Providing Administrative Control and Autonomy in Structured Peer-to-Peer Overlays

Alan Mislove and Peter Druschel

Rice University

Problem

- Structured p2p overlays designed so that
  - Participating organizations contribute resources
  - Use the overlay services in return

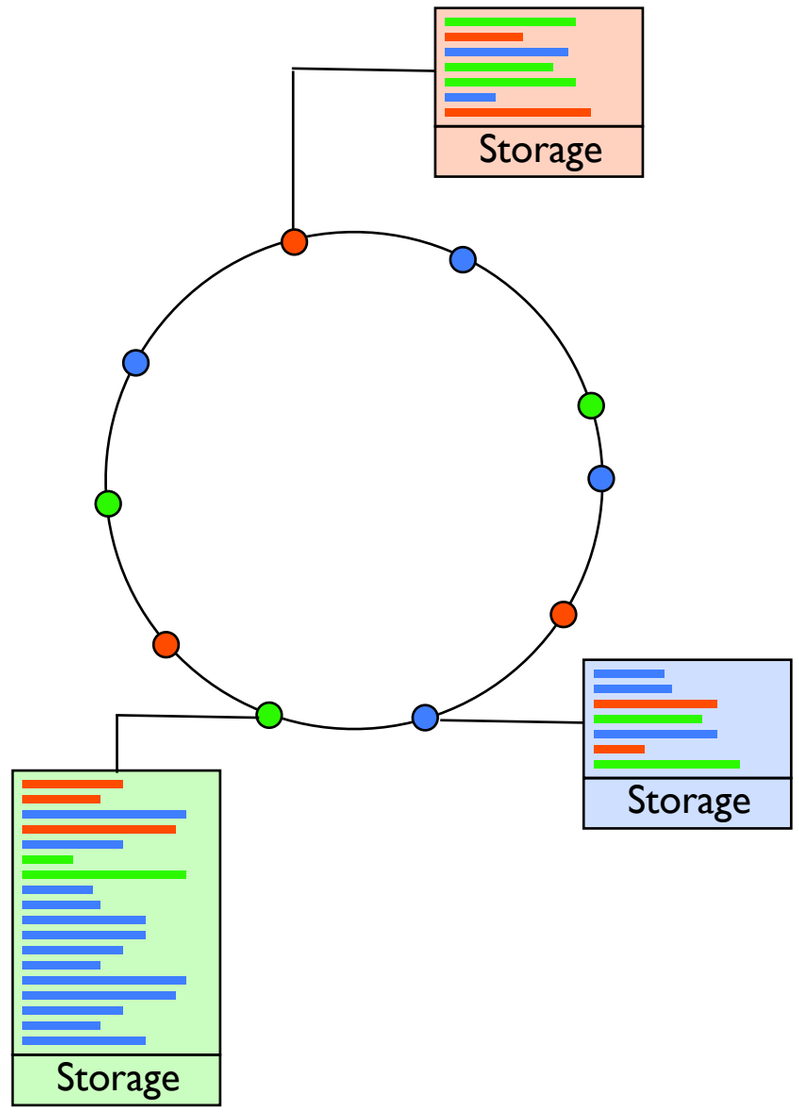
- Concerns over organizational autonomy
  - Unable to enforce membership policy
  - Unable to specify minimum node characteristics
  - Unable to choose protocol that best suites their needs

- Environment of interest is p2p system predominately consisting of large member organizations

2

Problem: Lack of Organizational Autonomy

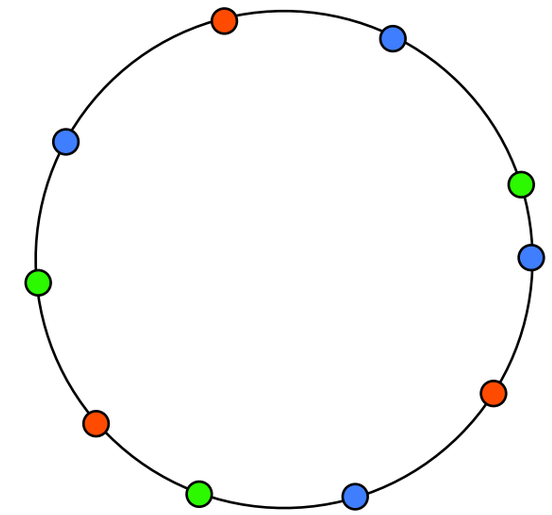
- Resource sharing at global scope
  - Good for load balancing and geographic diversity
  - Lack of organizational control may result in
    - Poor performance (slow nodes)
    - Reduced robustness (correlated failures and untrusted nodes)
    - No accountability
  - Poor write locality
  - Have to adopt system-wide protocol and parameters
    - Unable to choose protocol and parameters that best suit needs
- Lack of path locality



3

Problem: Lack of Organizational Autonomy

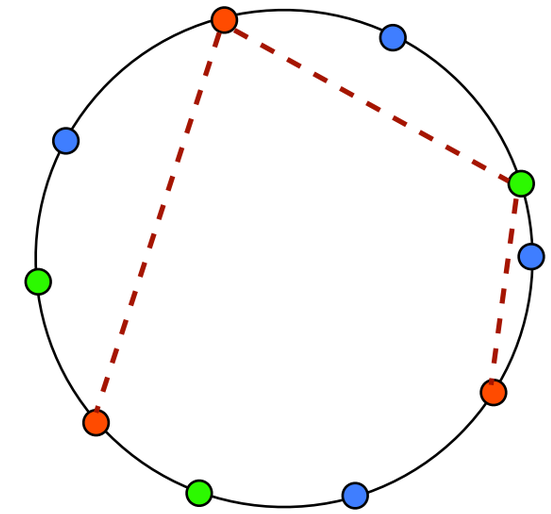
- Resource sharing at global scope
  - Good for load balancing and geographic diversity
  - Lack of organizational control may result in
    - Poor performance (slow nodes)
    - Reduced robustness (correlated failures and untrusted nodes)
    - No accountability
  - Poor write locality
  - Have to adopt system-wide protocol and parameters
    - Unable to choose protocol and parameters that best suit needs
- Lack of path locality



