

P2P Network

Structured Networks (III)

Distributed Hash Tables

Pedro García López

Universitat Rovira I Virgili

Pedro.garcia@urv.net

Departament d'Enginyeria



Informàtica i
Matemàtiques



UNIVERSITAT
ROVIRA I VIRGILI

SkipNet

Christian Schmidt-Madsen, csm@itu.dk

Peter Tiedemann, petert@itu.dk

Data locality

- Hashing distributes data uniformly across the overlay -> related data is not located in contiguous nodes
- Idea: group related data in contiguous nodes -> breaks load balanced distribution of keys to nodes
- Routing tables depend on the uniform distribution of ids to establish connections
- Solution: 2 ids, numID is uniform in the identifier space but nameID can be chosen related to the data.

SkipNet structure overview

Dual ID-space:

Name ID

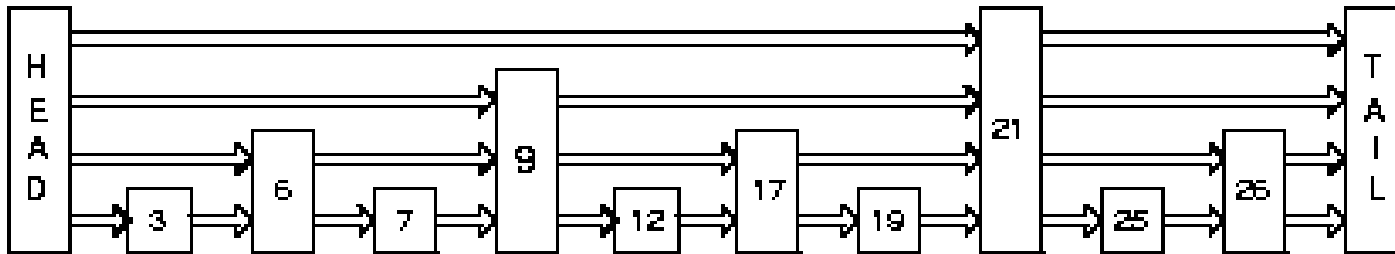
Numeric ID (similar to Chord, Pastry etc)

- **Node names are the reverse DNS name of the host (example: dk.itu.mycomputer)**
- **Name ID Objects are stored at the lexicographic predecessor node.**
(example: dk.itu.mycomputer/index.html will be stored at dk.itu.mycomputer)*
- **Numeric ID objects are stored at the numerically closest node**

*: assuming "/" is less than other symbols

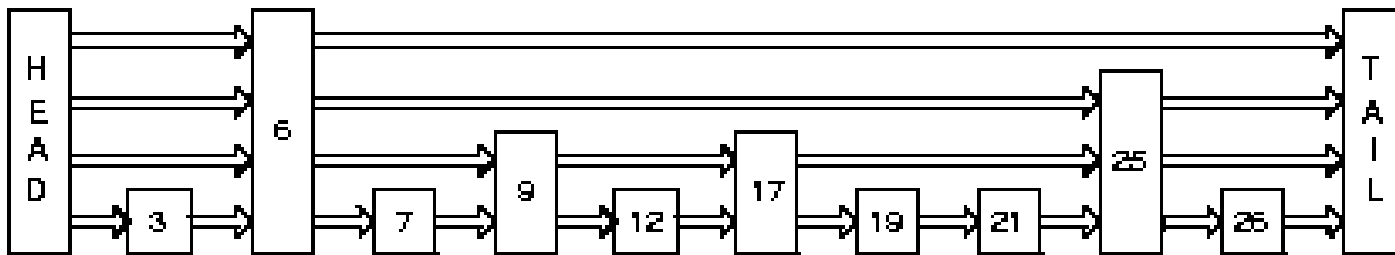
The SkipList datastructure

Deterministic:



Pointer at level h skips over 2^h elements

Probabilistic:



