Distributed Information Processing

15th Lecture

Eom, Hyeonsang (엄현상)
Department of Computer Science & Engineering
Seoul National University

Outline

- Peer-to-Peer Computing
  - Tapestry
- Q&A

Tapestry

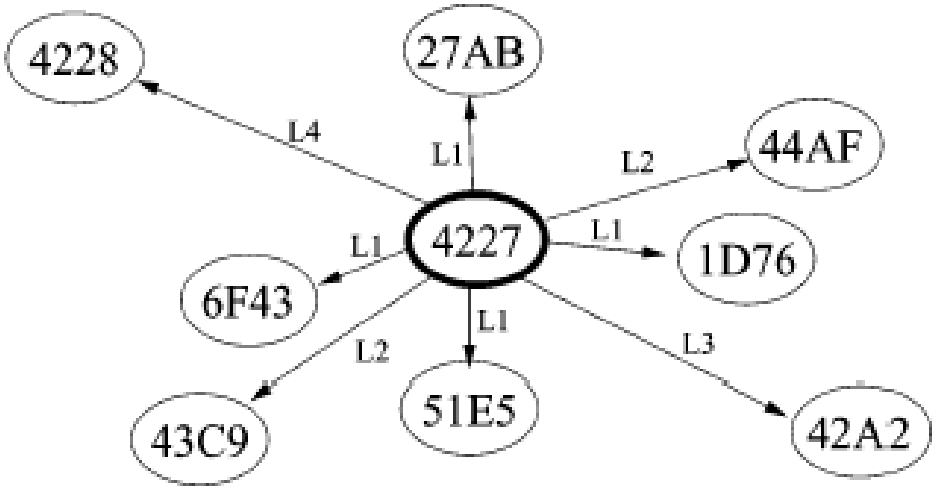
- Peer-to-Peer Overlay Infrastructure
  - Distributed Object Location & Routing (DOLR)
    - Efficient
    - Scalable
    - Self-administering
    - Fault-tolerant
    - Using locally optimal routing tables
      - C.f., Chord routing on the shortest overlay hops

Reducing Routing Stretch
– Ratio between the Actual Latency or Distance and the Shortest Distance

Routing Messages Directly to Nearby Copies of an Object or Service

Tapestry [Zhao04]

- Routing & Object Location
  - Routing Meshes
    - Routing overlay messages to dest ID digit by digit
      - E.g., 4*** => 42** => 42A* => 42AD



Closest in Network Distance

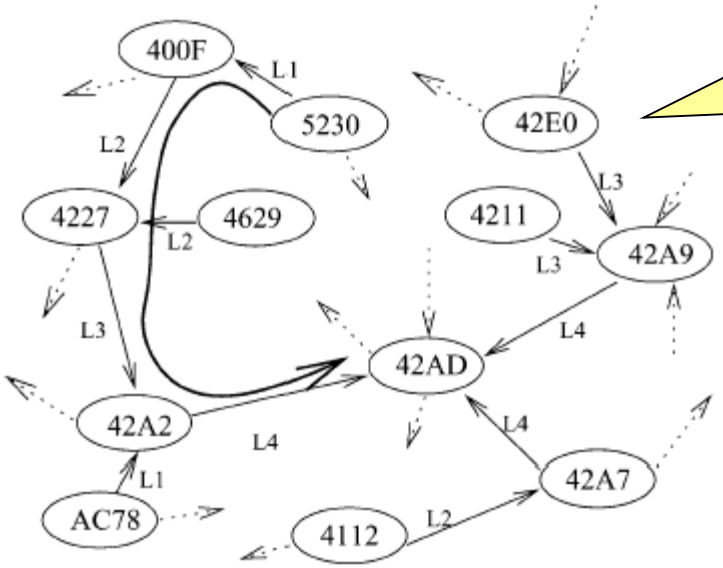
Level	0	1	...	F
1	0***	1D76		F***
2	40**	41**		4F**
3	420*	421*		42F*
4	4220	4221		422F

Neighbor Map Held by Node 4227

Tapestry routing mesh from the perspective of a single node. Outgoing *neighbor links* point to nodes with a common matching prefix. Higher level entries match more digits. Together, these links form the local routing table.

