

Overlay & P2P Systems Background

Brighten Godfrey
cs598pbg Nov 4 2010

slides ©2010 by Brighten Godfrey unless otherwise noted





- **Peer-to-peer system:** participants have the same functionality and role in the system
 - ...as opposed to client-server architecture
 - Commonly used to imply file sharing but also used in other contexts (e.g., “BGP peering session”)
- **Overlay network:** a virtual network whose links are end-to-end paths in another network
- **Peer-to-peer networks:** Intersection of the above two terms
 - Or, can also mean “file sharing systems”

In the beginning...



- **Napster (1999)**
 - Centralized index server to find the right peer
 - Peer-to-peer file transfer
- **Gnutella (2000)**
 - Fully decentralized P2P indexing: scoped flooding
 - Problems?
- **Freenet (1999)**
 - Goal: censorship-resistant key-value content store
 - Routing: heuristic clustering of similar keys

Key properties of a DHT

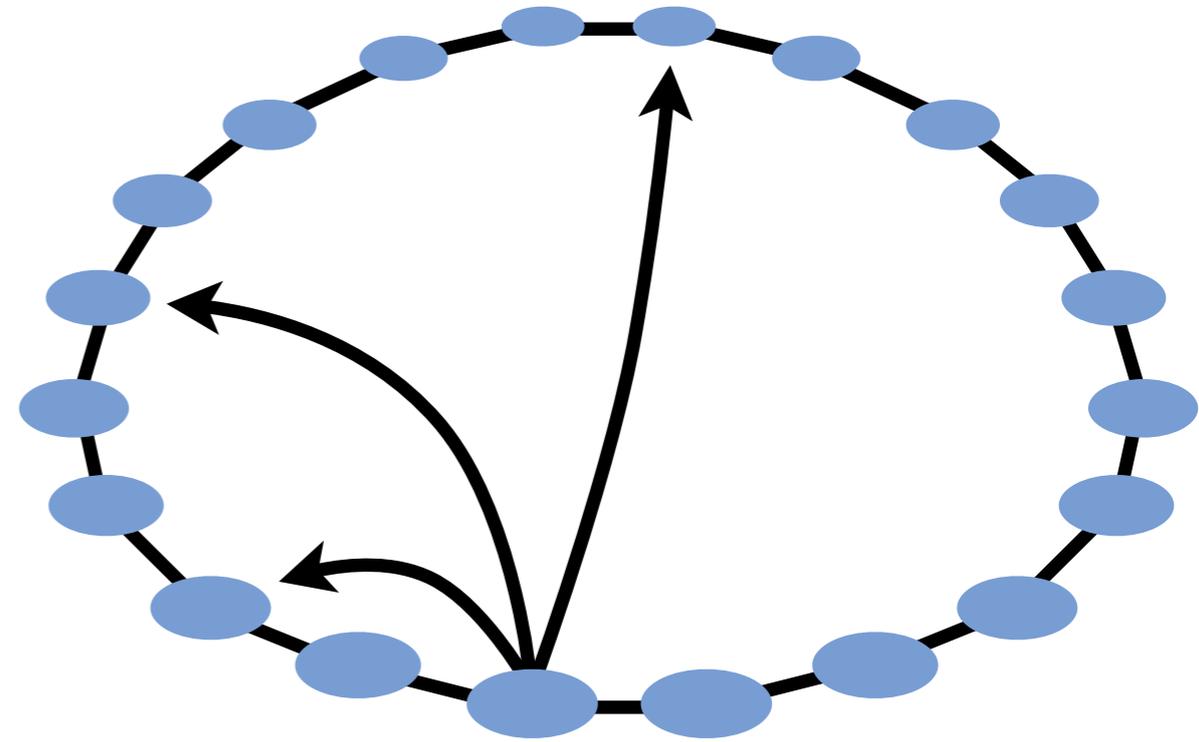


- **Hashtable interface** (fast $put(k,v)$, $get(k)=v$)
 - Unlike Freenet, guarantees results (and quickly)
- **Scalable**
 - Low memory / communication
 - Uses **consistent hashing**: transfers in expectation $1/n$ of objects when a node leaves/joins
- **Resilient**
 - Still works if, say, 50% of the nodes suddenly fail
- **Decentralized**

DHTs: carefully structured



- Greedy routing based on distance in keyspace
- (Where did we see greedy routing before?)
 - Geographic routing
 - Small world models
 - Grid / torus
- What do we need the DHT topology to look like for this to work? ... work well?