Node at level h with probability $1/2^h$

SkipList -> SkipNet

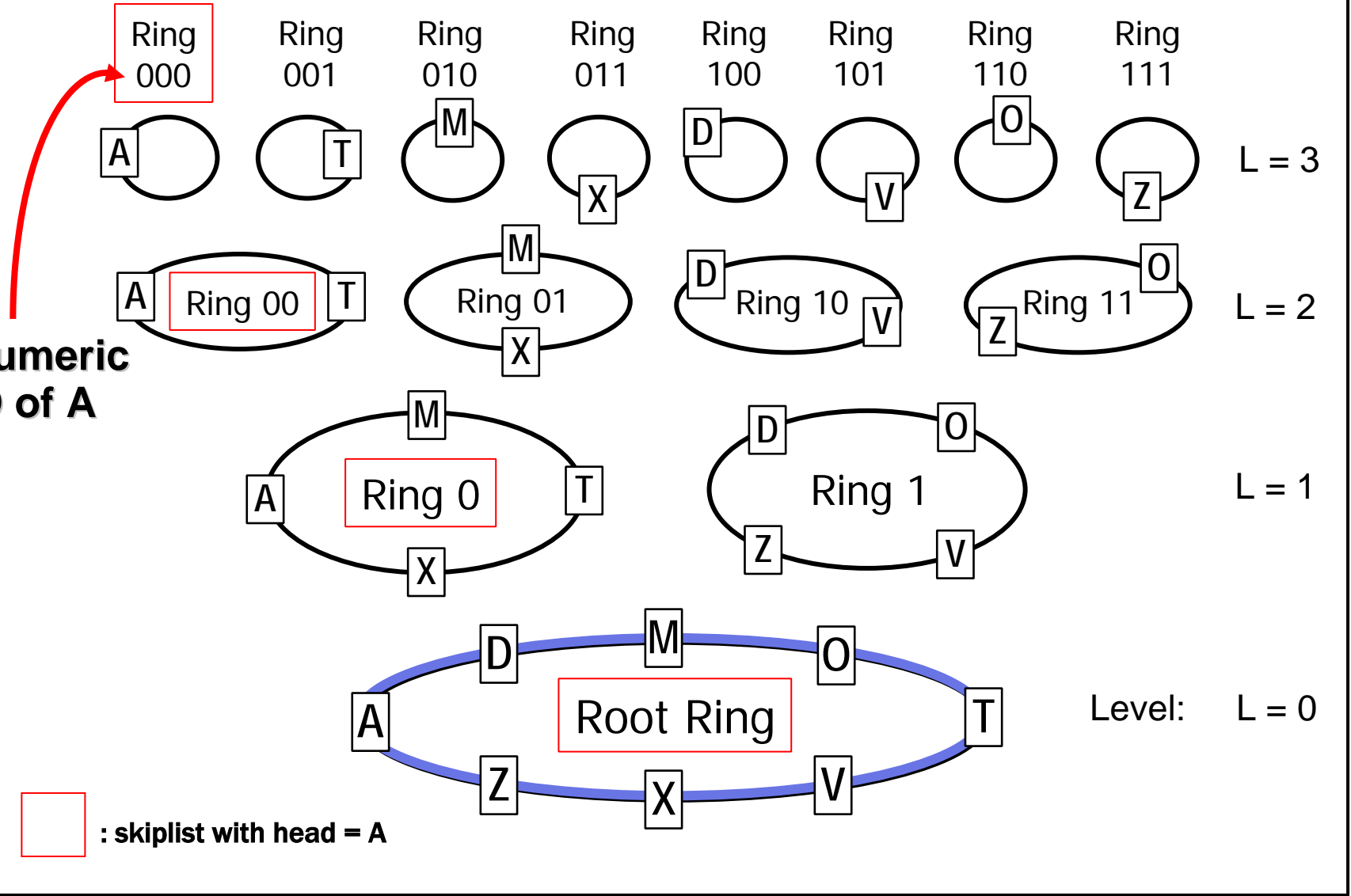
Problems with SkipList:

- Efficient search only possible from head
- Some nodes more likely to be in routing path

