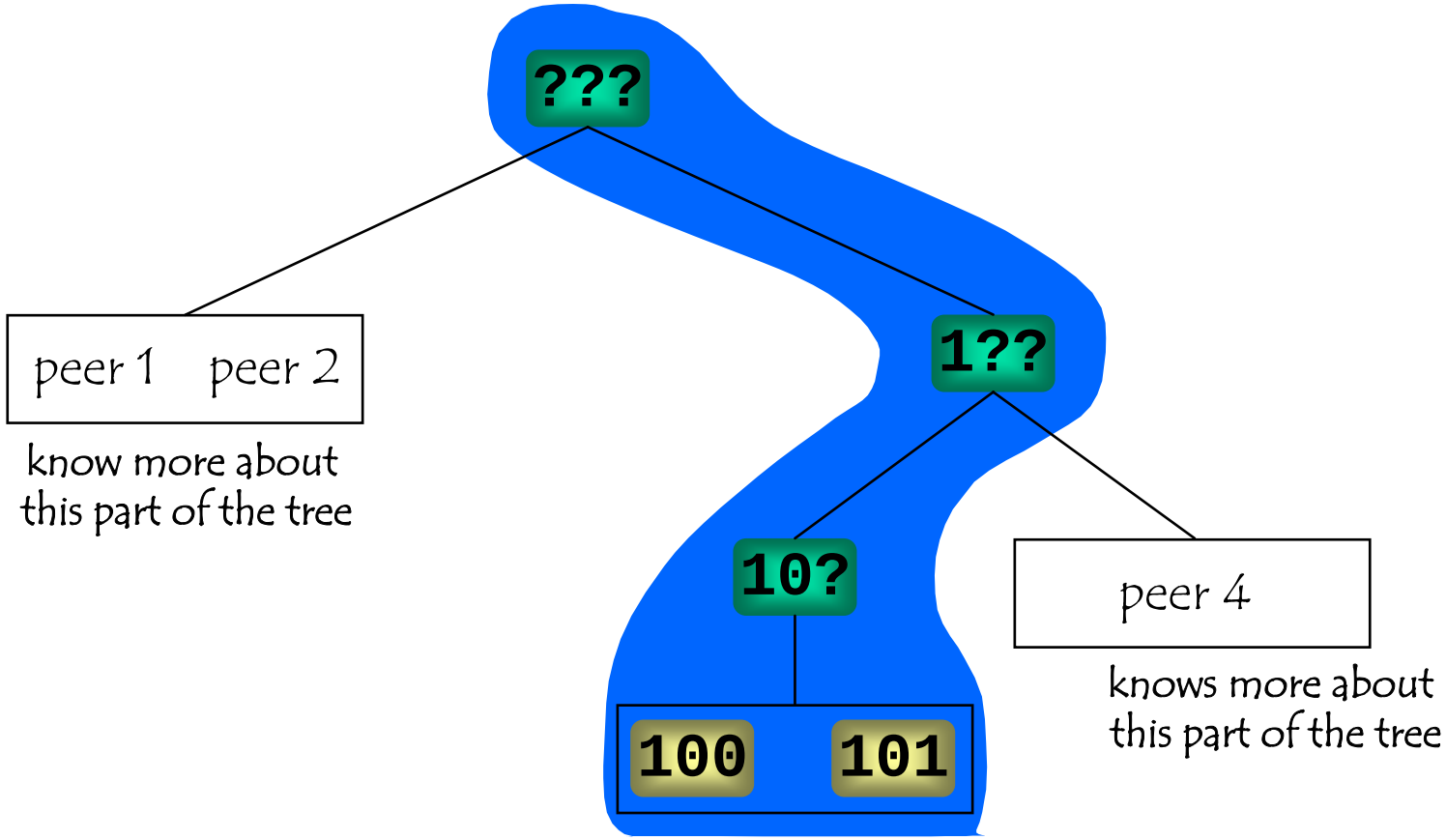


Detour: Scalable data access structures - 3

Associate each peer with a complete path

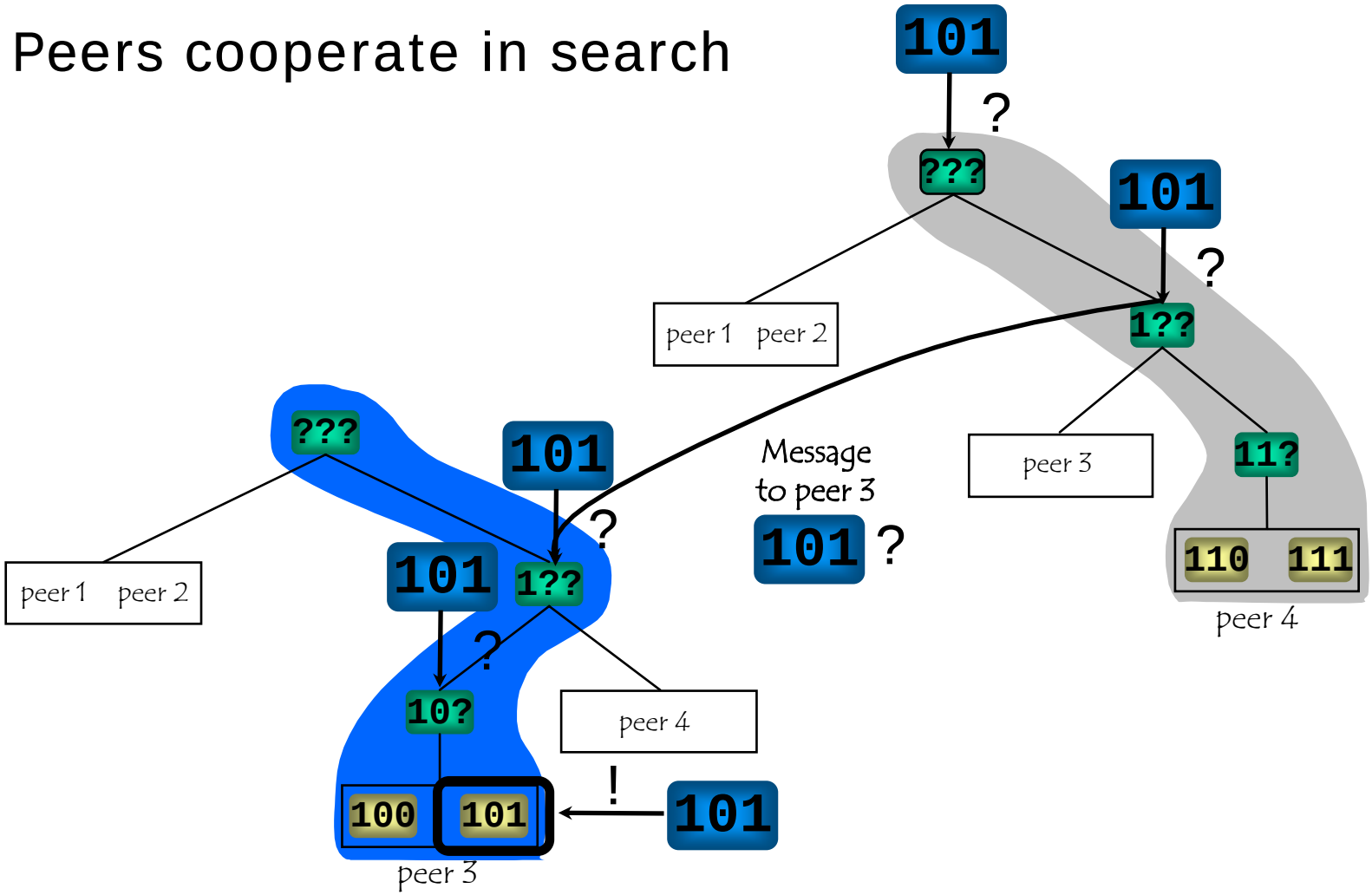


Detour: Scalable data access structures - 4