3

Problem: Lack of Organizational Autonomy

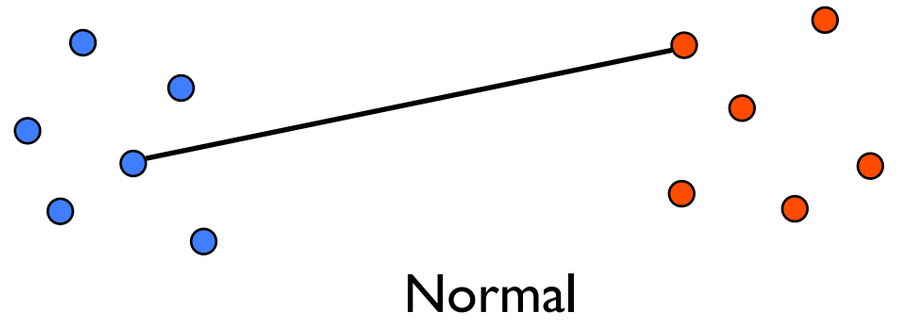
- Resource sharing at global scope
  - Good for load balancing and geographic diversity
  - Lack of organizational control may result in
    - Poor performance (slow nodes)
    - Reduced robustness (correlated failures and untrusted nodes)
      - No accountability
  - Poor write locality
  - Have to adopt system-wide protocol and parameters
    - Unable to choose protocol and parameters that best suit needs
- Lack of path locality

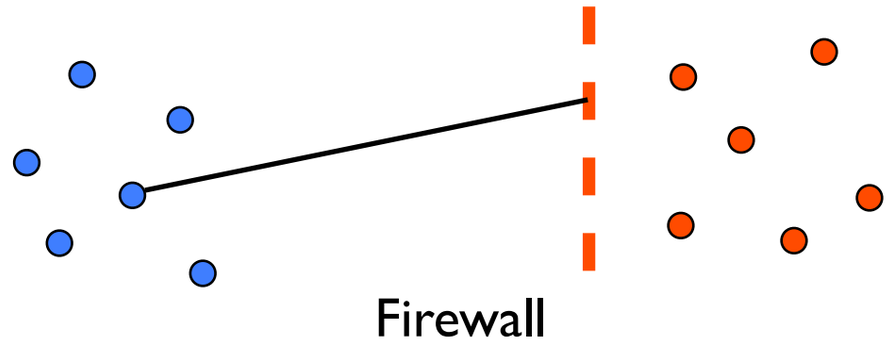


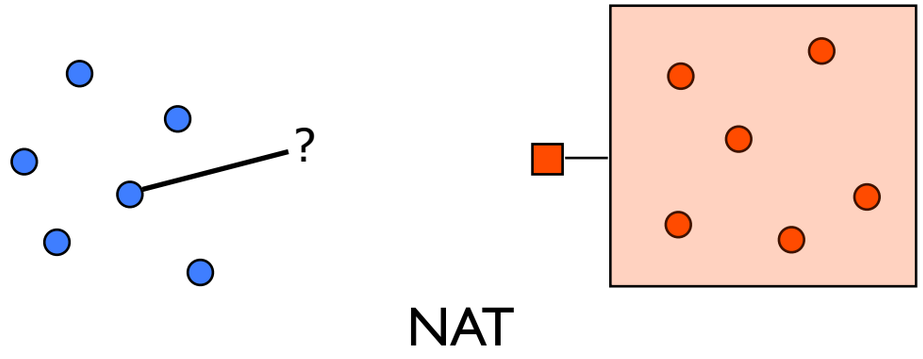
3

Problem: Connectivity Constraints

- In the general Internet connectivity is often constrained
  - Firewalls at at organizational boundaries
  - Network Address Translation
- Deploying overlays currently requires additional engineering
  - Rendez-vous points
  - Pushing
  - Tunnels

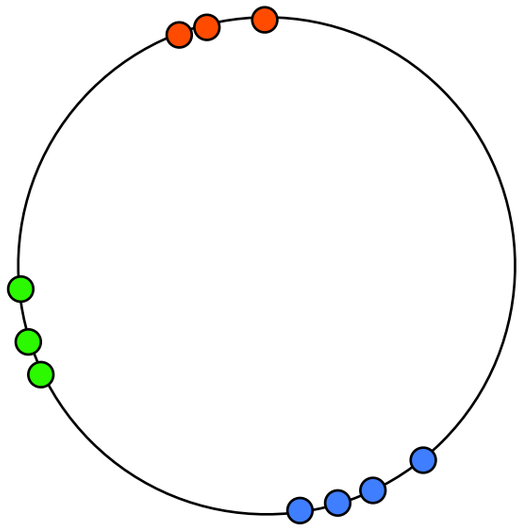






4

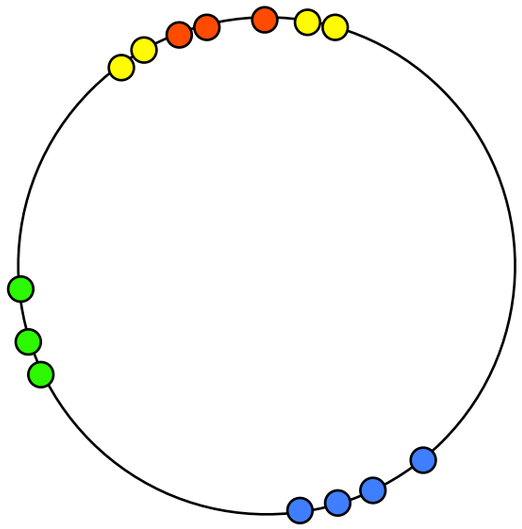
SkipNet



- SkipNet
  - Achieves content and path locality
  - Uses location-based id assignment
  - Need for explicit load balancing constrains design space
  - Security problems
  - Can't leverage existing work on other overlay protocols (e.g. secure routing)
  - Still requires static choice of overlay and parameters