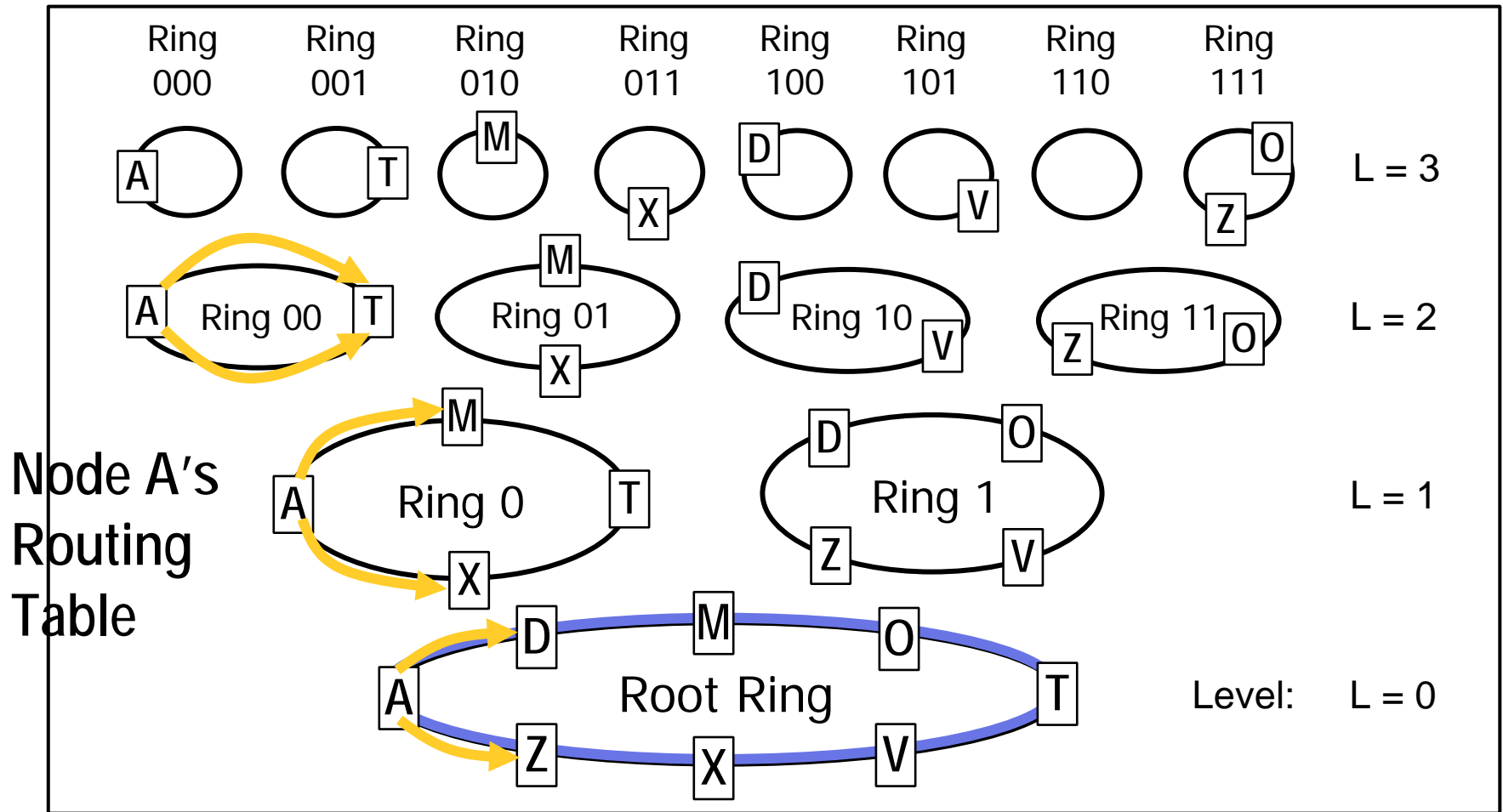
Basic principle of SkipNet:

- A probabilistic skiplist where every node is a head.
- The keys are just the names of the nodes
- The ordering is lexicographic

