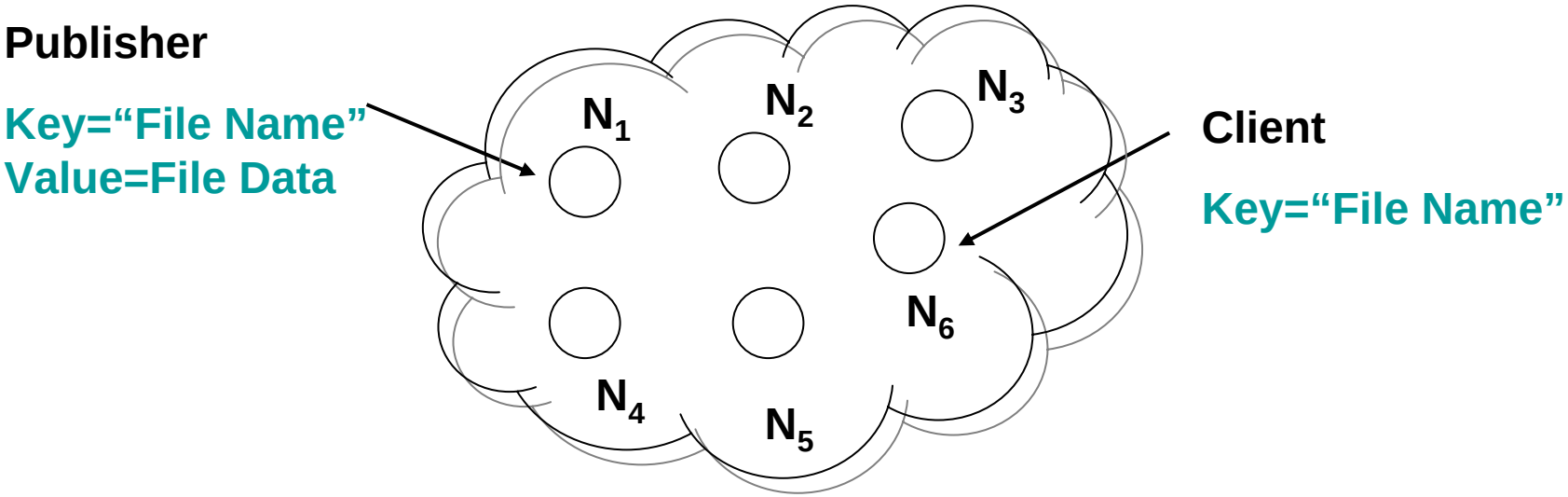
LinkNet: A New Approach for Searching in a Large Peer-to-Peer System

Kunlong Zhang, Shan Wang

School of Information, Renmin University of China
zhangkl@ruc.edu.cn, swang@ruc.edu.cn

2005-03-30

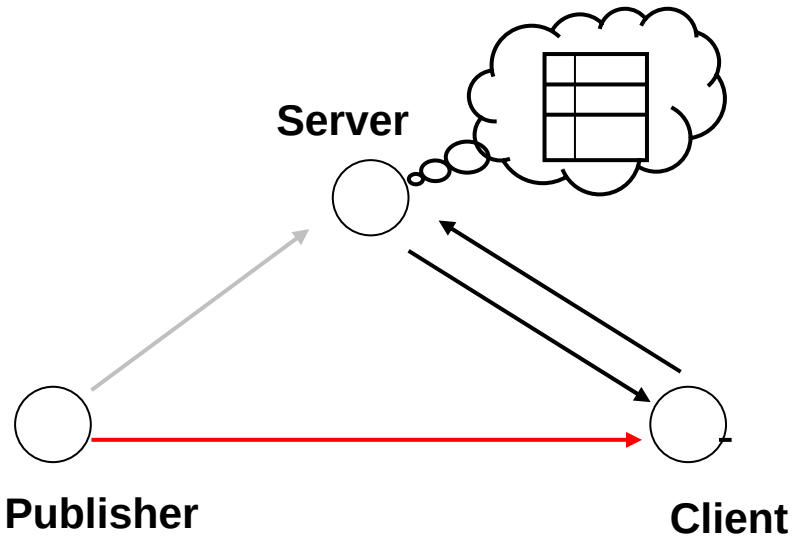
The Lookup Problem



How to find a file by its name in a large p2p file sharing system?