5

SkipNet



- SkipNet
  - Achieves content and path locality
  - Uses location-based id assignment
  - Need for explicit load balancing constrains design space
  - Security problems
  - Can't leverage existing work on other overlay protocols (e.g. secure routing)
  - Still requires static choice of overlay and parameters

5

Goals

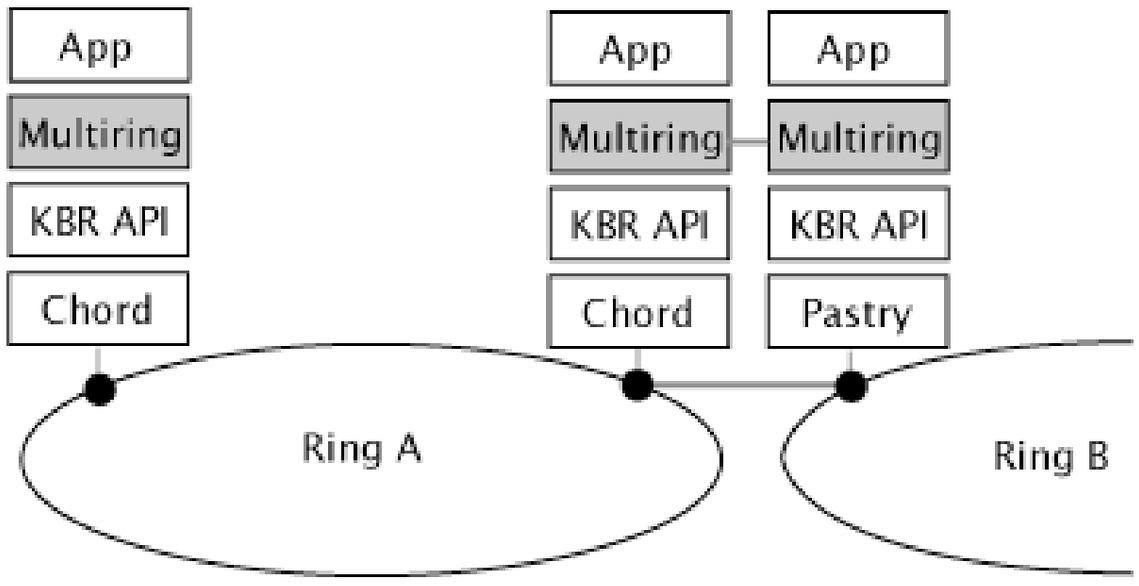
- Provide a layer above existing protocols
  - Organizational autonomy
    - Organizational choice over protocol
    - Choice of parameters (e.g. leafset size, maintenance frequency)
      - Local membership policy
      - Local hardware mix
      - Local churn rate
  - Support for NATs and firewalls

- Thus, delegate authority over resources while providing global overlay connectivity
  - Leverage work on existing overlays (e.g. secure routing)
  - Provide global lookup capability among autonomous organizational rings

6

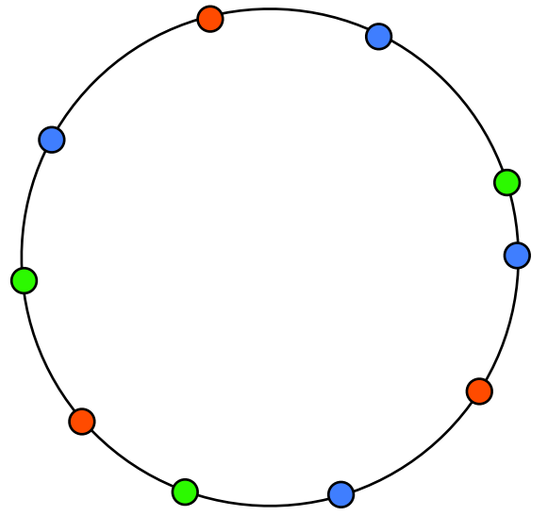
Overview

- Provide a transparent layer above existing structured overlay protocols
  - Support any overlay which is compatible with the KBR API *(IPTPS'03)*
  - Interface into our layer will also be the KBR API
  - Use anycast communication (Scribe) based on the KBR API
  - Can stitch together rings with different protocols