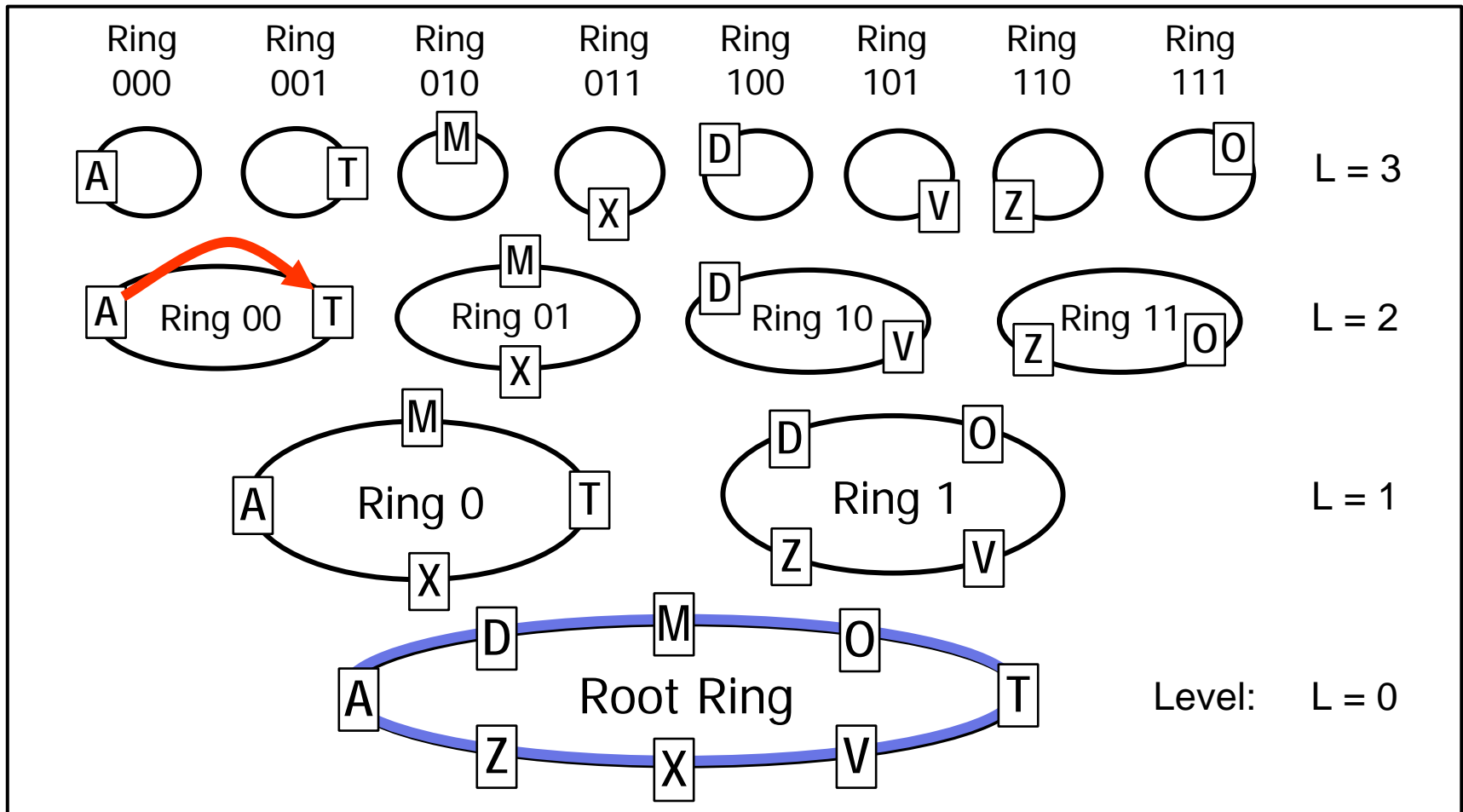
Numeric ID of A



SkipNet Global View – Routing table

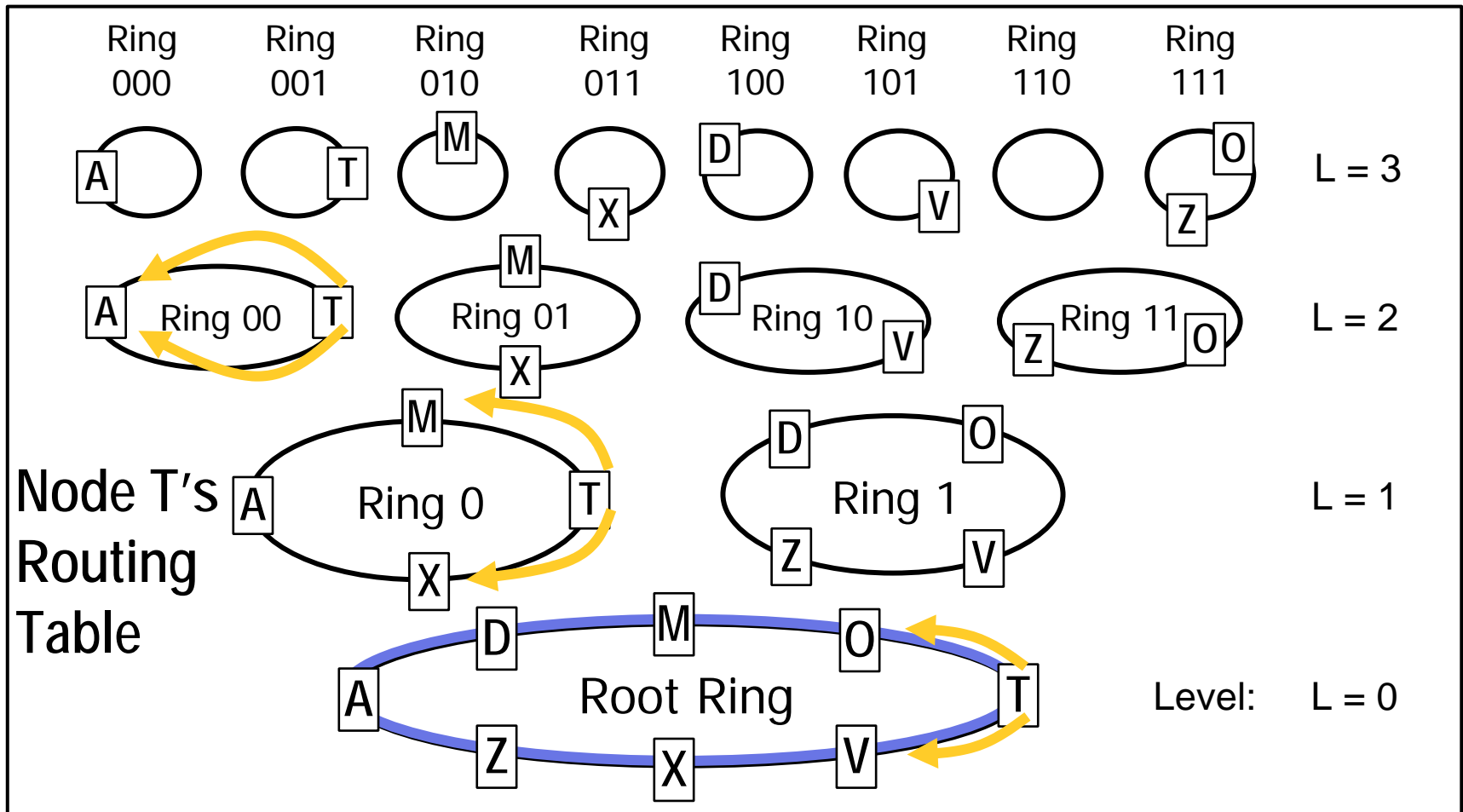


Routing by Name ID to node