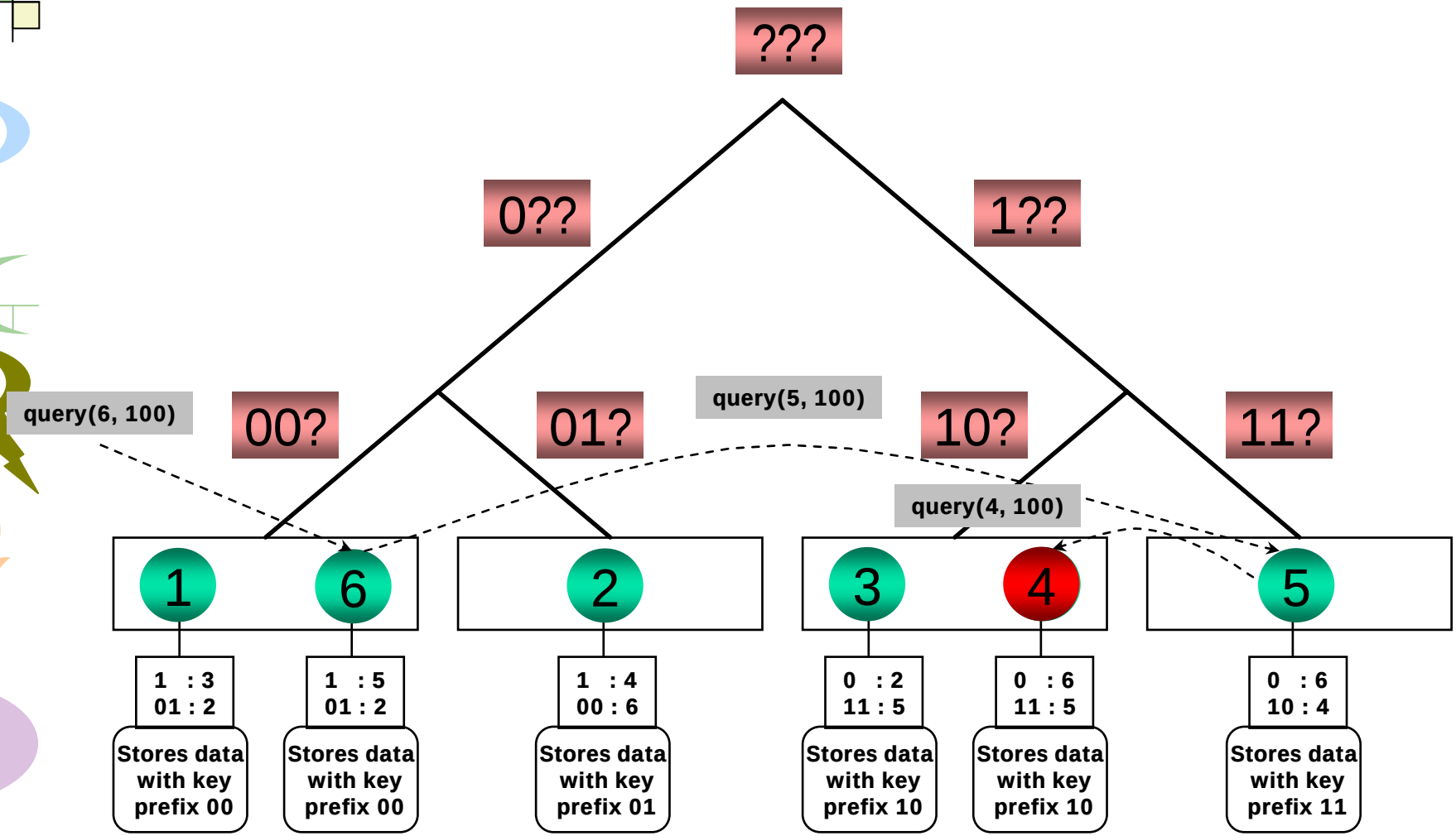


Detour: The result is P-Grid

- Peers cooperate in search

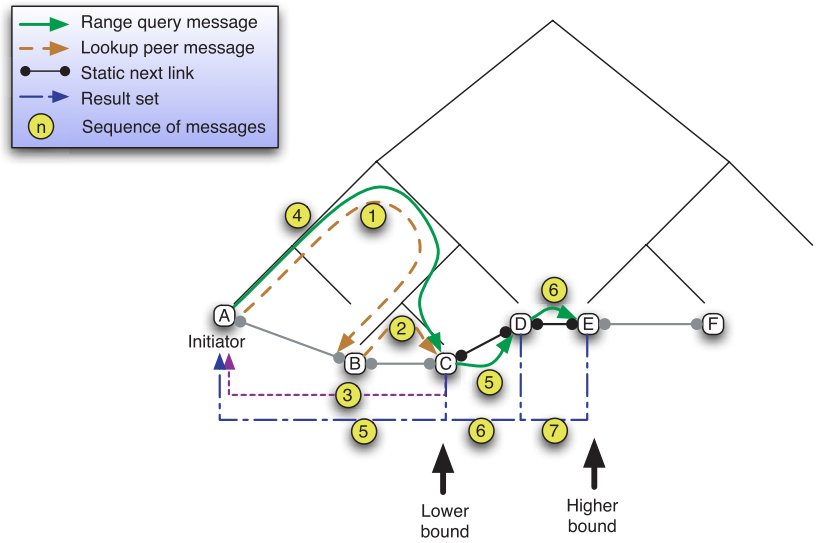


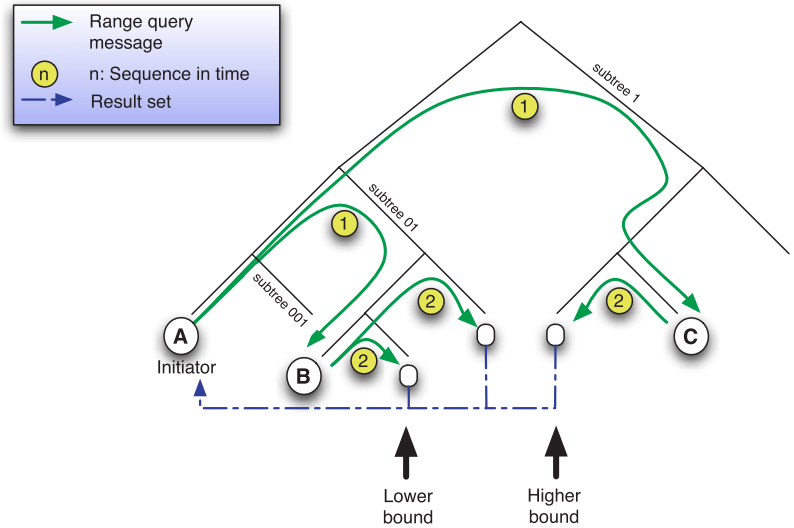
Detour: P-Grid queries



P-Grid overview - 1

- Efficient search in *O(log n)* steps (*n* nodes) even for skewed distributions
- Exact search, substring search, and efficient range queries [IEEE P2P 2005] (simple XPath is already supported as well [ODBASE 2005])
  - 2 range-query algorithms: min-max, shower





- Efficient, epidemic update algorithm for highly unreliable environments [ICDCS 2003]

