Router assisted overlay multicast

Konstantinos Katsaros, Nikolaos Bartsotas, George Xylomenos

Mobile Multimedia Laboratory
Department of Computer Science
Athens University of Economics and Business
Greece

Outline

- Motivation
- Overlay Multicast with Scribe
- Router assisted overlay multicast
- Simulation environment
  - Topologies
  - Scenarios
- Results
  - Simulation scalability
  - Overlay multicast performance
- Dynamic application behavior
- Conclusions

Motivation

- Internet becoming *information centric*
  - Major part of Internet traffic content distribution
- Peer-to-Peer (P2P) distribution beneficial
  - Content providers: reduced resources
  - Users: reduced download times (e.g. BitTorrent)
- But, inefficient in terms of network resources
  - Redundant unicast transmissions not avoided
- Multicast seems the appropriate choice
  - IP multicast practically not available
- Overlay multicast revisited …
  - … in our case, overlay multicast with Scribe over Pastry (DHT)

Pastry overview

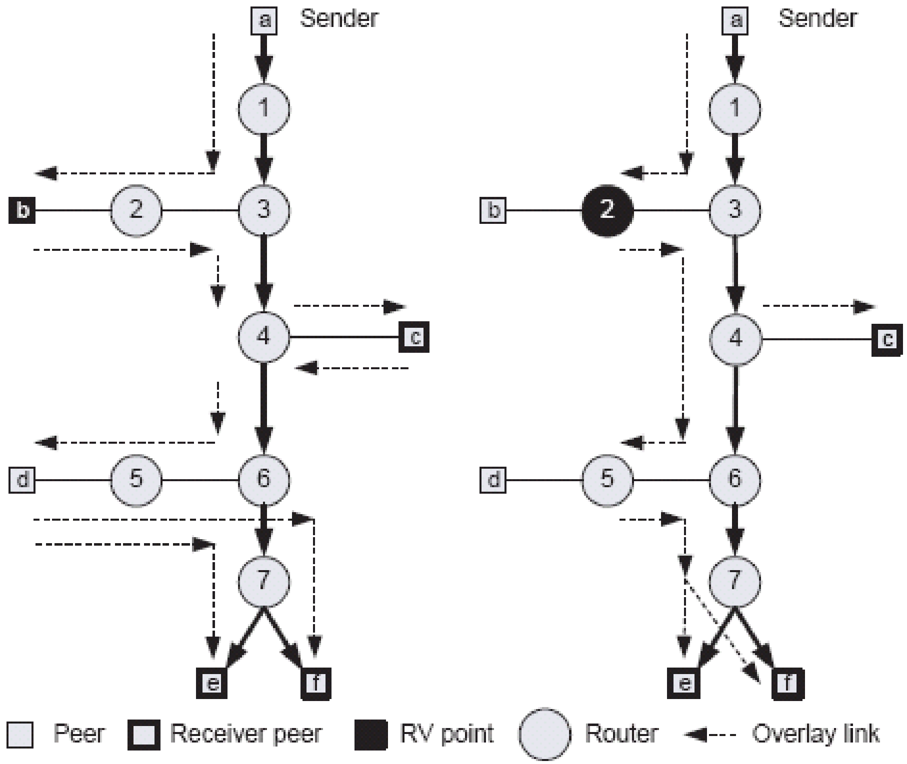
- DHT based, overlay routing infrastructure
- Nodes have unique 128-bit IDs
  - Responsible for the keys (numerically) closer to their ID
- Proximity awareness (e.g. RTT)
- Routing of incoming message with key *K*
  - Send to the closest node sharing one more digit
- Logarithmic number of overlay hops
  - Between any two hosts
- Logarithmic (overlay) routing state size

Scribe overview

- Distributed pub/sub multicast infrastructure
  - Topic-based
    - RV point = Pastry owner of OverlayKey(topic)
- Subscribers send JOIN messages towards a RV point
  - Each node forwards the message in the overlay
    - UNLESS it has been already subscribed to the group
  - Reverse path routing
  - Any node can send data to the RV point
- Scalable
  - No global knowledge of group membership required
    - Each node only aware of its immediate ancestors and descendants
  - No limitations in group identifier space
  - Uniform distribution of forwarding load

Scribe overview (cont.)