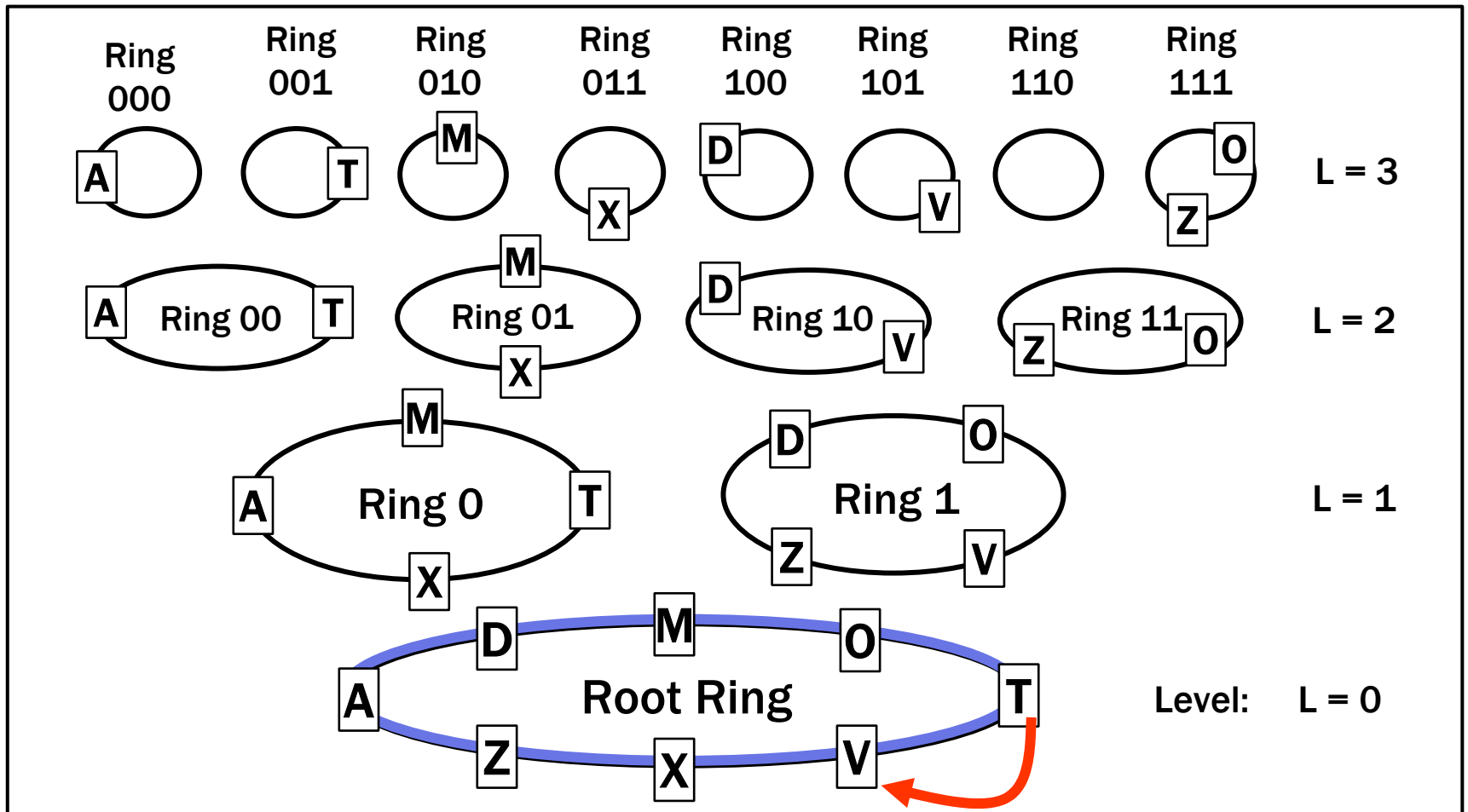


- **Example: route from A to V**
- **Simple Rule: Forward the message to node that is closest to dest, without going too far.**