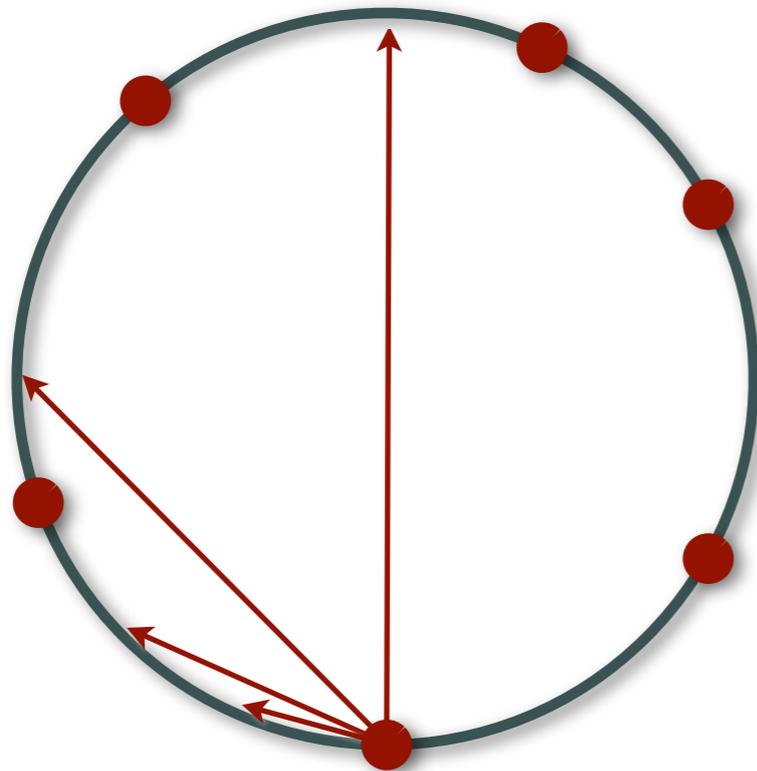


What's missing
from this
Chord topology?