P-Grid overview - 2

- Load-balancing of memory and replication load (availability)
- Prefix-preserving hash function for key generation
  - $s_1 < s_2 \Rightarrow h(s_1) < h(s_2)$
  - $\Rightarrow$ clustering of similar information
- P-Grid's trie only exists virtually, in fact the system is "flat" and all nodes are equal
- Self-organized construction of the index
- Individual P-Grids can be split and merged
- Available from http://www.p-grid.org/ under a modified GPL

Our approach: Using P-Grid for resource discovery

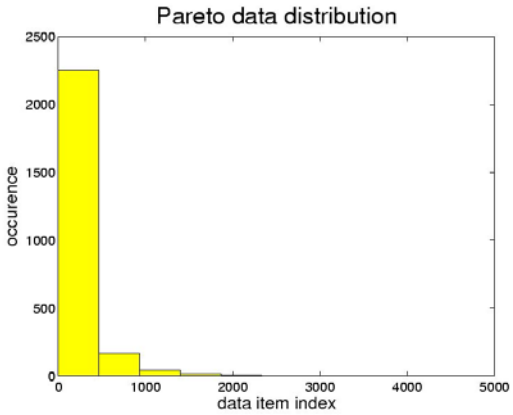
- Instead for resources and their states, job requirements are advertised, for example,

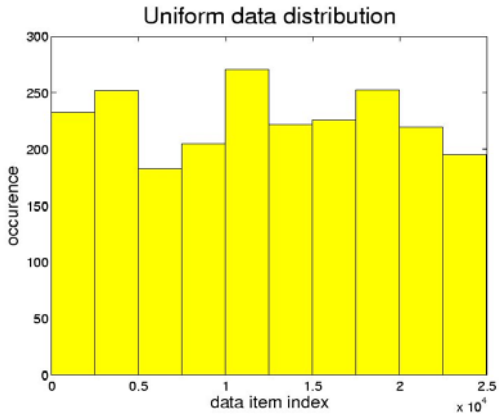
```
CPU_cycles=3500,disk=50MB,mem=1024MB,
advertiser=http://need.cpu.com/job42
```

- Providers actively look for jobs (exact search or range queries) and accept the ones they want to
  - ⇒ less updates required
  - ⇒ resource provider is in control
- Specific problems to address
  - key distributions may be highly skewed, for example, if most job advertisement are at the maximum of possible values and then sharply decrease.
  - but also uniform distributions have to be supported
  - robustness, scalability and efficiency

