Scalable Routing for Topic-based Publish/Subscribe Systems under Fluctuations

Volker Turau & Gerry Siegemund

37th International Conference on Distributed Computing Systems
June 8th, 2017

Institute of Telematics
Hamburg University of Technology
TUHH

1

Publish/Subscribe Paradigm

Publish/Subscribe Paradigm

- Publish/subscribe: A loosely coupled distributed information dissemination middleware
- Publishers distribute data (a.k.a. publication) to subscribers asynchronously and anonymously
  - Senders are unaware of number and addresses of subscribers
- Subscribers define their interest in topics by a categorization done by publisher
- Well established paradigm in Internet

Publish/Subscribe Paradigm

- Interface

  subscribe(t)
  unsubscribe(t)
  publish(m, t)

- Publish/subscribe middleware takes care of forwarding publications to subscribers
- **Our focus**: Low-power wireless networks with limited resources $\longrightarrow$ IoT
- Challenge: Low memory routing structure

2

Memory Constrained Routing

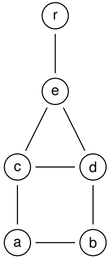
Memory Constrained Routing

- **Fact 1**: To accomplish shortest path routing, the routing table of each node needs to grow as $\Omega(n)$
- **Path stretch** of protocol P: Ratio of path length achieved by P, divided by shortest path length
- **Fact 2**: Any protocol that keeps path stretch in worst case strictly below 3, requires a $\Omega(n)$ bit state at each node [Gavoille et al.]
- Routing over spanning tree:
  Upper bound for average path stretch for spanning trees is in $\Omega(\log n))$ [Alon et al.]
  - Size of routing table $O(\Delta)$

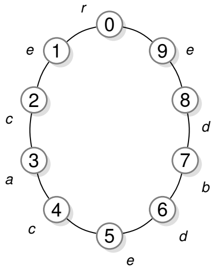
Memory Constrained Routing

- **Virtual Ring**: Directed closed path involving each node of the graph, possibly several times
  - Routing: Publisher forwards message around ring and each subscribing node reads it. Upon return to sender message is discarded
  - Constant space routing tables
  - May incur a linear path stretch
- To lessen stretch additional edges – a.k.a. chords – are used as shortcuts at cost of increased routing tables

Virtual Ring Routing

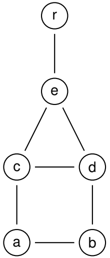


Communication Graph

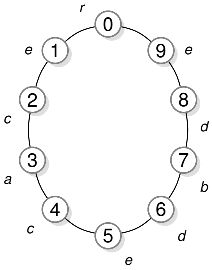


Virtual ring

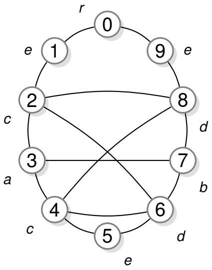
Virtual Ring Routing with Chords



Communication Graph



Virtual ring

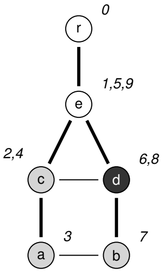


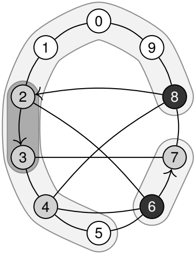
Virtual ring graph

Routing on Virtual Rings with Chords

- Publications are routed around ring as before
- Chords are used as short cuts
  - A node can skip a ring segment if does not contain a subscriber
- Greedy routing on virtual ring graph
  - Each node maintains for each of its positions p and for each topic t the subscriber's position that is counter clockwise closest to p
  - This is called the **forwarding position**
- Concurrent forwarding into disjoint segments

Example



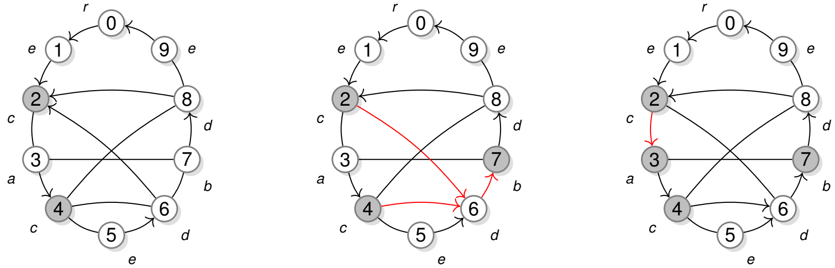


- Subscribers at nodes a,b,c (pos. 2,3,4, and 7)
- Publisher at node d (pos. 6,8)

position	0	1	2	3	4	5	6	7	8	9
forwarding position	2	2	3	4	7	7	7	2	2	2
closest position	1	2	3	4	6	6	7	8	2	0
next neighbor	e	c	a	c	d	d	b	d	c	r

Subscribing

Subscribing is a local operation: Information is not forwarded beyond first subscriber in each subtrees



c is only subscriber *b* subscribes *a* subscribes

Unsubscribing

- Strictly speaking no maintenance operations are required, but then routes become longer
- A forwarding position at v that corresponds to position of an unsubscribing node w is replaced by forwarding position at w
- Straightforward implementation can lead to race condition
- Solution: Adaptation of Raymond's mutual exclusion algorithm

Routing on Virtual Rings with Chords

Theorem

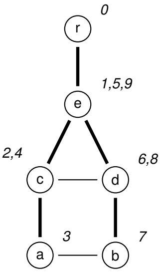
In a fault-free network our algorithm satisfies the following properties.