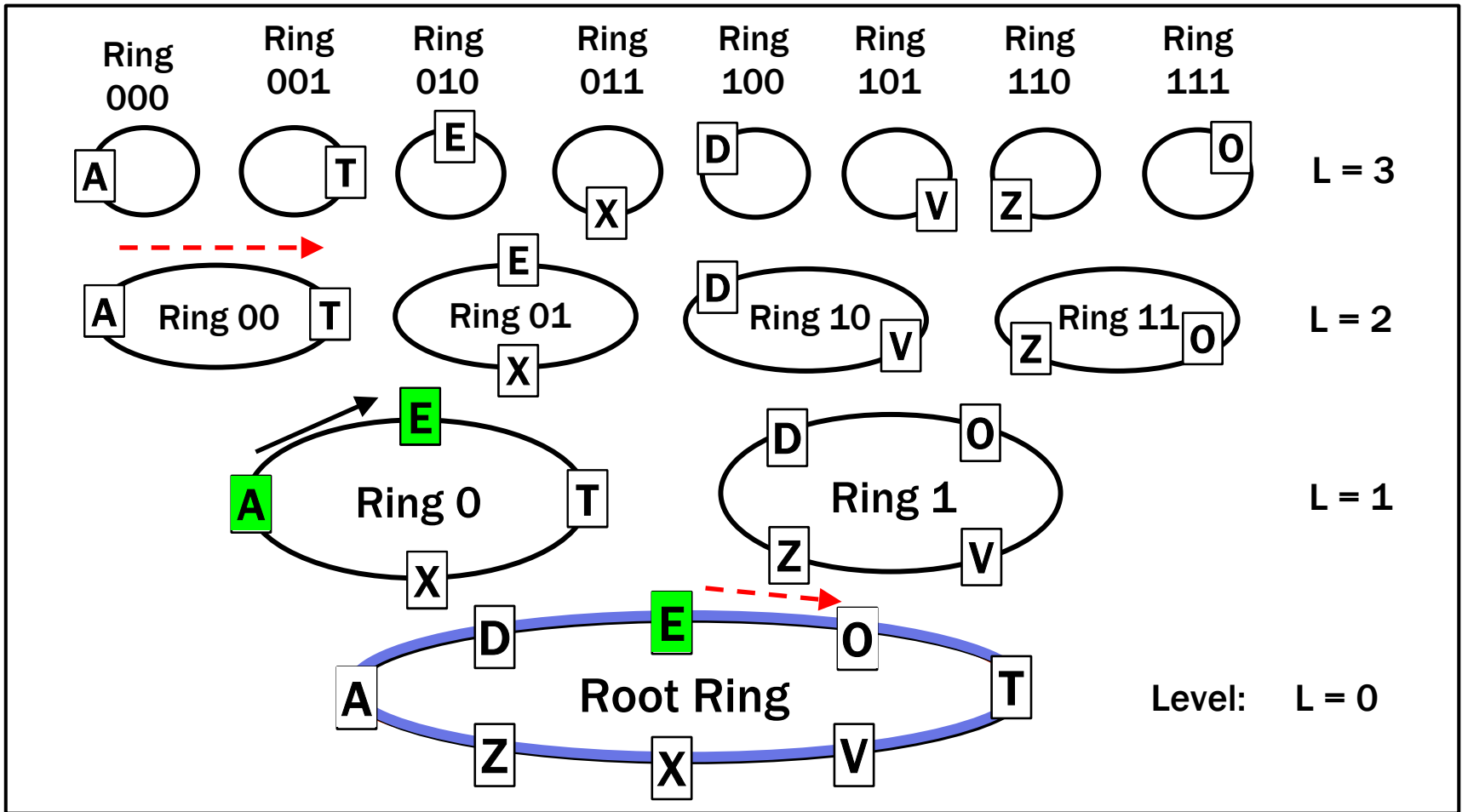
Routing by Name ID to node



- **Example: route from A to V**
- **Simple Rule: Forward the message to node that is closest to dest, without going too far.**



- Example: route from A to V
- Simple Rule: Forward the message to node that is closest to dest, without going too far.