Experimental evaluation on PlanetLab

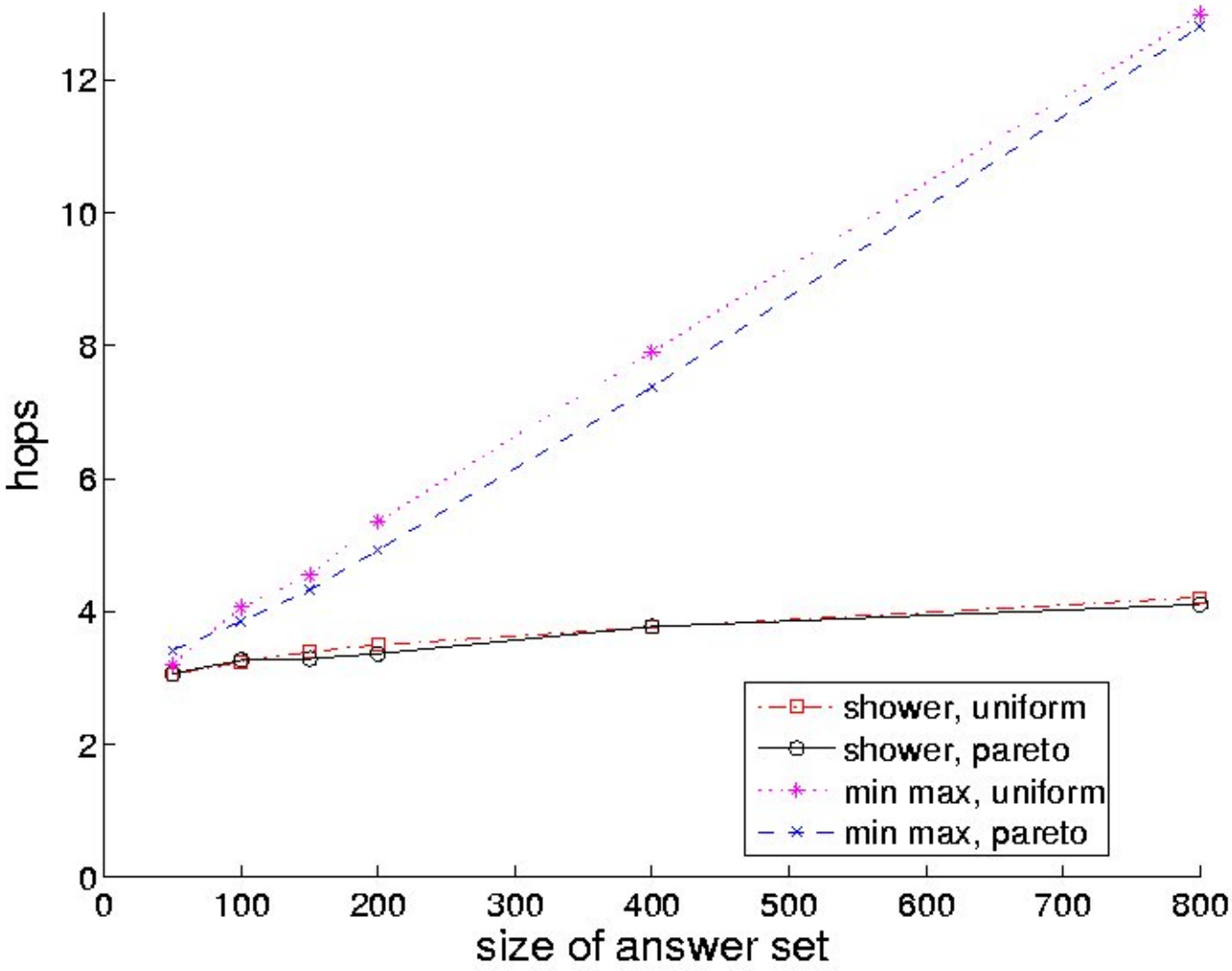
- PlanetLab: World-wide testbed for distributed applications
  - approx. 450 nodes
  - wide range of network connectivity (T1, DSL, etc.)
  - large number of experiments in parallel
- 250 peers, each running on a dedicated PlanetLab node
- 2500 unique data keys (Pareto and uniformly distributed), each peer selects 10, average replication factor was set to 5 ⇒ 18750 keys in the system, each peer is responsible for 50-100 keys