1. *The algorithm is free of race conditions, deadlocks, and livelocks.*
2. *A new subscriber will receive all publications sent at least d rounds later by nodes that have distance d to v in the order of sending.*
3. *Forwarding of publications for topic t is suspended at most D (diameter of G) rounds after last subscriber unsubscribed from t.*

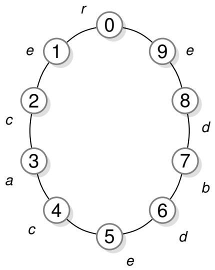
3

Construction of Virtual Rings

DFS-Construction of a Virtual Ring



Spanning tree

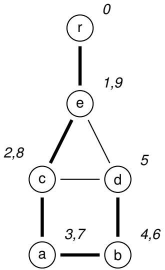


Virtual ring

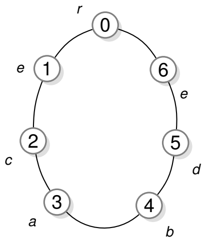
- Traverse any tree and assign positions to nodes
- Can be integrated in DFS
  - $O(n)$ rounds
  - Length of ring: $2(n-1)$ (independent of tree)

Shorter Virtual Rings

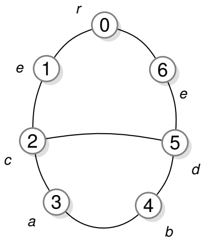
- Apply aggressive backtracking
- Instead of backtracking via dfs-tree use back edges and skip nodes



Spanning tree



Virtual ring



Virtual ring graph

Shorter Virtual Rings

- Algorithm is a modification of Awerbuch's dfs algorithm
- $O(n)$ rounds, $O(m)$ messages, message size $O(\log n)$

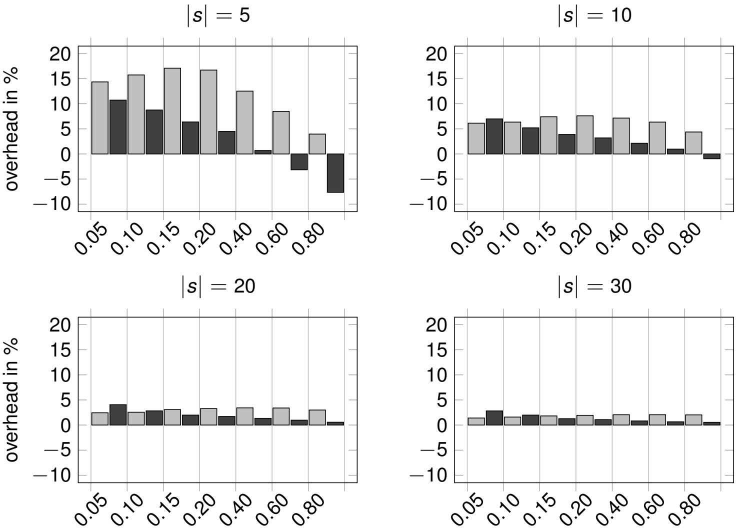
4

Evaluation

Evaluation

- Our algorithm is a compromise between ease of maintenance of routing structure and lengths of forwarding paths
- Comparison with
  - $\mathcal{MT}$: A bfs routing tree rooted at each node, recursively pruned leaves not corresponding to subscribers
  - $\mathcal{ST}$: A single bfs routing tree rooted at a central node
- Note: In both cases, changes of subscriptions require a complete recomputation of trees

Evaluation: $\mathcal{MT}$ (gray) and $\mathcal{ST}$ (black)



5

Conclusion

Conclusion

- Distributed routing algorithm for pub/sub systems in resource-constrained networks
- Compromise between efficient maintenance of routing structure and lengths of forwarding paths
- Sub- and unsubscriptions require message exchange in a local region only
- Implementation tested in real network in Fit-IoT Lab in France

Scalable Routing for Topic-based Publish/Subscribe Systems under Fluctuations

Volker T

37th International Conf

Volker Turau

Professor

Phone +49 / (0)40 428 78 3530

e-Mail turau@tuhh.de

http://www.ti5.tu-harburg.de/staff/turau