European Union
European Social Fund
OPERATIONAL PROGRAMME
EDUCATION AND LIFELONG LEARNING
investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
NSRF 2007-2013
programme for development
EUROPEAN SOCIAL FUND
Co- financed by Greece and the European Union

Towards an Error Control Scheme for a Publish/Subscribe Network

Charilaos Stais, Alexios Voulimeneas and G. Xylomenos
xgeorge@aueb.gr
Mobile Multimedia Laboratory
Department of Informatics
Athens University of Economics and Business

IEEE ICC NGN Symposium
Budapest, June 12, 2013

Outline

- Motivation
- Background
- Operation
- Simulator setup
- Performance
- Conclusions

Motivation

- ICN: Information Centric Networking
  - Focuses on information rather than on endpoints
  - Most importantly, it supports native multicast
- Reliable multicast transport over ICN
  - Some multicast applications do not require reliability
    - Example: Multimedia streaming
  - But others really depend on it
    - Example: Software updates, sensor readings
  - Lots of work on reliable IP multicast, e.g. PGM
  - We adapt these ideas to the ICN case
    - Specifically, over the PSI architecture

Background

- Publish Subscribe Internet (PSI) architecture
  - Publishers advertise available data
  - Subscribers express interest in data
  - The Topology Manager creates paths between them
- Stateless forwarding in PSI
  - Paths are encoded as source routes
  - A Bloom filter includes the corresponding links
- Relay points used for forwarding scalability
  - Bloom filters cannot cover large trees
  - So we break them into connected subtrees
  - Relay points used to switch Bloom filters

Operation

- NACK based protocol
  - Only missing packets are NACKed
  - Relay points serve as NACK aggregation points
- Reverse Bloom filters used for the NACKs
  - Simply use the reverse links of the subtree
- Phase 1: Setup
  - Calculation of necessary Bloom filters
    - Both forward and reverse, for each subtree
  - Initial message from publisher
    - Propagates downstream
    - Relay points store Bloom filters

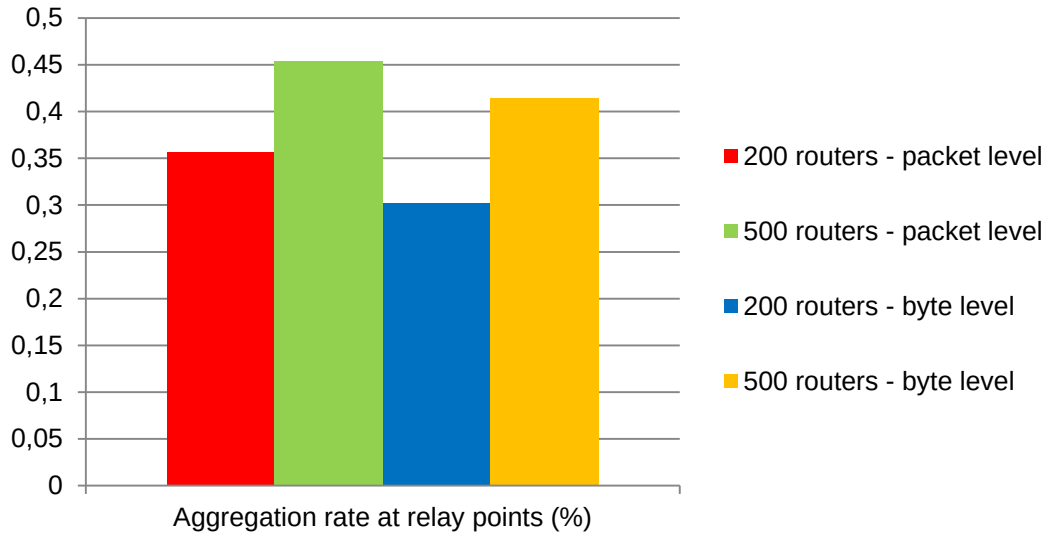
Operation

- Phase 2: Initial content distribution
  - Publisher sends entire content
  - Receivers send NACKs for missing packets
- Phase 3: Recovery
  - Repeats Phases 1 and 2
  - Some subscribers leave at the end of each cycle
  - Eventually no subscribers left
- NACK aggregation
  - Relay points receive NACKs
  - Wait to get more NACKs for a little while
  - Then, received NACKs are merged and propagated