- Reliance on end-hosts leads to inefficiencies
  - A **forwarding end-node** may become a bottleneck
    - ADSL uplinks
    - E.g. node *d*
  - Multiple downloads by neighboring
    - E.g. nodes *e* and *f*
  - Increased churn
    - Dynamic participation …

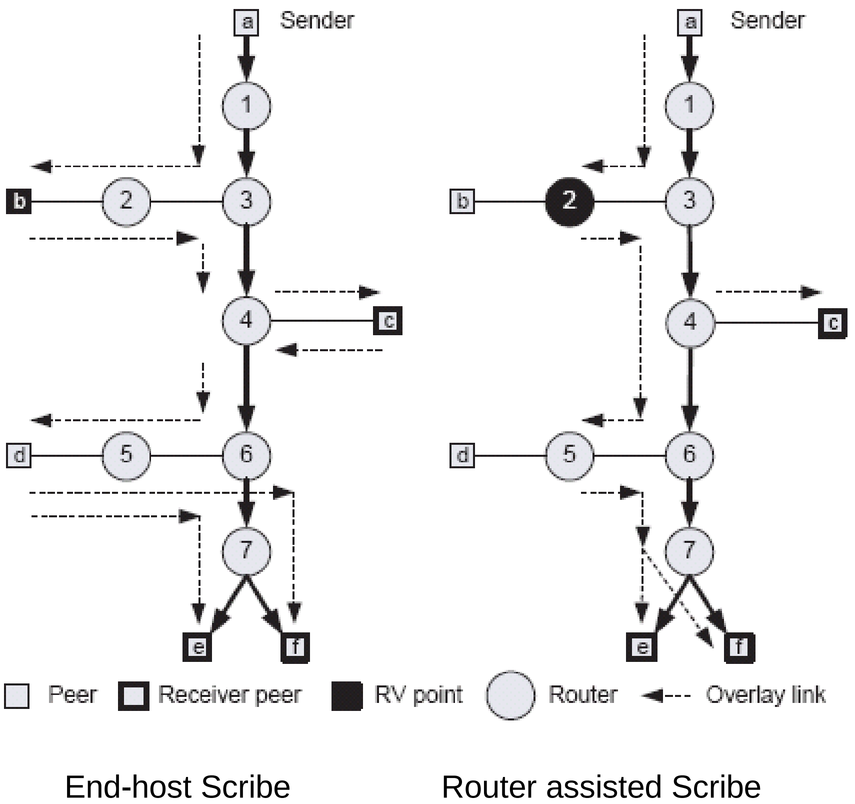


End-host Scribe

Router assisted Scribe

Router assisted overlay multicast

- Access routers acting as proxies:
  - Participate in DHT substrate (Pastry)
  - Overlay multicast functionality (Scribe)
- As long as interested end-hosts exist
  - Only necessary signaling overhead
  - Subject to end-host dynamics…
- One overlay entity per router
  - Single place in DHT
  - Aggregating attached end-hosts demands, i.e. group memberships
- May act as forwarders for end-hosts attached to other access routers



Router assisted overlay multicast (Cont.)

- Why should routers support overlay multicast?
  - Network operator direct gains: reduced traffic load
  - End user gains: lower latencies and higher bandwidths
- But, IP multicast promised the same
- Overlay multicast is different
  - Logarithmic scalability
  - Not *all* routers must participate in the overlay
    - Overlay multicast incrementally deployable
  - Promising for mobility …
- But, routers do not support higher layer functionality
- Overlay multicast proxies
  - Co-located with access routers

    e.g. Scattercast, web-proxies, etc.