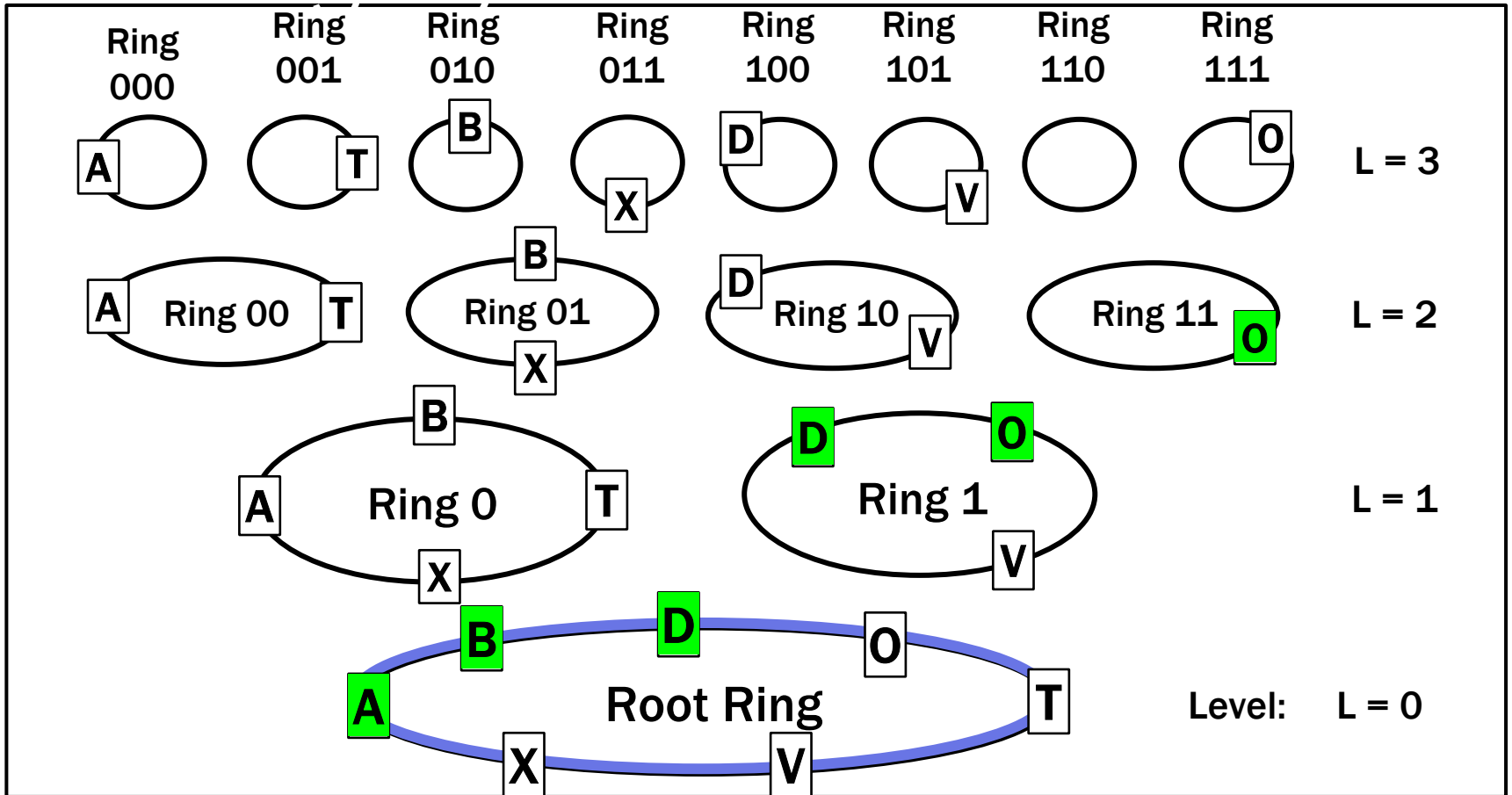


- Example: route from A to F -> Terminates at E
- **Simple Rule: Forward the message to node that is closest to dest, without going too far.**

Routing by Num ID to node

RouteTo(numID):

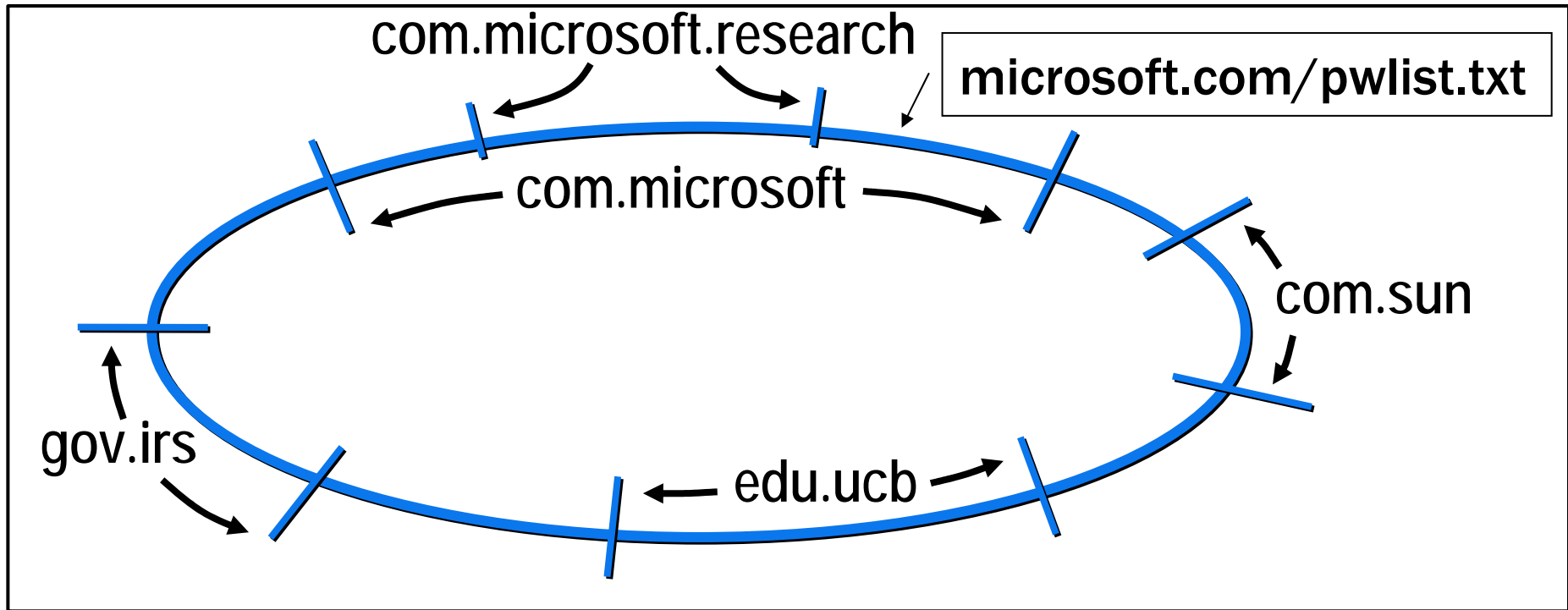
- Current ring = root ring
- Move around the current ring until you meet a node which is in the same next higher ring as numID
- Continue recursively until a top ring is reached