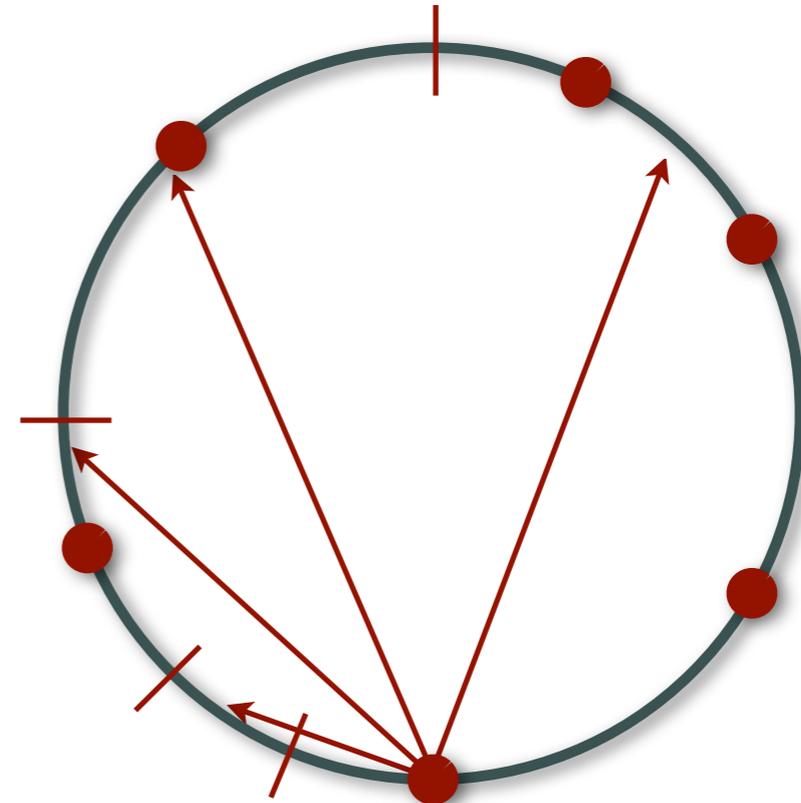
DHTs: check your understanding



Standard Chord topology



Randomized topology



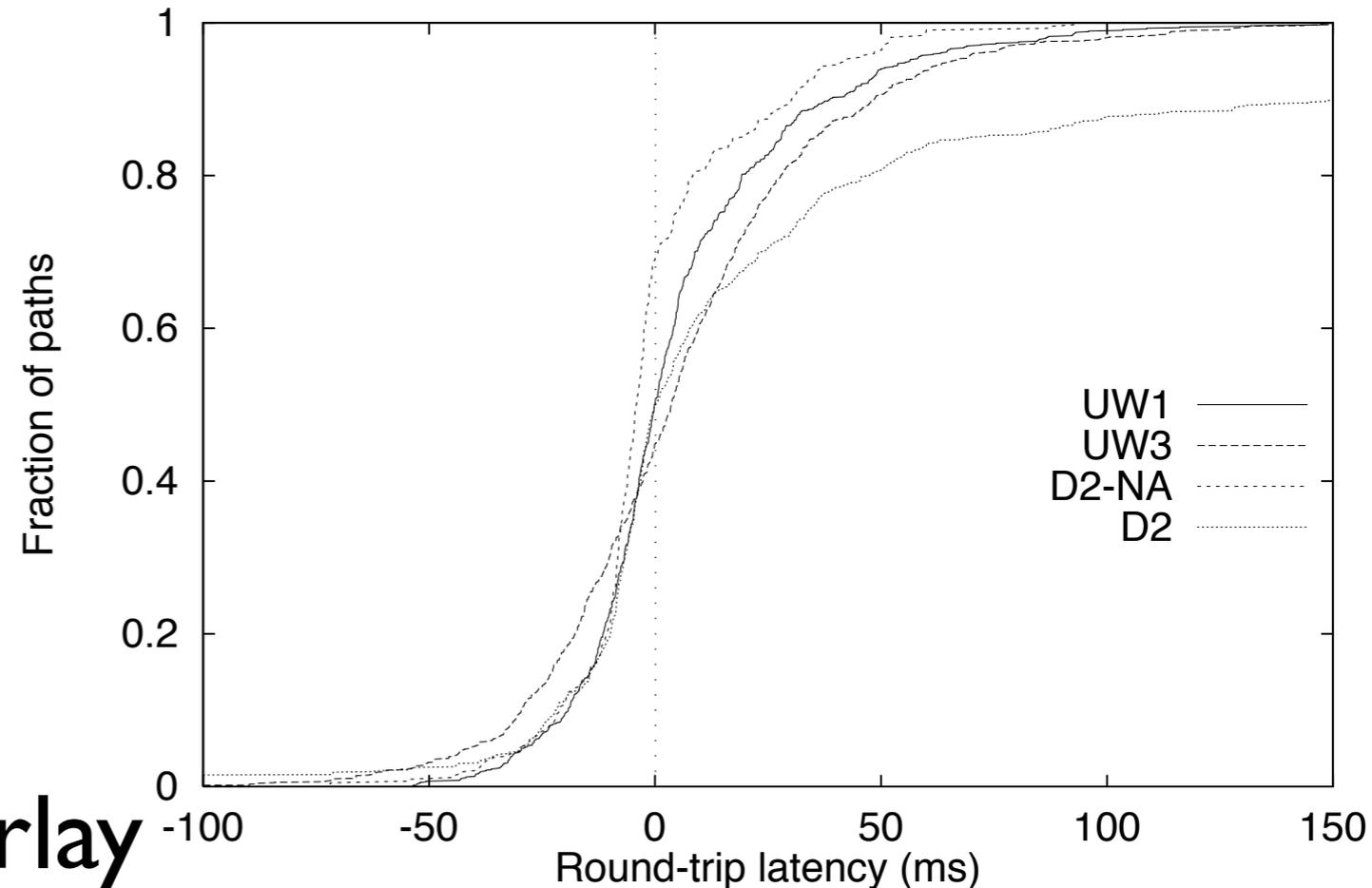
- Randomized topology still guarantees $O(\log n)$ lookup hops. Why?
- Why would you do something other than the deterministic Chord topology?

In the *other* beginning...



- Internet routing is suboptimal

[Savage et al '99, "E2E effects of Internet path selection"]



- Improve it with an overlay
- Key: Internet does not obey the triangle inequality
 - Can happen: $d(a,b) + d(b,c) < d(a,c)$



- **Common theme of many overlay networks: provide more advanced services than the Internet provides**
 - Much easier to deploy new functionality at hosts
 - The Internet doesn't even know what's happening to it
- **Examples**
 - **RON**: more reliable, efficient routing
 - **DHT**: flat name routing and key-value store
 - **i3**: indirection, mobility, middlebox support, ...
 - **Content distribution**: a kind of time-delayed multicast

DHT & overlay in the real world



- Content distribution: Akamai, CoralCDN
- Swarming: DHT for BitTorrent distributed tracker
- File sharing: DHTs in Kad, Overnet/eDonkey
- Amazon Dynamo
- Storm botnet: command & control delivered via DHT
- ...and big impact on many research systems & papers
- Many ideas from DHT / overlay research incorporated into other work, if not entire DHT system

Questions from Tuesday



- Why can RON improve on the Internet's routing latency? Is this a good thing?
- When the objects stored in a DHT are large, it takes a lot of bandwidth to move the objects when a node joins or leaves. Is this a problem? How would you fix it?
- Suppose we use Random Replacement for Chord's fingers. What kind of node misbehavior is possible, and how would you fix it?