Simulation environment

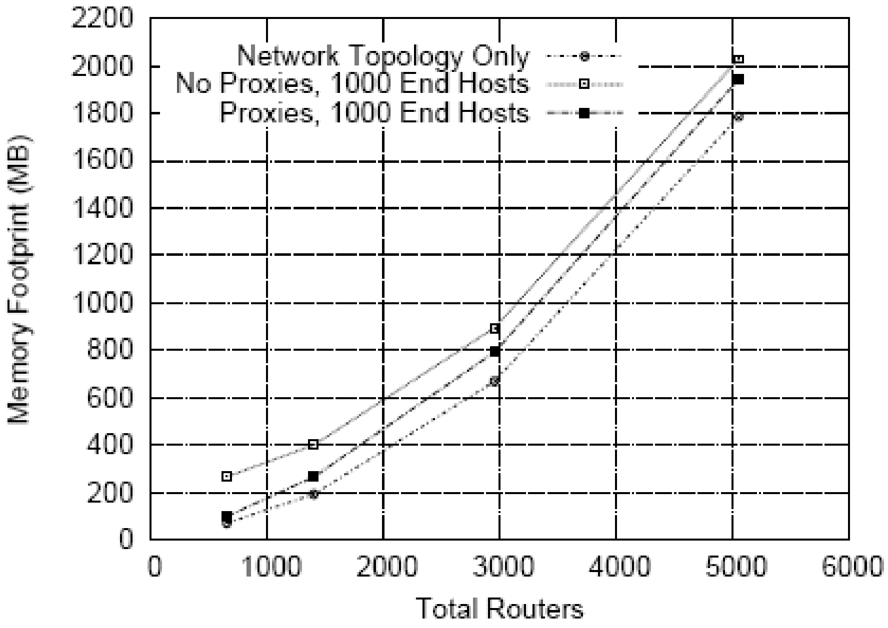
- OMNet++ Simulation environment
  - INET Framework providing TCP/IP protocol stack
  - OverSim providing Pastry/Scribe
- Realism
  - Simulating GT-ITM topologies
    - Topology conversion tool for OverSim…

Topology Name	Transit domains	Average nodes per transit domain	Stub domains per transit node	Average nodes per stub domain	Stub Routers	Transit Routers	Total Routers
Topo-0	5	5	5	5	625	25	650
Topo-1	7	4	7	7	1372	28	1400
Topo-2	9	4	9	9	2916	36	2952
Topo-3	10	5	10	10	5000	50	5050

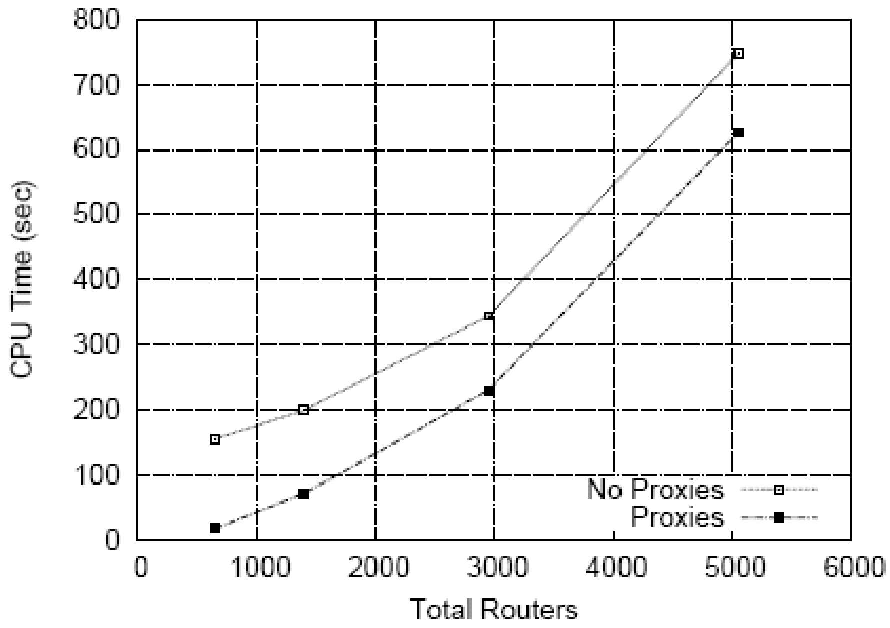
  - Full protocol stack in operation
- Zipf-like distribution of group sizes
- End-host population
  - 500 (Sparse), 1000 (Medium), 4000 (Dense)

Results: simulation scalability

- Impact of network topology on memory consumption
  - E.g. 5000 access routers & 50 backbone routers: 1800 MB
  - Previously neglected