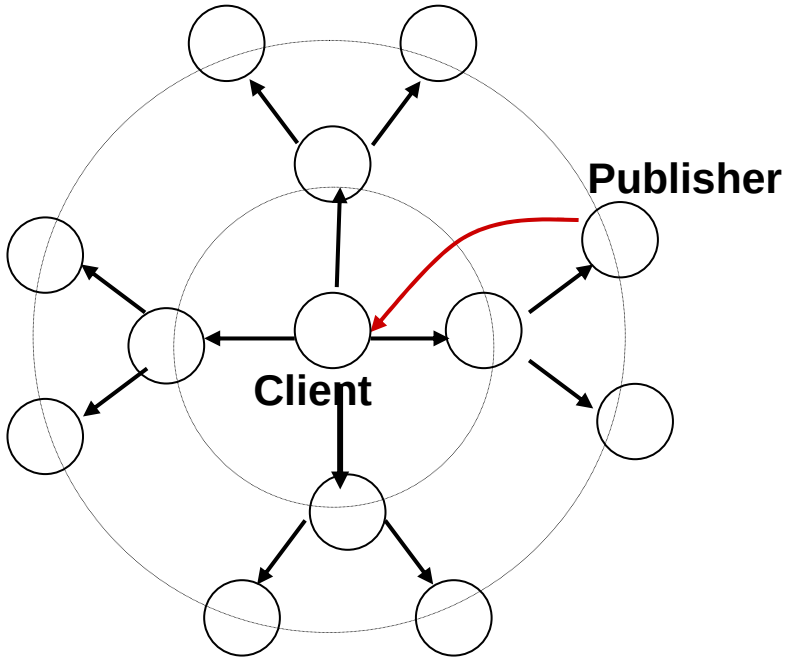
Early Systems: Napster and Gnutella

Napster



Central server bottleneck

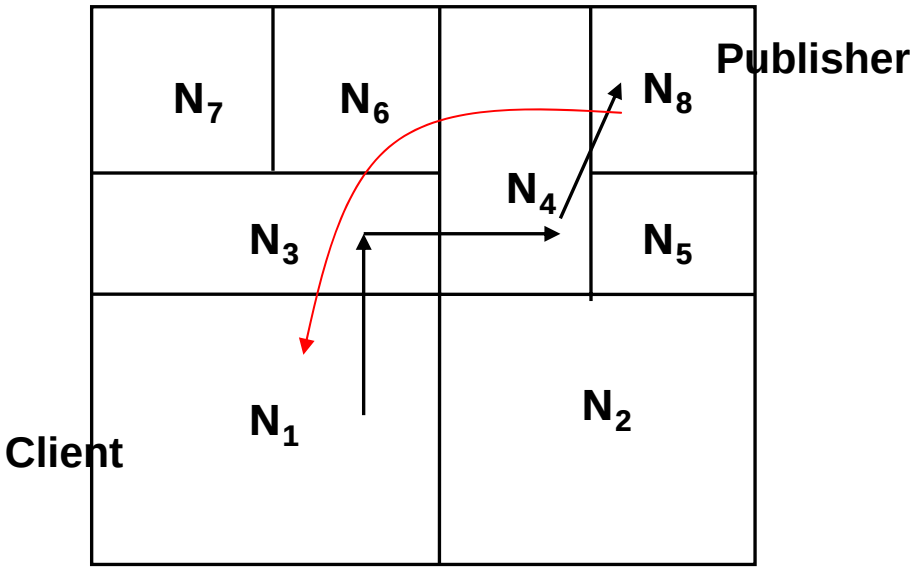
Gnutella



Inefficient flooding