Tapestry (Cont'd)

- Routing & Object Location
  - Message Path
    - Looking in the next level matching the next digit
      - E.g., 4*** => 42** => 42A* => 42AD

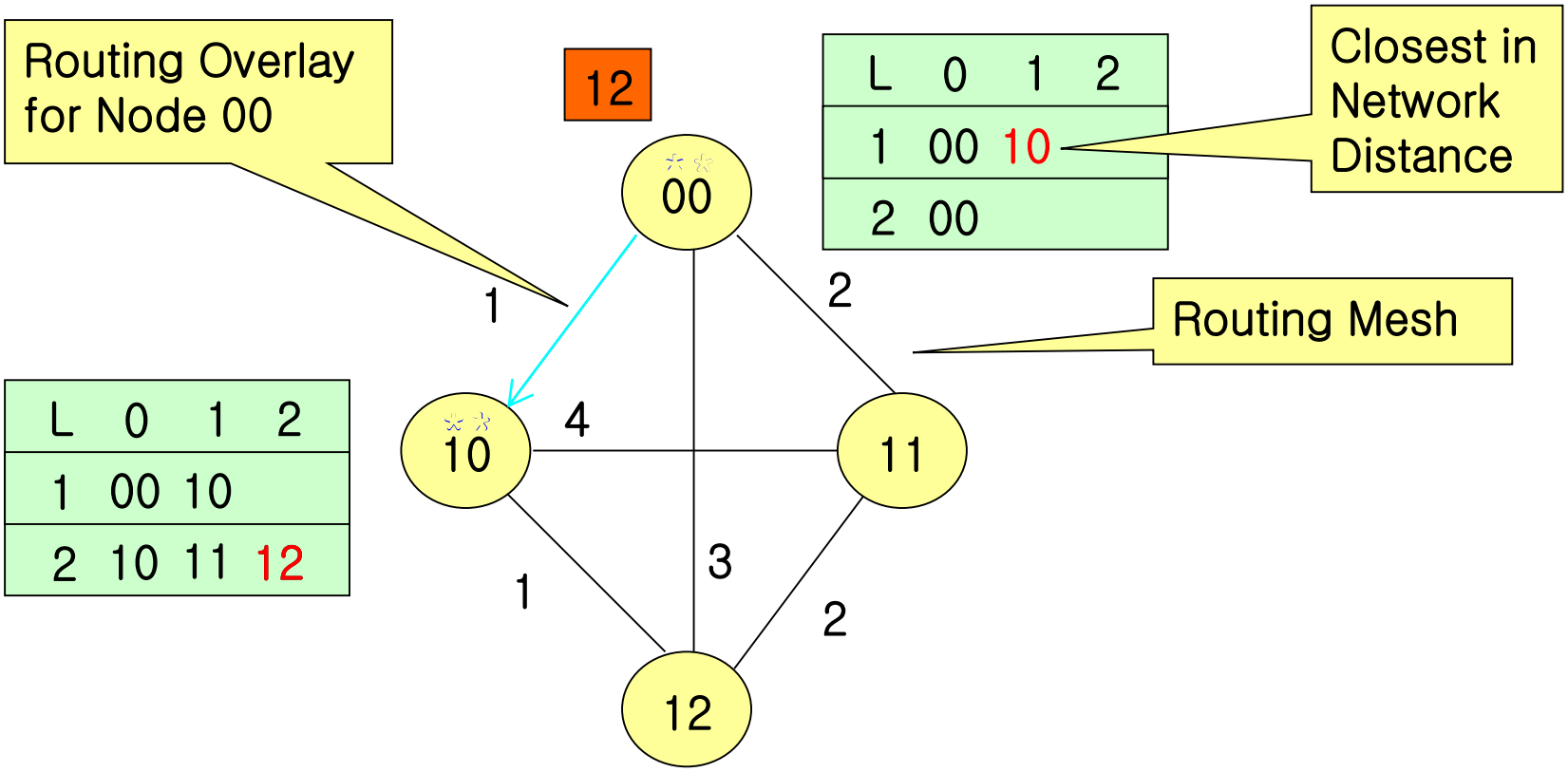


Any Node Can Be Reached in at Most $\log_{\beta} N$, β Is the Base for IDs and N is the Namespace Size

Path of a message. The path taken by a message originating from node 5230 destined for node 42AD in a Tapestry mesh.

Neighbor Map Construction

- Example (β =3, N = 9)



Tapestry [Zhao01] (Cont'd)