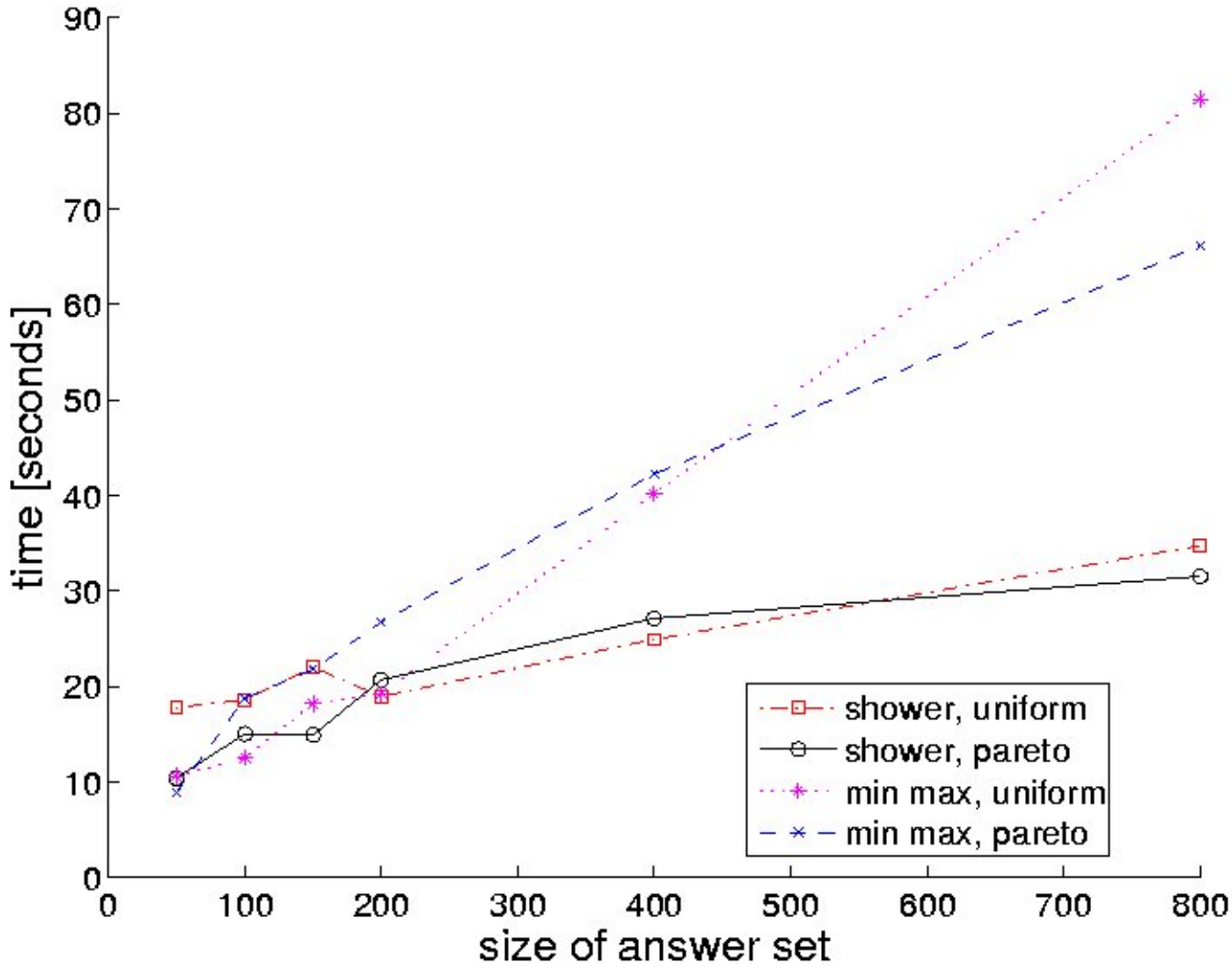


- Each node performs a query with a random lower bound for each distribution, with 2 different algorithms, and for each of the answer set sizes (50, 100, 150, 200, 400, and 800), i.e., a total of 250 * 2 * 2 * 6 = 6000 queries