DHT-based Systems: CAN, Chord, Pastry and Tapestry

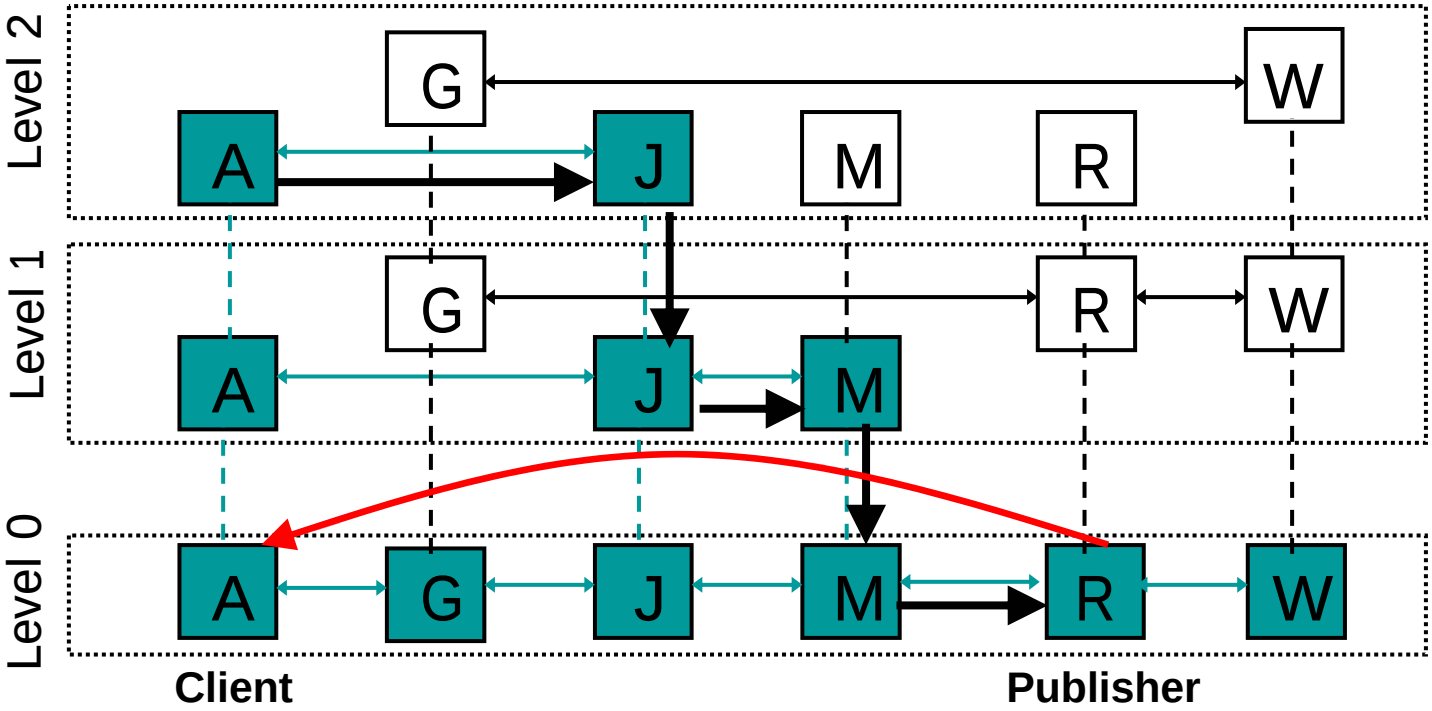
CAN



- Greedy routing: forward the request to node in routing table closest to target
- Hashing does not keep the order of the keys