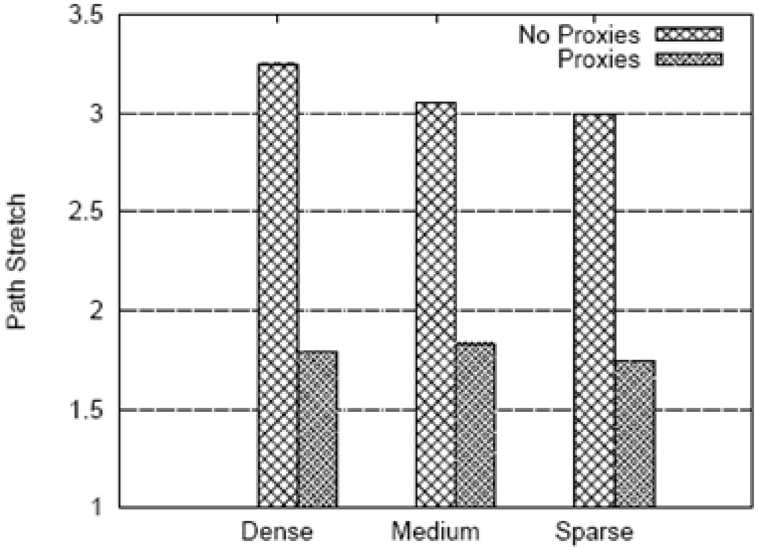
Memory requirements

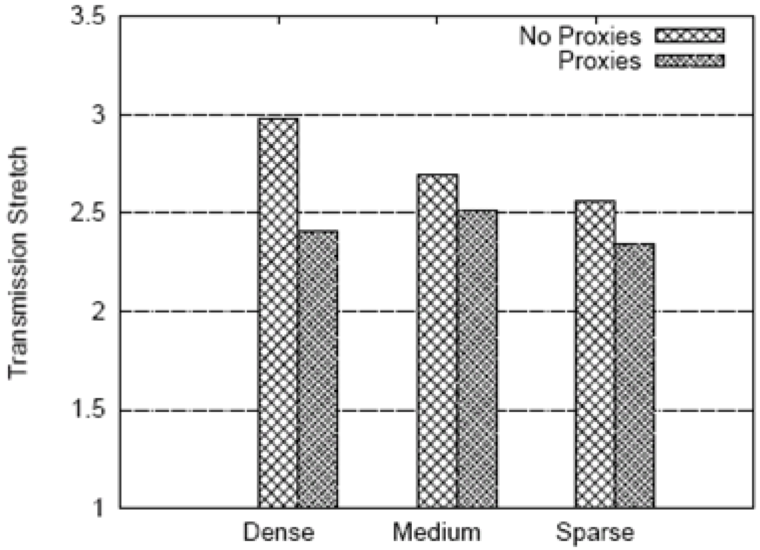


Simulation processing time

Results: overlay multicast performance

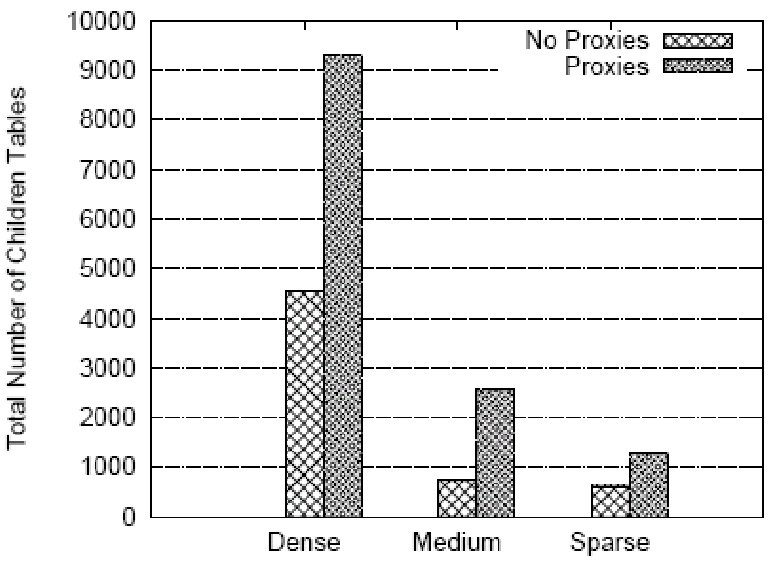
- *Path stretch*
  - # IP hops in the overlay /
    # IP hops of shortest path tree
  - Over all trees and receivers
  - 40-45% decrease
- *Transmission stretch*
  - # hop-by-hop transmissions /
    IP multicast # transmissions
  - Over all trees
  - 7-19% decrease