Example: route from A to 110 -> ends at node O

Log(N) rings to visit

In expectation only 2 nodes visited before encountering one belonging to correct ring

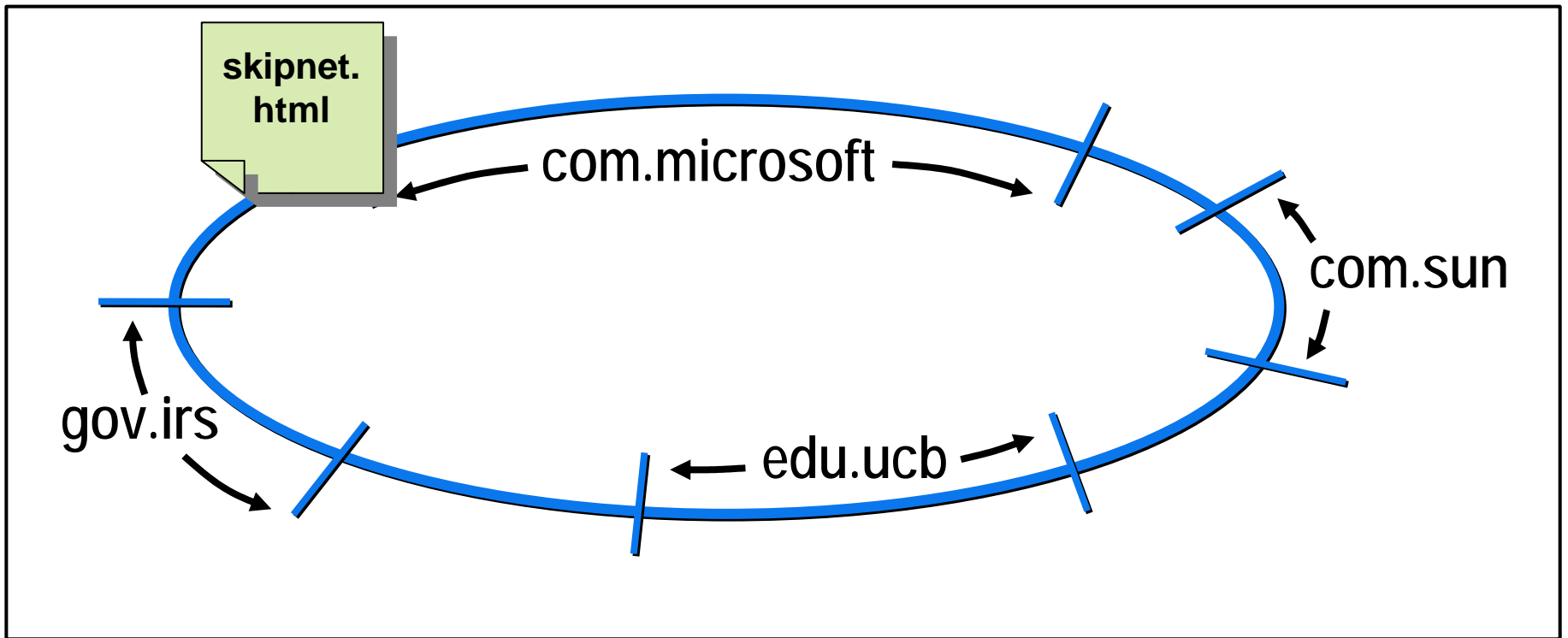


- Organizations correspond to contiguous SkipNet segments
 - Internal routing by nameID remains internal
 - Objects are stored according to nameID

Constrained load balancing

- DHT over a limited domain
- We combine the two id spaces
- Calling a file <CLB Domain>!<CLB Suffix> will store the file on the node beginning with <CLB Domain> and having the closest numeric ID to hash(<CLB Suffix>)





- To get object "microsoft.com!skipnet.html"
 - Route by nameID to "microsoft.com"
 - Route by numericID to Hash("skipnet.html") within the "microsoft.com" constraint

SkipNet conclusions

- Data locality is a very interesting feature for achieving range queries over the overlay
- Distributed indexes and data location can really benefit from skipnet overlay structure
- Retrieval of related data does not imply contacting far away nodes
- **But:** contiguous data could be in contiguous nodes in different countries (USA node – China Node – Australia node)
 - Skipnet proposes a P-Table to fix this (proximity-aware routing) -> questionable !