List-based Systems: Skip Graphs and SkipNet

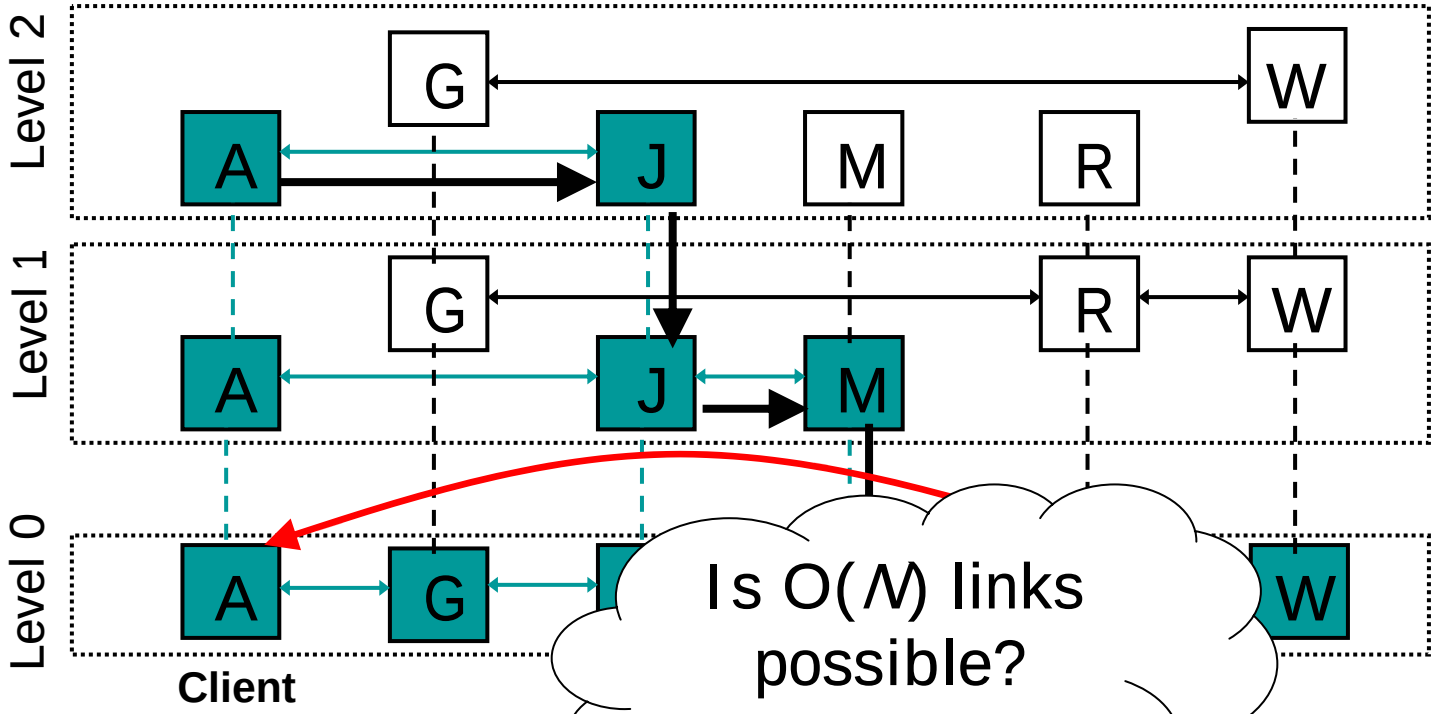
A Skip Graph



- Skip-list-based search: restricting to the lists containing the starting element of the search, we get a skip list.
- Expected hops per search is O(log*N*), and the number of links is **O(NlogN)**

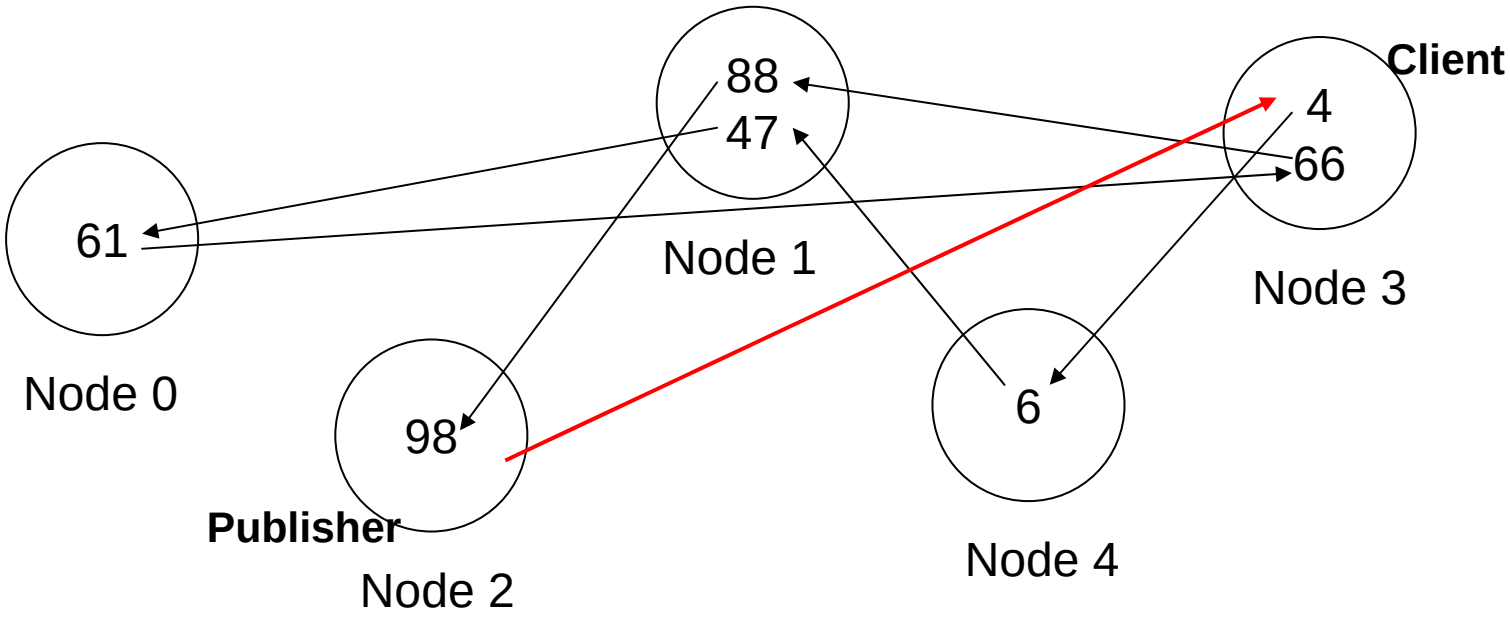
List-based Systems: Skip Graphs and SkipNet