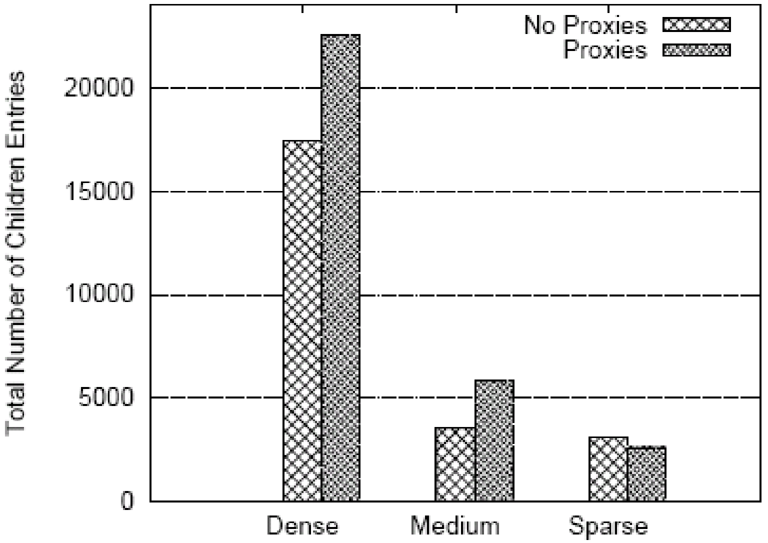




Results: overlay multicast performance

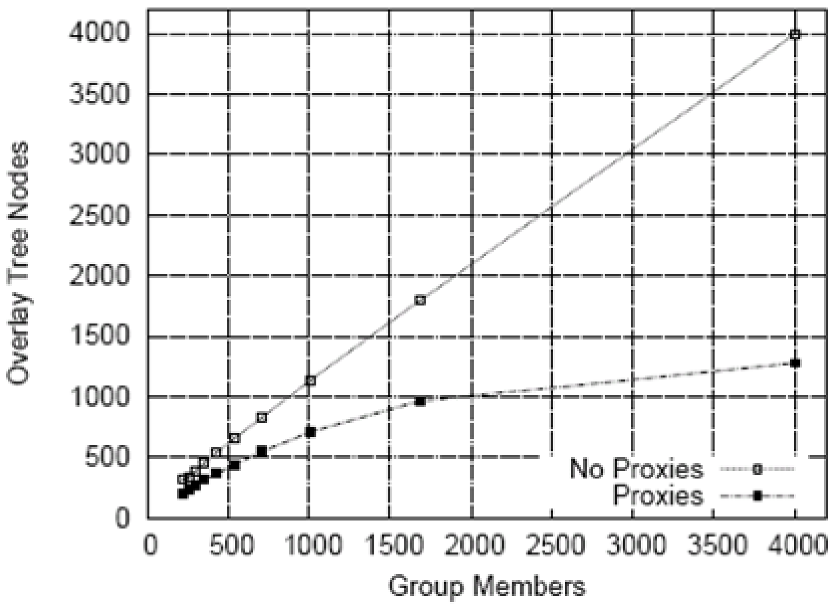
- Node stress
  - Forwarding load
- Number of children tables
  - Number of groups served
- Number of children entries
  - Number of nodes served
- Overhead increase
  - Proxy routers aggregate overhead from multiple end-nodes