Multiple Rings

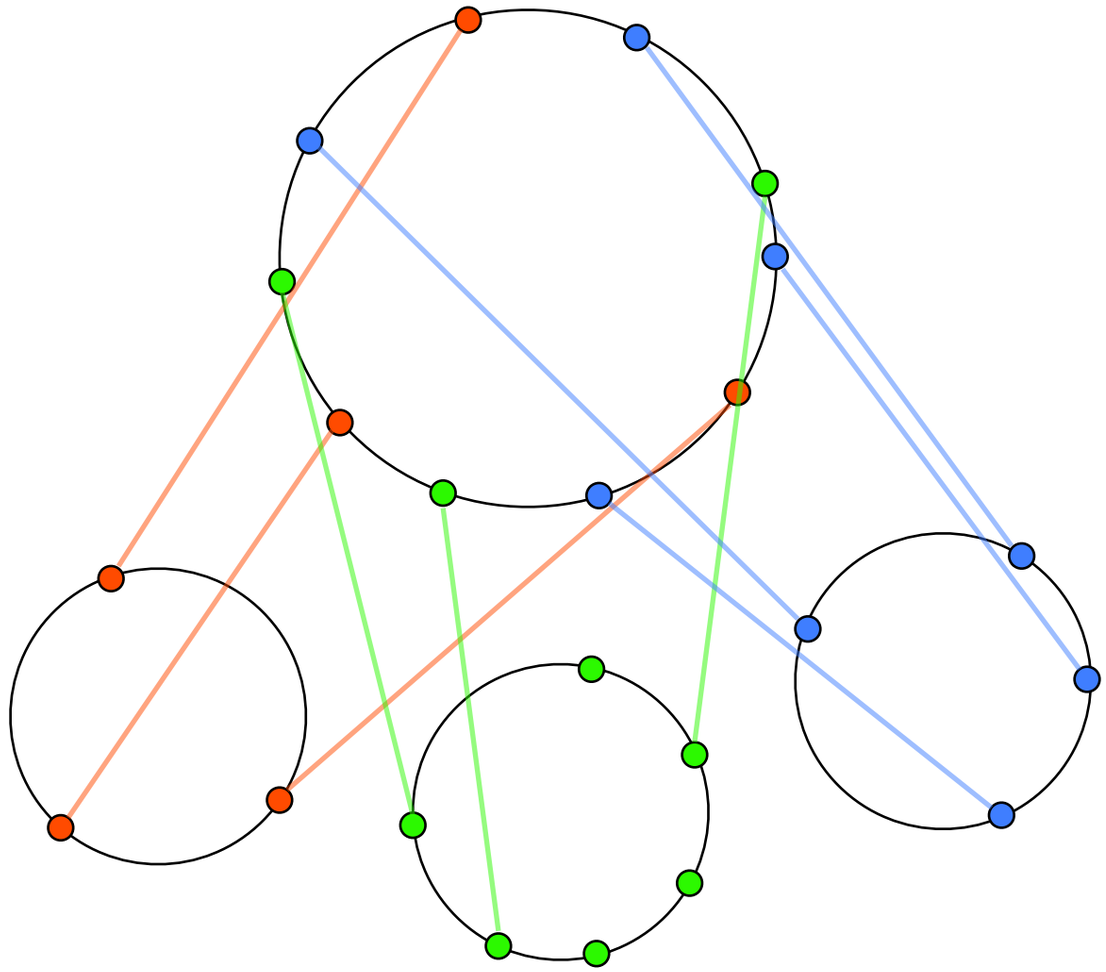
- Move existing ring to a tree of rings
  - Each organization or locality has its own ring
  - Nodes join multiple rings as separate overlay nodes
- Ring boundaries aligned with domains and firewalls/NATs
- Organizations can specify policies for their local ring
  - Insertion into a DHT
  - Subscription to a multicast group
- Global ring enables global key lookup