A Skip Graph



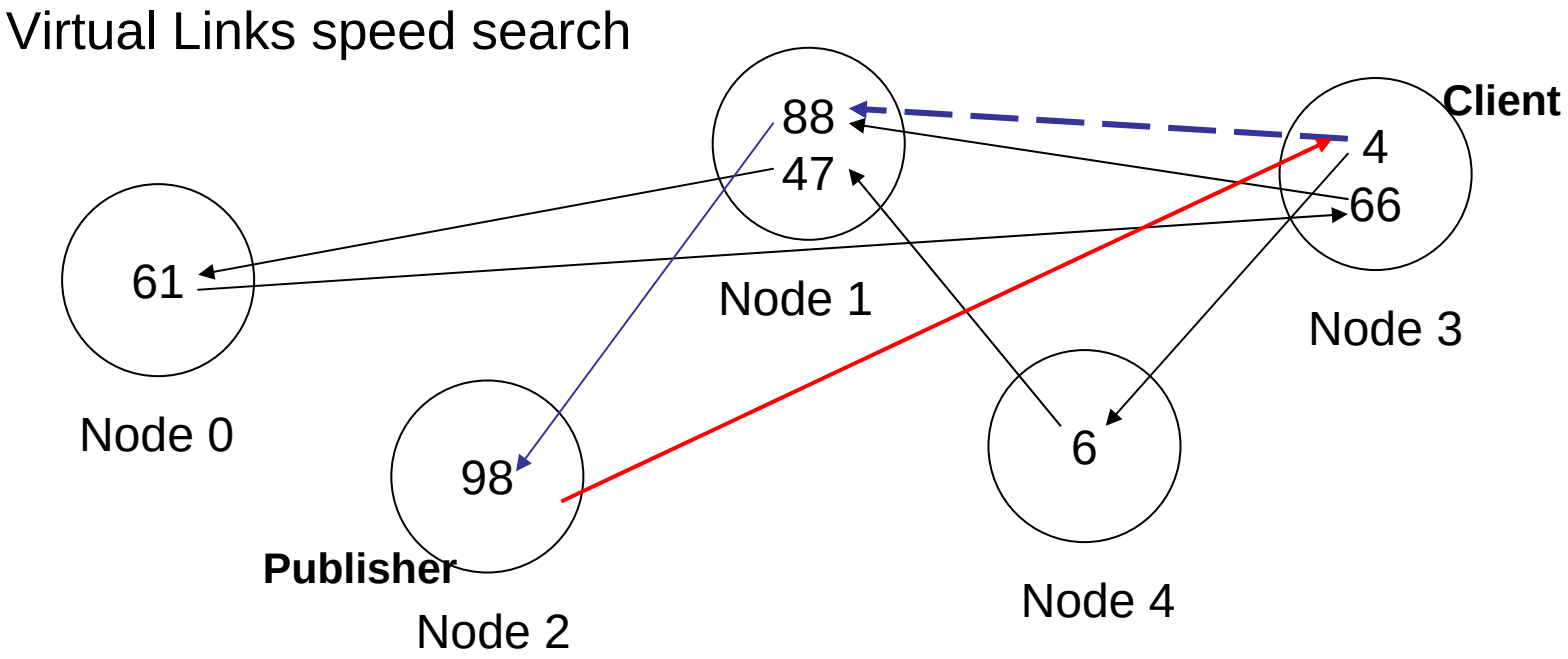
- Skip-list-based search[illegible]s containing the starting element of the [illegible] get a skip list.
- Expected hops per search is O(log*N*), and the number of links is **O(NlogN)**