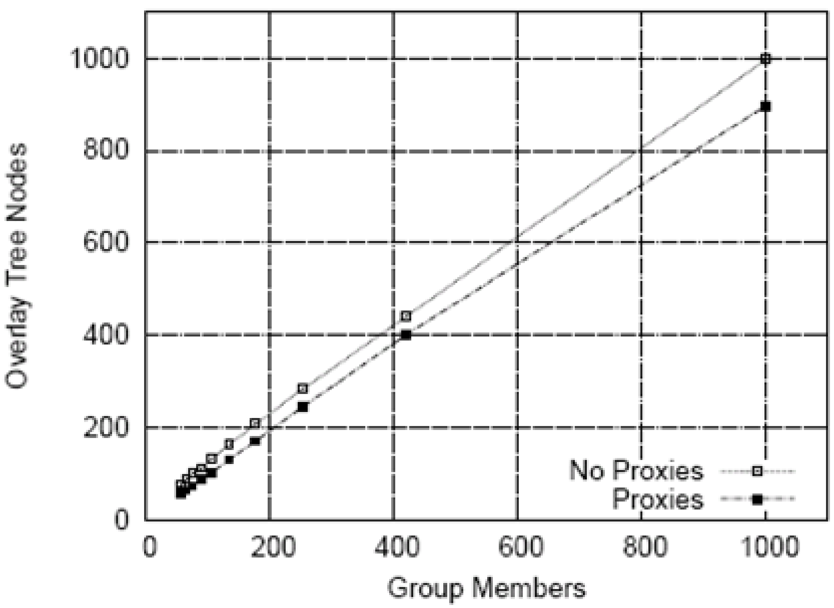


Results: Overlay multicast performance

- Overhead depends on network/end-user density
  - Increased due to the fewer overlay nodes



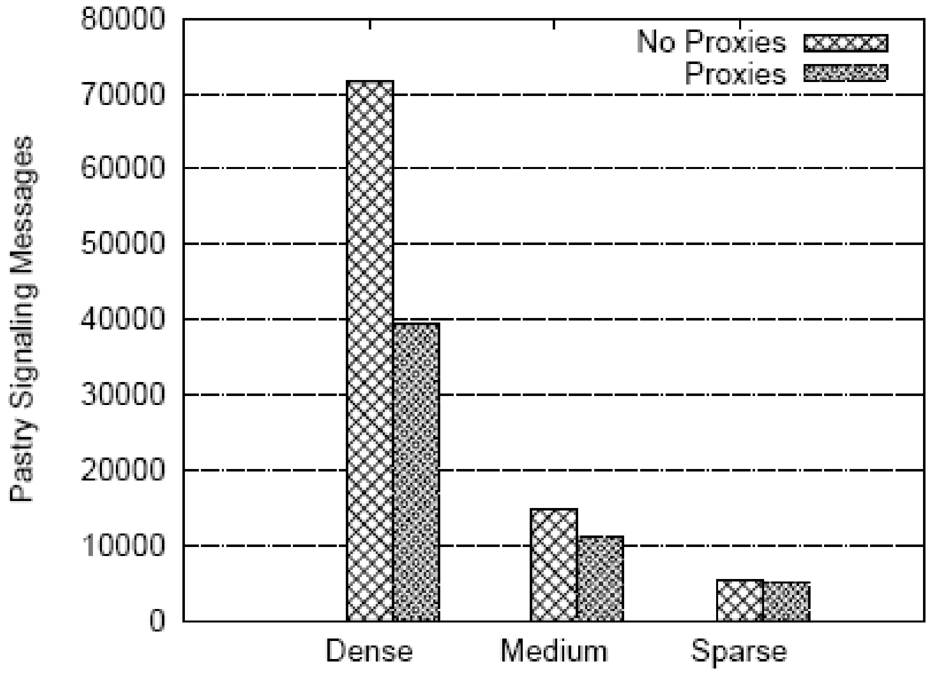
Topo-1 (1400 routers), Dense



Topo-3 (4000 routers), Sparse

Results: Overlay multicast performance

- Reduction in DHT signaling
  - Exactly due to the fewer overlay nodes



Dynamic application behavior

- Data volumes, group sizes
  - Regular Scribe may be sufficient for small volumes to a few end-hosts
- Group membership dynamics
  - Most current studies assume initialized DHT
  - DHT/tree maintenance/adaptation overhead
    - Expecting smaller impact in our scheme
- Global, permanent access router participation
  - Increased stability
  - Increased signaling

Conclusions & Future Work

- Established realistic simulation environment
  - Good for overlay pub/sub & multicast
  - Explored simulation scalability limitations
    - Significant impact of network topology
- Proposed router assisted overlay multicast
  - Improved path stretch and link stress
  - Increased forwarding state/load
  - Reduced overlay overhead
- Incremental deployment of proxies
  - *Intra-domain* level …
  - *Inter-domain* level …
- Dynamic application behavior
  - Investigating content-distribution, caching, …

Thank you!

Konstantinos Katsaros

Mobile Multimedia Laboratory

Department of Informatics/Computer Science
Athens University of Economics and Business

ntinos@aueb.gr
http://mm.aueb.gr/