8

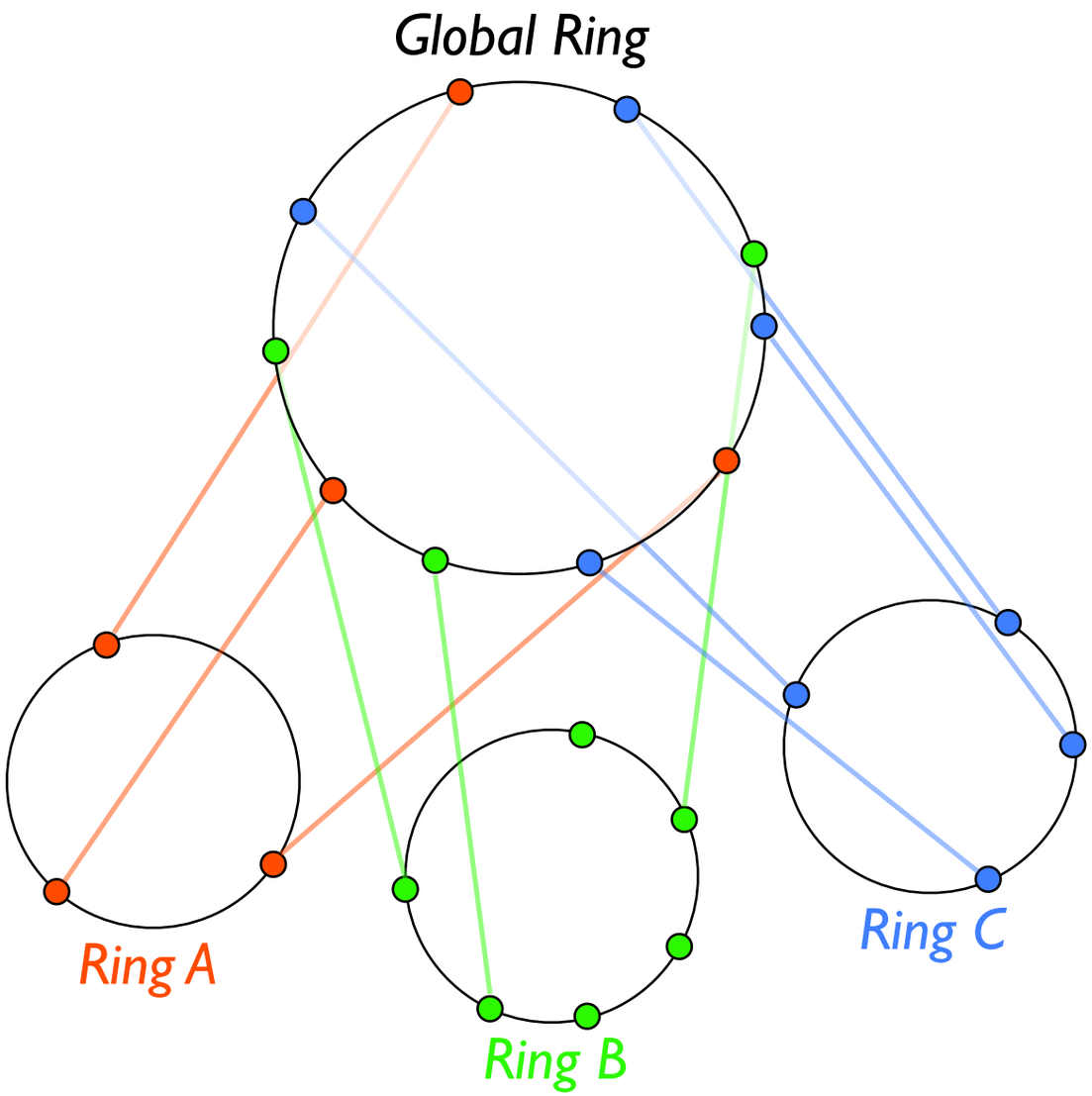
Multiple Rings

- Move existing ring to a tree of rings
  - Each organization or locality has its own ring
  - Nodes join multiple rings as separate overlay nodes
- Ring boundaries aligned with domains and firewalls/NATs
- Organizations can specify policies for their local ring
  - Insertion into a DHT
  - Subscription to a multicast group
- Global ring enables global key lookup



8

RingIds

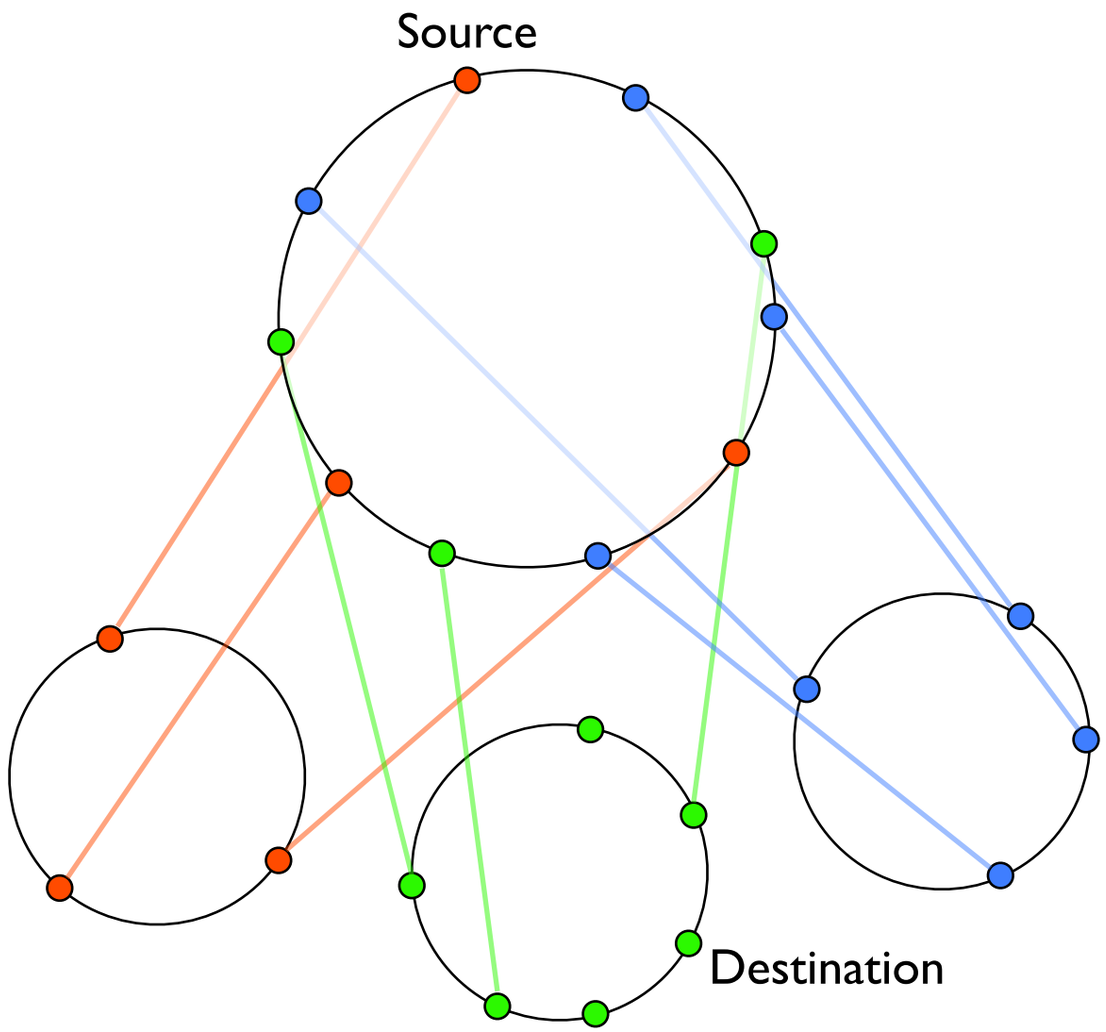


- Each ring is given a globally unique ringId
  - Root, or global, ring has the null ringId
- RingIds are included in a node's certificate
- Keys for routing are now tuples *(ringId, id)*

9

Routing

- Delivering a message to another rings involves finding a gateway node
  - Nodes advertise ring memberships by joining anycast groups
- If a node is a member of ring *A* as well as the global ring, it joins
  - Group *A00...0* in the global ring
  - Group *000...0* in ring *A*
- Other nodes can then anycast to these groups to find gateway nodes
  - Locates a close gateway node in the physical network