Simulator setup

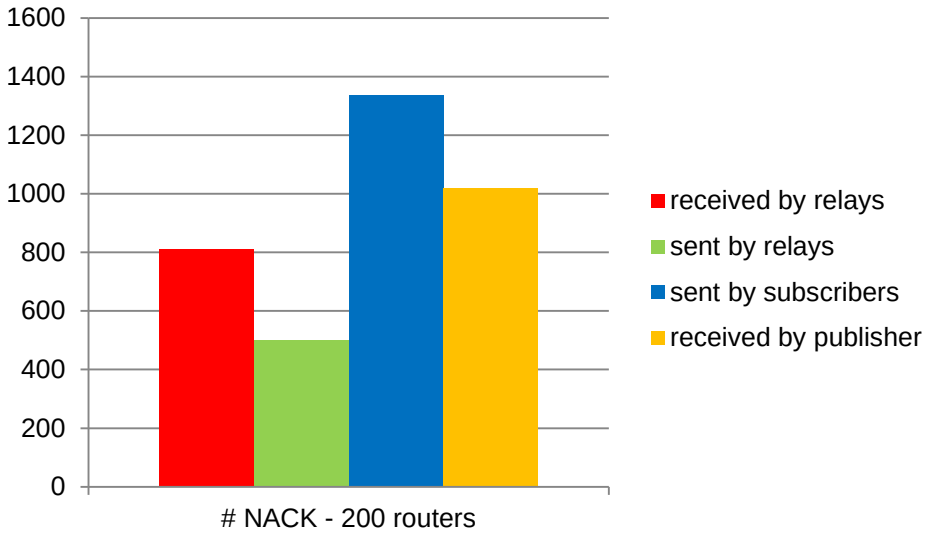
- NS-3 based simulations
  - Entire PSI architecture implemented
  - Relaying and transport protocol added
  - Single publisher transmits 20 MB
  - Scale free topologies with 200 and 500 routers
  - 50 and 100 subscribers randomly attached to routers
  - 3% of the packets are reported lost in first round
    - Random loss model used
    - Roughly 600 packets need to be retransmitted
    - Manual setting of link loss to achieve target rate

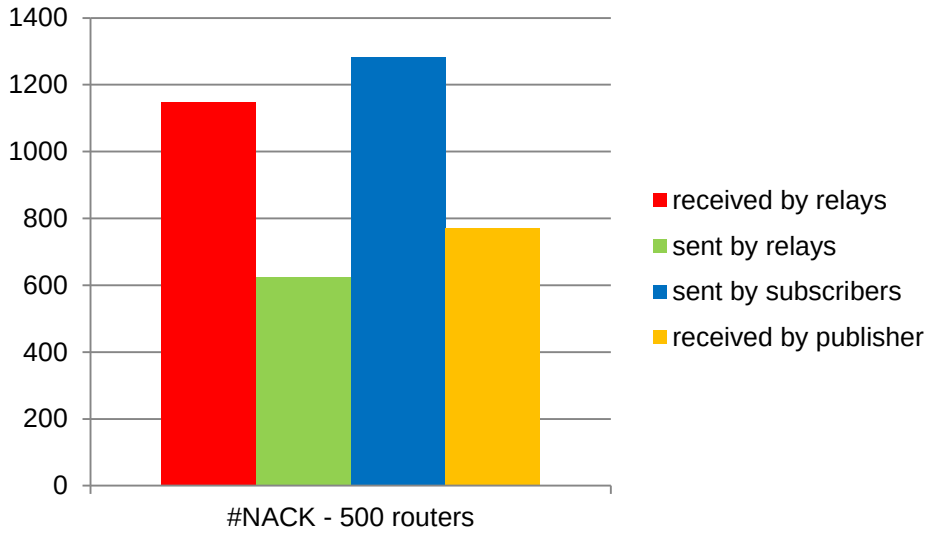
Performance: aggregation rate



- Aggregation rate of NACKs
  - Fraction of NACKs ***not*** reaching the publisher
  - Grows with larger topology
    - More opportunities for aggregation
  - Slightly lower in bytes