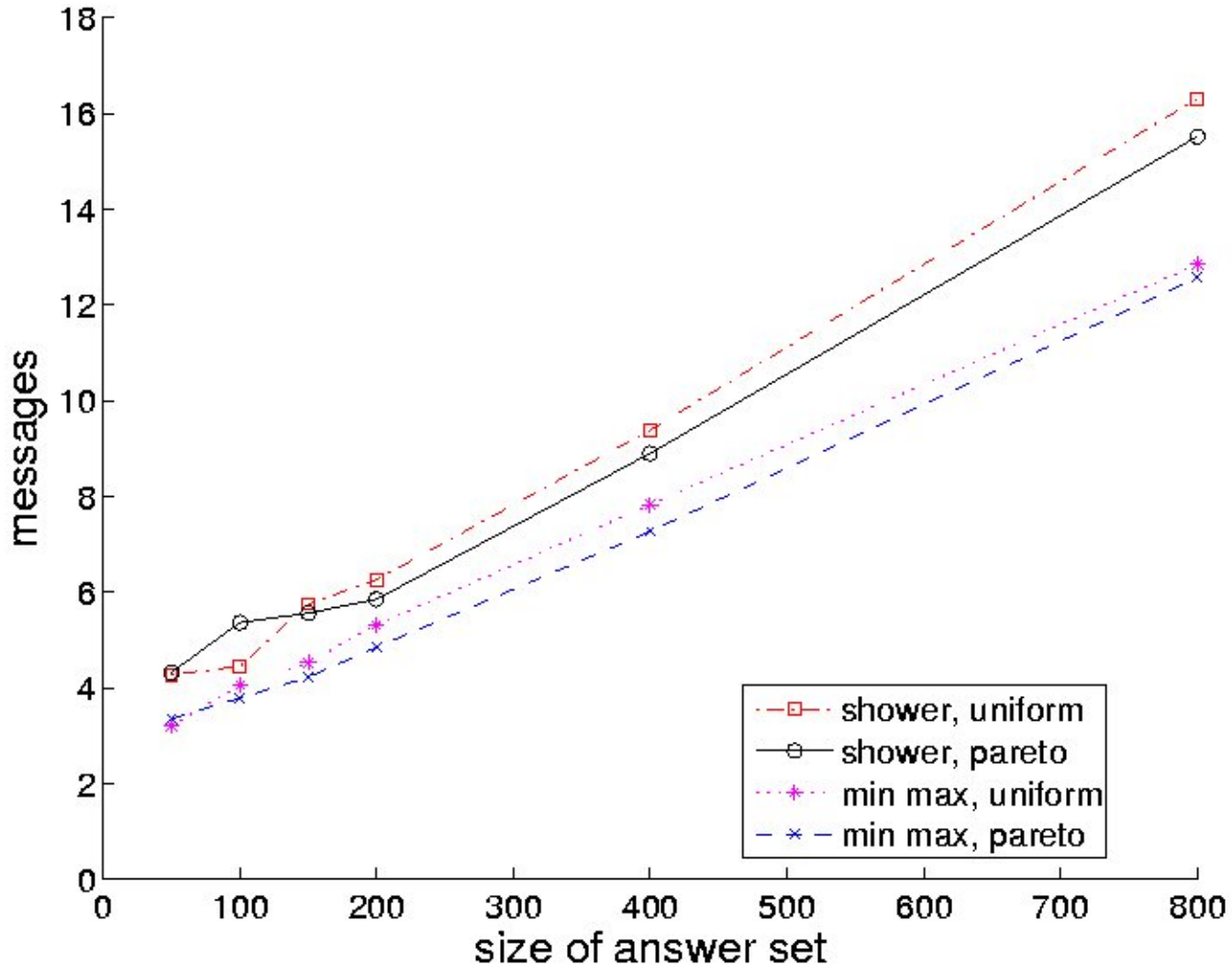
Experimental results: Message latency (hops)



Experimental results: Latency (time)

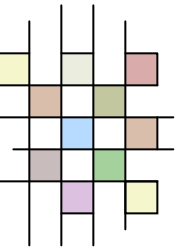


Experimental results: Message costs



Conclusions

- Overlay networks for resource discovery could be applicable in very large scale Grids
- Base overlay technologies exist but a more in-depth investigation of applicability is necessary (latency, updates, etc.)
- Job advertisements instead of resource advertisements may be also interesting for other Grid discovery approaches to strengthen the autonomy and control of the resource provider
- P-Grid overlay was tested under worst-case conditions as infrastructure for discovery with promising results
  - We can expect much better results in Grid environments which are more stable
- More cooperation between Grid and P2P communities may be necessary and fruitful



Thank you!

Questions?