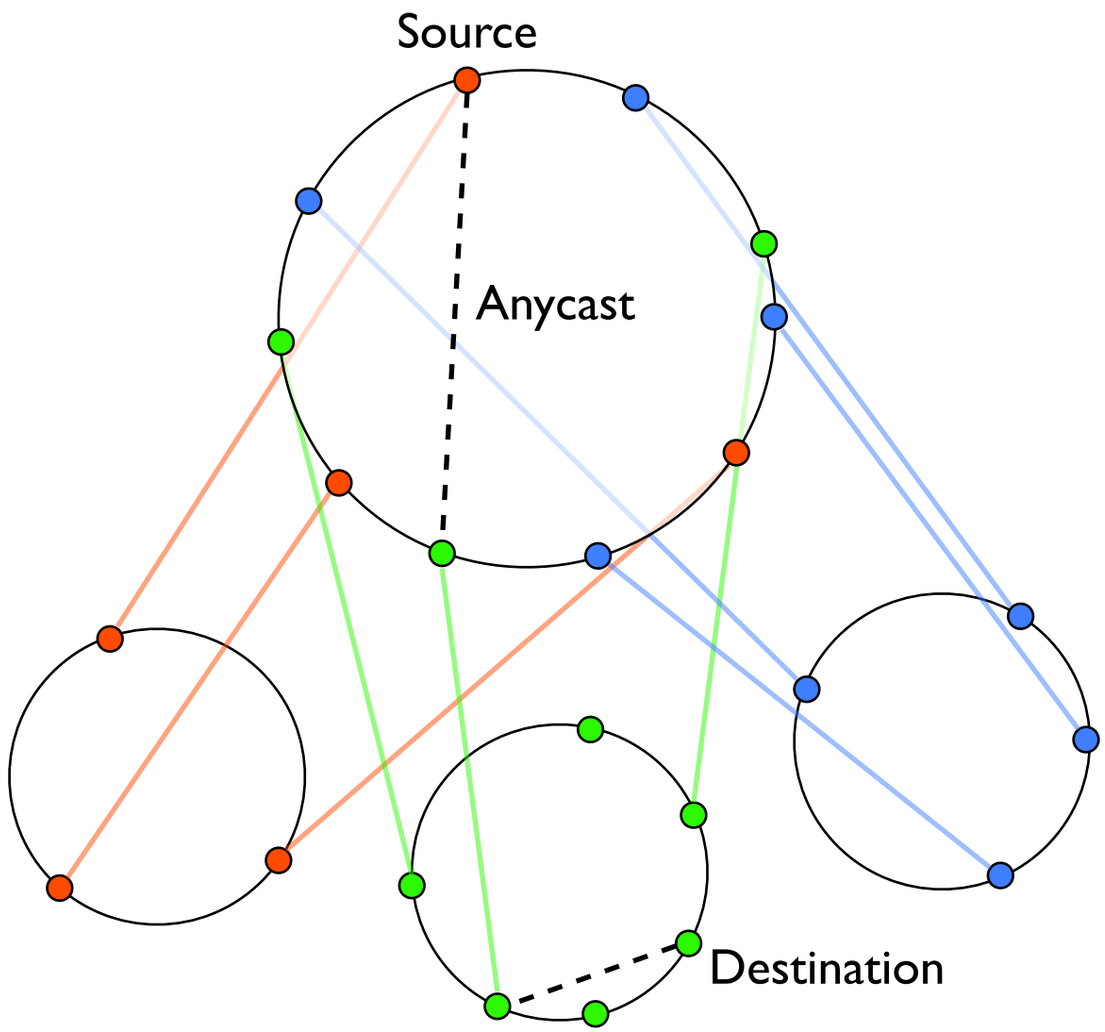
10

Routing

- Delivering a message to another rings involves finding a gateway node
  - Nodes advertise ring memberships by joining anycast groups

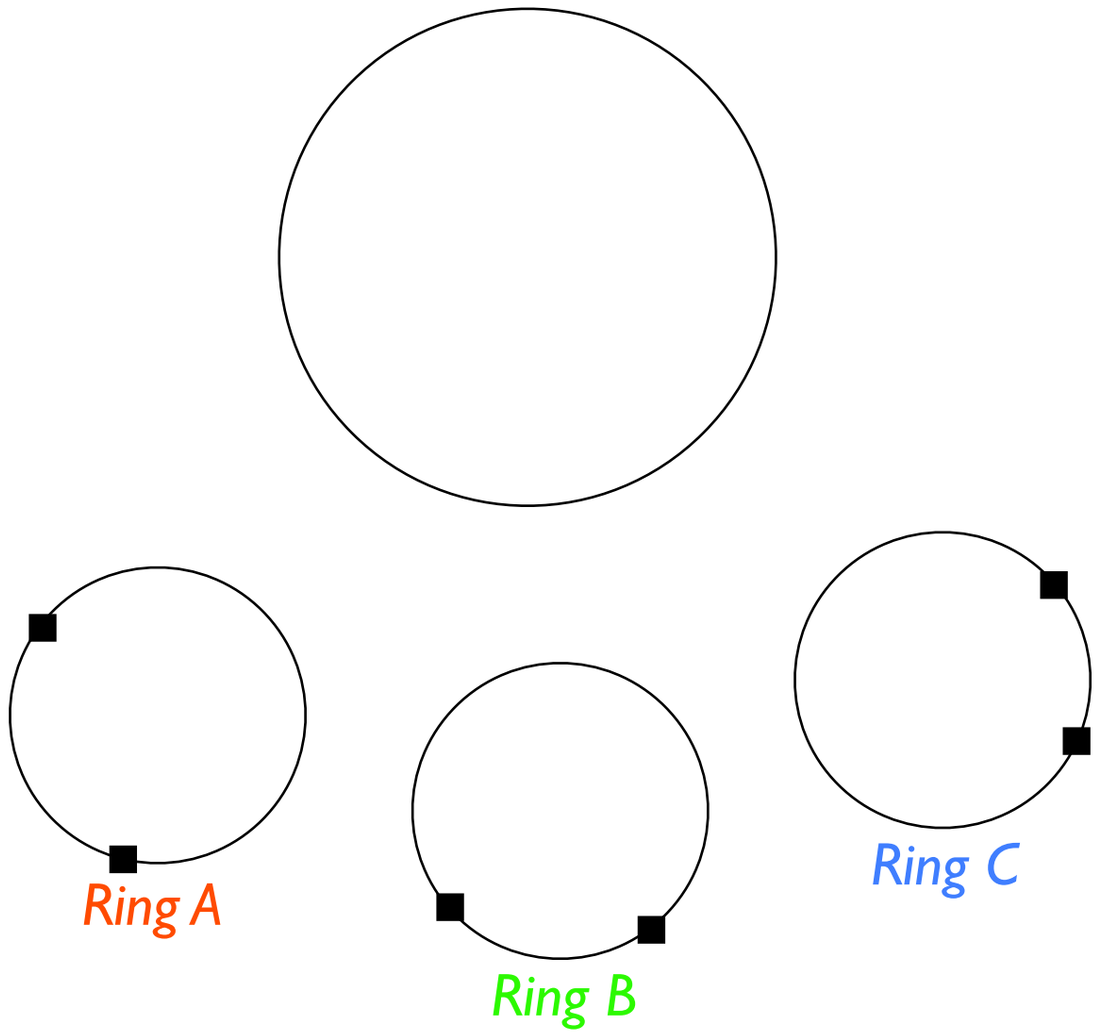
- If a node is a member of ring *A* as well as the global ring, it joins
  - Group *A00...0* in the global ring
  - Group *000...0* in ring *A*