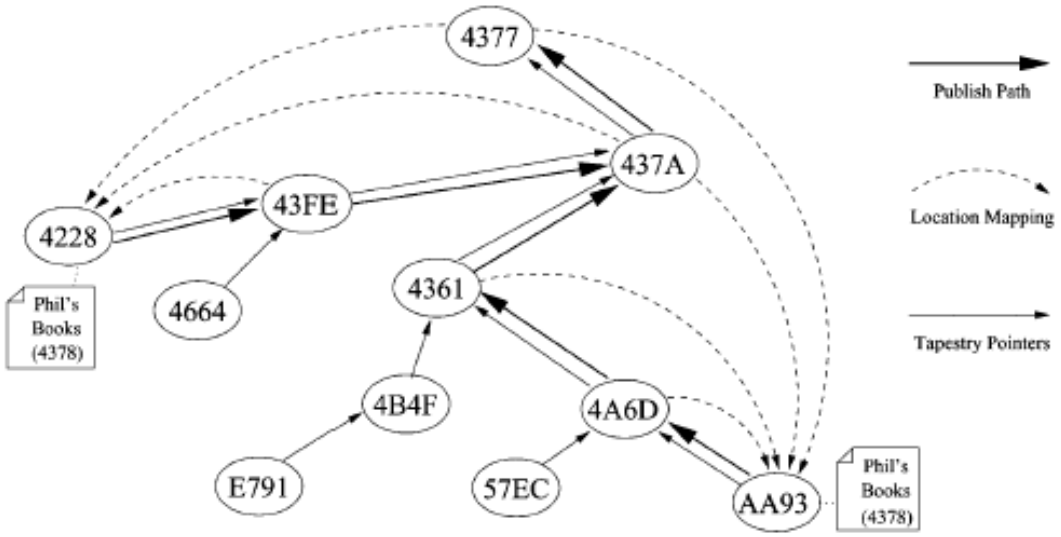
- Routing & Object Location
  - Surrogate Routing
    - To incrementally compute a unique root node for an object, as a guaranteed or surrogate one from which the object can be located
      - Tentatively using the object ID as the root ID
      - Deterministically selecting an existing (alternative) link
        - Looking for a "close" digit
      - Terminating when a neighbor map is reached where the only non-empty entry belongs to the current node, which is then designated as the surrogate root

This Selection Must Be Deterministic, Scalable, and Consistent

Quiz: Does Surrogate Routing Provide a Technique by Which Any ID Can Be Uniquely Mapped to an Existing Node in the Network?

Tapestry [Zhao04](Cont'd)

- Routing & Object Location
  - Object Publication & Location
    - Server's periodic publication of its object
      - Routing a publish message toward the object's root

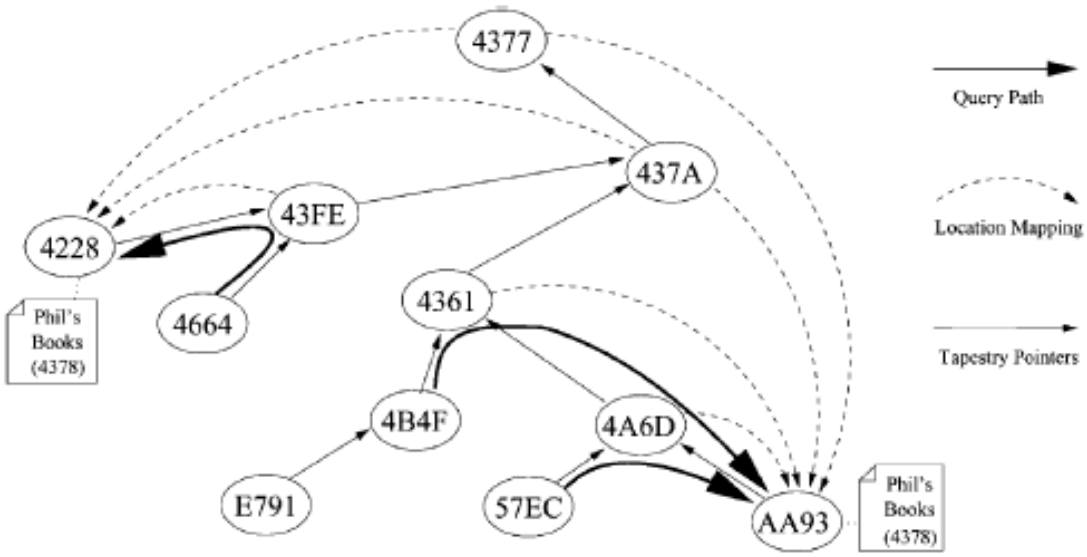


Tapestry object publish example. Two copies of an object (4378) are published to their root node at 4377. Publish messages route to root, depositing a location pointer for the object at each hop encountered along the way.

Tapestry (Cont'd)

- Routing & Object Location
  - Route to an Object
    - Locating a message to the object
      - Redirection

Queries Are Routed to Nearby Object Replicas



Tapestry route to object example. Several nodes send messages to object 4378 from different points in the network. The messages route toward the root node of 4378. When they intersect the publish path, they follow the location pointer to the nearest copy of the object.

Tapestry (Cont'd)

- Properties
  - Locality Awareness
    - Satisfying queries with a nearby object replica
  - Fault Tolerance
    - Routing with two backup neighbors
    - Locating with multiple roots
    - Publishing location information periodically
  - Scalability
    - Routing only with locally available data
  - Adaptability
    - Using dynamic algorithms for node insertion, neighbor-map population, & neighbor notification

References

- [Zhao04] B.Y. Zhang, *et al.*, "Tapestry: A Resilient Global-Scale Overlay for Service Deployment," IEEE JSAC, Vol. 22, No. 1, January 2004
- [Zhao01] B.Y. Zhao, *et al.*, "Tapestry: An Infrastructure for Fault-Tolerant Wide-Area Location and Routing," Tech. Rep. CSD-01-1141, University of California at Berkeley, April 2001