Physical Link vs. Virtual Link



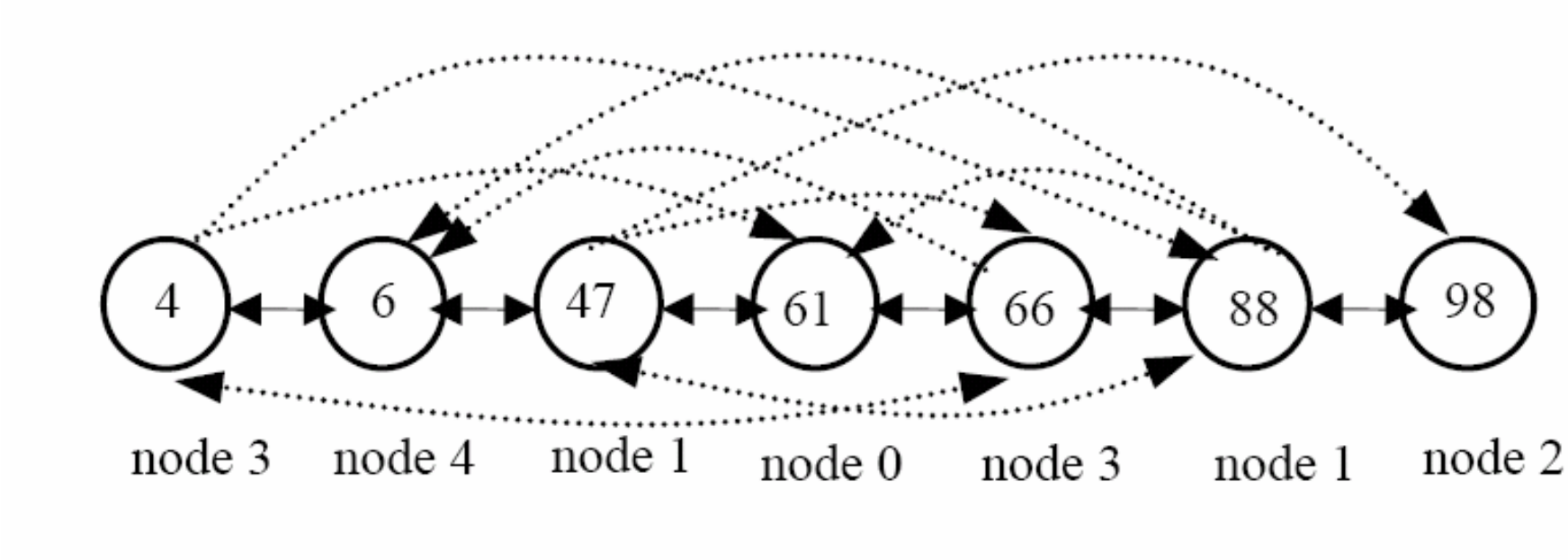
There are 5 nodes and 7 keys. How to search key 98 starting from key 4?