- Other nodes can then anycast to these groups to find gateway nodes
  - Locates a close gateway node in the physical network



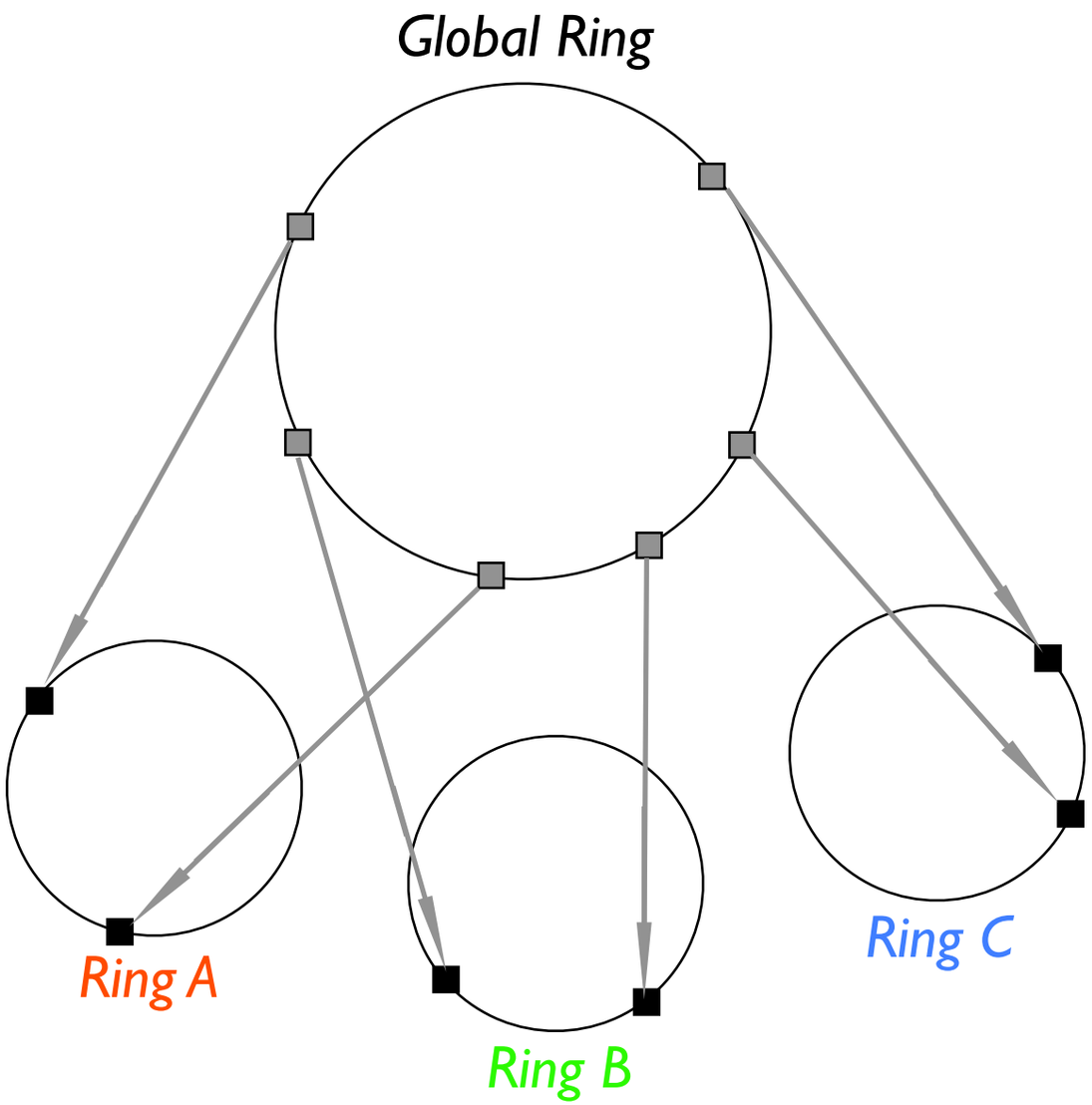
10

Indirection Service



- Still provide for global lookup by key only
- To aid these, an indirection service is run in the global ring
  - Contains pointers with the ringIds of objects
- When inserting an object which should have global scope
  - Pointer is inserted into the indirection service

Indirection Service



- Still provide for global lookup by key only
- To aid these, an indirection service is run in the global ring
  - Contains pointers with the ringIds of objects
  - When inserting an object which should have global scope
    - Pointer is inserted into the indirection service