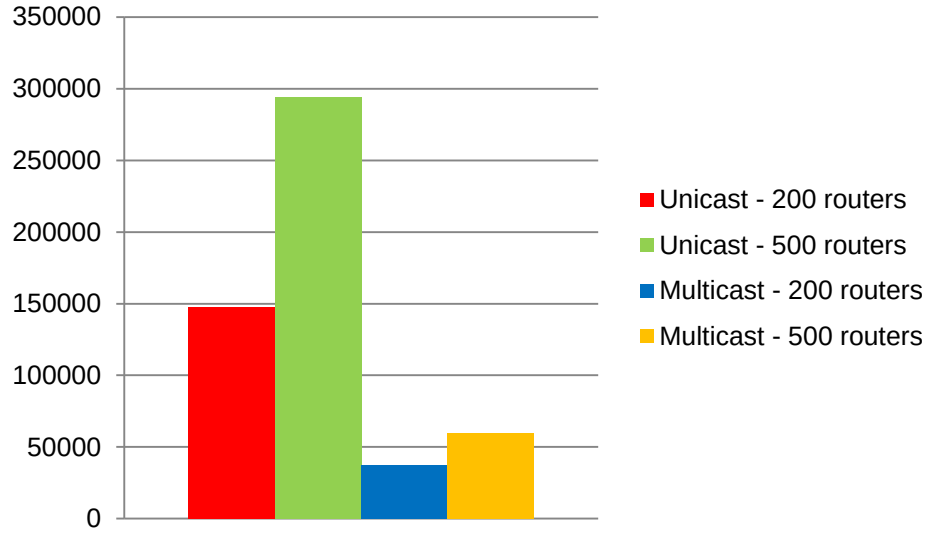
Performance: NACKs handled

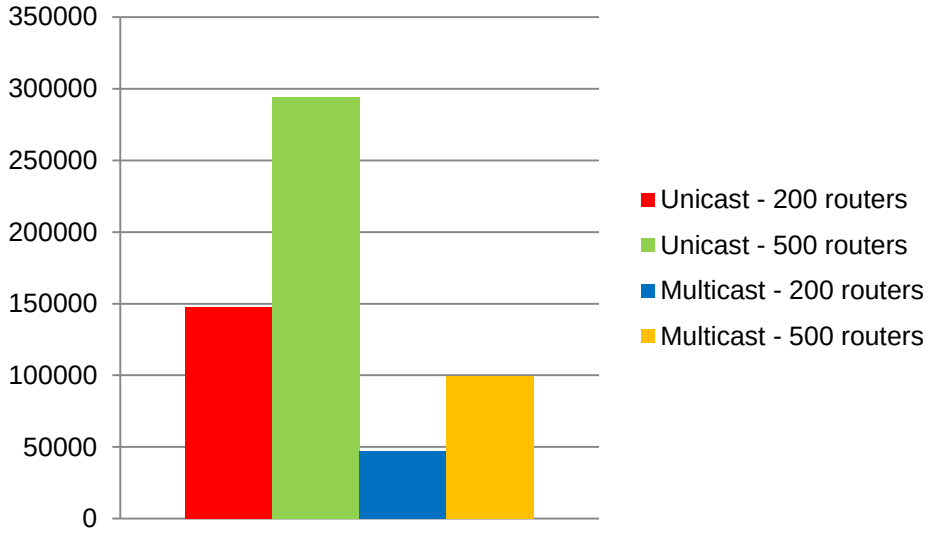




- Number of NACKs handled by each entity
  - Similar number of original NACKs
    - Both topologies start with the same loss rate
  - More NACKs handled by relays with larger topology
  - But, more aggregation overall with larger topology

Performance: Uni vs. Multicast





- Unicast vs. Multicast recovery
  - NACKs and retransmissions
    - Calculated over all network elements
  - Left: Number of NACK transmissions
  - Right: Number of packet retransmissions

Conclusions

- Reliable multicast transport in PSI
  - Takes advantage of relay points
  - Very good feedback aggregation
  - Beats unicast by a wide margin
- Ongoing work
  - Detailed comparison with PGM
  - PGM exploits unidirectional multicast
    - Relay point multicasts NACK confirm
    - Good for “correlated” losses
  - We exploit reverse Bloom filters
    - Relay point just aggregates NACKs
    - Good for “uncorrelated” losses