Physical Link vs. Virtual Link



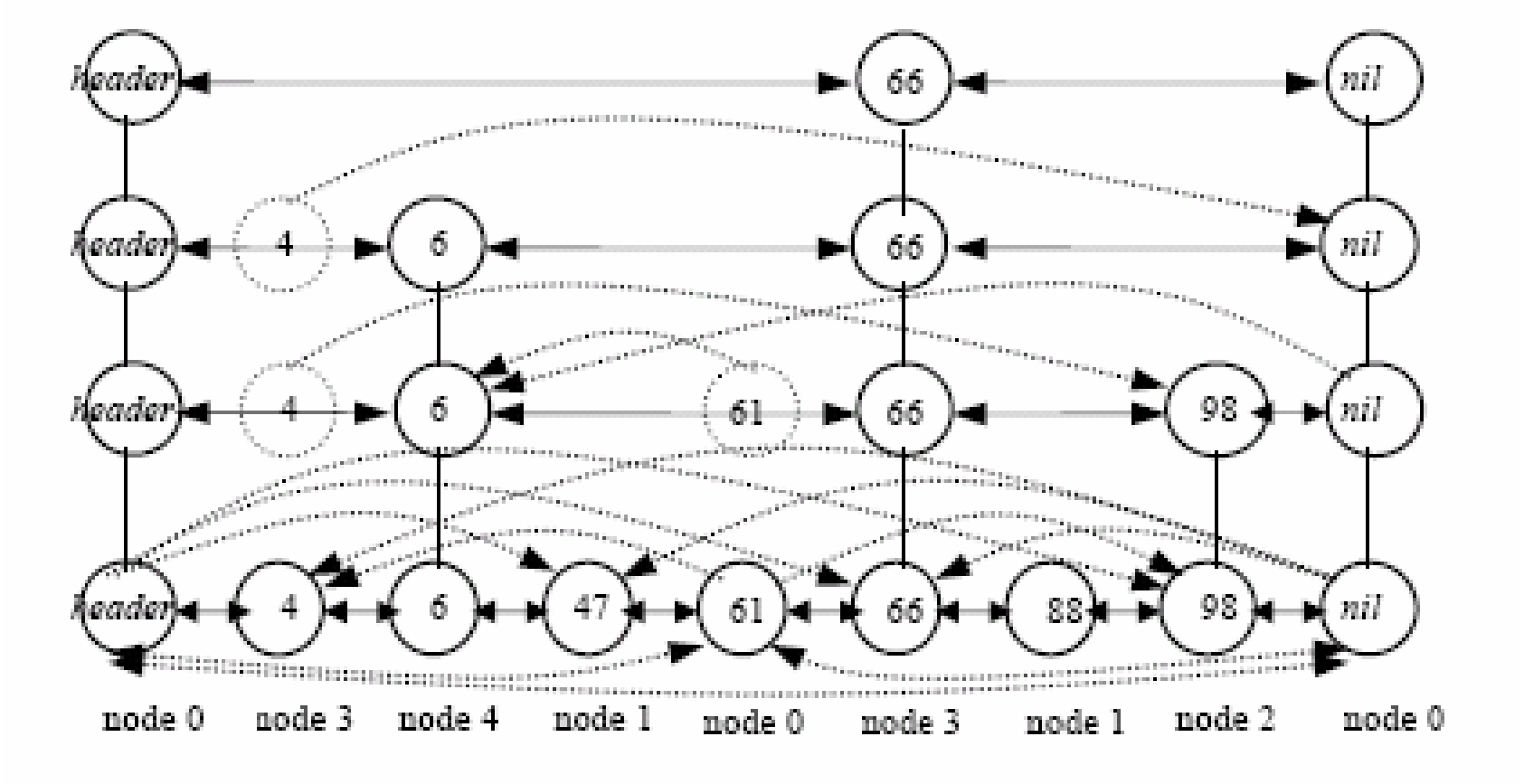
There are 5 nodes and 7 keys. How to search key 98 starting from key 4?

LinkNet: A New List-based data structure



Virtual Links speed search and enhance fault tolerance

Skip-list-based LinkNet



Performance Evaluation

- In an *N* nodes *M* elements network, the expected total space Skip-list-based LinkNet takes is O(*M*), and when *M* is big enough, the search operation takes expected O(log*N*) messages among nodes.

System	Order Preserved?	Expected hops per search	Expected total space
Chord	**No**	**O(log*N*)**	**O(*N*log*N*)**
Skip Graphs	**Yes**	**O(log*M*)**	**O(*M*log*M*)**
LinkNet	**Yes**	**O(log*N*)**	**O(*M*)**

References

- H. Balakrishnan, M. Frans Kaashoek, D. Karger, R. Morris, and I. Stoica. Looking Up Data in P2P Systems. Communications of the ACM, 46(2), February 2003.
- J. Aspnes and G. Shah. Skip Graphs. In Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms, January 2003.
- N. J. A. Harvey, M. B. Jones, S. Saroiu, M. Theimer, and A. Wolman. SkipNet: A Scalable Overlay Network with Practical Locality Properties. In Proceedings of the 4th USENIX Symposium on Internet Technologies and Systems (USITS), March 2003.
- W. Pugh. Skip Lists: A Probabilistic Alternative to Balanced Trees. Communications of the ACM, 33(6):668-676, June 1990.
- W. Pugh. Concurrent Maintenance of Skip List. Technical Report CS-TR-2222, Department of Computer Science, University of Maryland, June 1990