11

Overhead

- The overhead is comprised of routing overhead and maintenance overhead

- Routing overhead is proportional to the number of hops
  - If no NATs or firewalls
    - Overhead is one extra anycast and one extra overlay route
    - Anycast caching can reduce this to one extra overlay route
  - Otherwise, overhead can be reduced to an extra overlay route per ring layer

- Maintenance overhead is due to multiple rings
  - Organizational ring maintenance is completely internal
  - Recent work has reduced maintenance to < 1 message/second/node
  - Overhead from multicast group maintenance is small

12

Deployment

- Deciding on ring structure is a balance between fault tolerance and locality/ autonomy

- Each organization ring can control their diversity through
  - Separate Internet connections
  - Independent power sources
  - Nodes in different buildings or cities

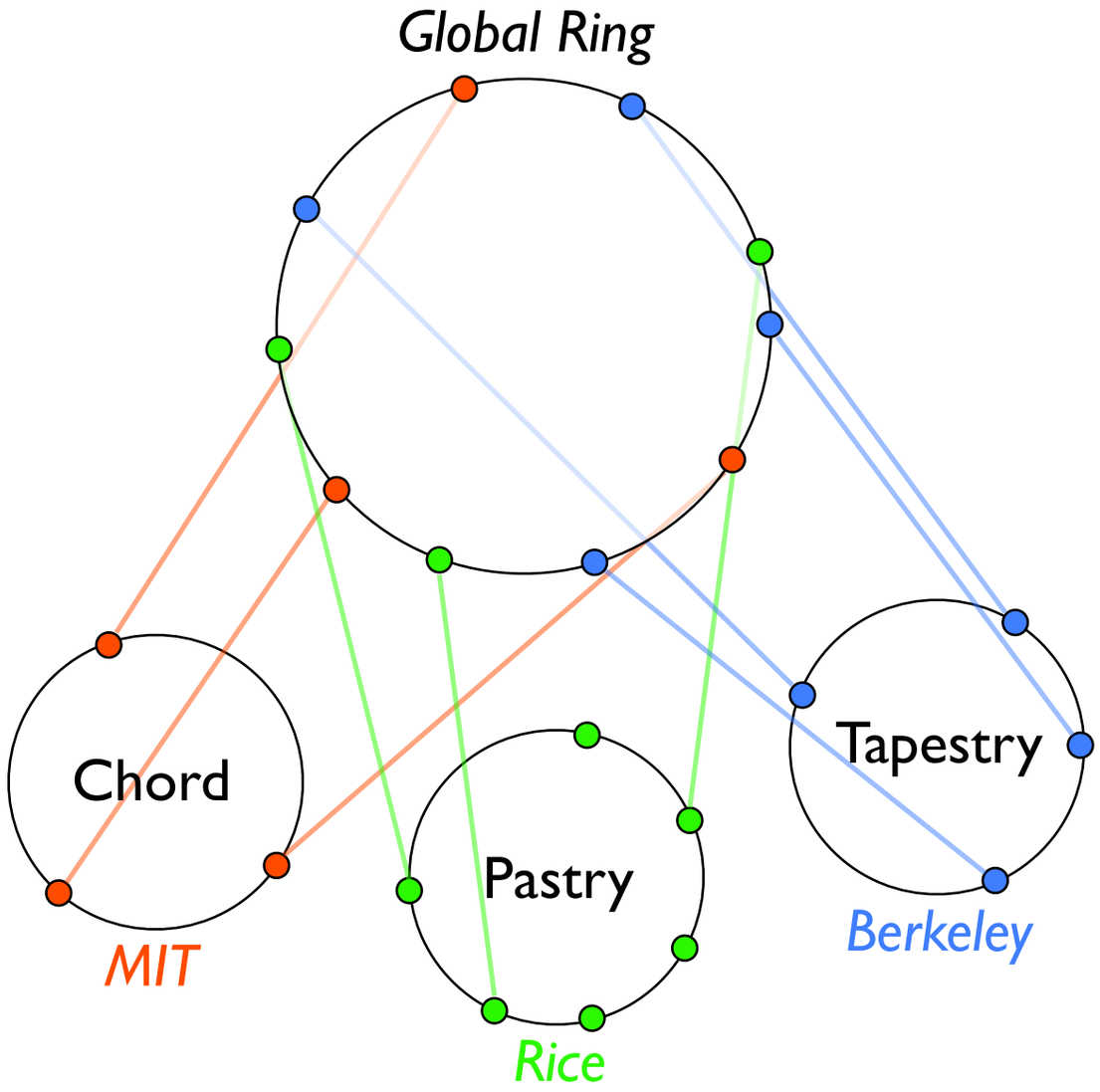
- All nodes which can should join the global ring
  - Provides robust global ring and gateways
  - Imposed extra ring routing only when required by underlying physical network

- Multiple levels of hierarchy can be supported
  - Details are in the paper

13

Example Application: POST

- POST is a serverless, decentralized platform for collaborative applications
  - ePOST is an email service on POST
  - Email delivery only a small notification, data fetched later
- Current uses multiple rings to scope data insertion
  - Data only inserted into local ring is a local user wants it
- Benefits
  - Spam prevention
  - No space-filling attacks



14

Conclusion

- We have provided a layer on top of current structured overlays
  - Provides content and path locality guarantees
  - Gives organizations autonomy over their local ring
  - Allows overlays to work with firewalls and NATs
  - Able to leverage existing structured overlay work (e.g. secure routing)

- Thus, organizations can have autonomous rings stitched together via the global ring
  - Organization rings can run different KBR API protocols
  - Use different protocol and replication parameters

- We have an implementation on the KBR API
  - Will be released in FreePastry 1.4
  - Provides compatibility for applications unaware of the hierarchy

15

Questions?

16