A $O(\log n)$ Distributed Algorithm to Construct Routing Structures for Pub/Sub Systems

Volker Turau

20th Int. Symposium on Stabilization, Safety, and Security of Distributed Systems
November 6th, 2018

Institute of Telematics
Hamburg University of Technology
TUHH

1

Publish/Subscribe

Publish/Subscribe Paradigm

- Loosely coupled distributed information dissemination middleware
- Subscribers declare interest in topics by subscribing to topic
- Highly scalable
  - Data is distributed asynchronously and anonymously
  - Senders are unaware of number and addresses of subscribers

Publish/Subscribe Paradigm

- Interface

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

- Publish/subscribe middleware forwards publications to subscribers
- **Focus of our work**: Low-power wireless networks with limited resources $\longrightarrow$ IIoT
- Challenge: Low memory routing structure

Memory Constrained Routing

- **Fact 1**: Shortest path routing requires routing tables of size $\Omega(n)$
- **Path stretch** of protocol *P*: Ratio of path length achieved by *P*, divided by shortest path length
- **Fact 2**: Protocols with path stretch below 3 require $\Omega(n)$ size routing tables [Gavoille et al.]

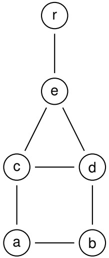
- Goal: Routing tables of poly-logarithmic size, e.g., $O(\log^2 n)$

Memory Constrained Routing

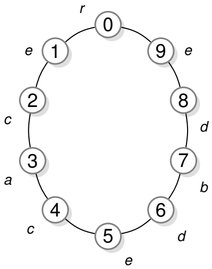
- **Virtual Ring**: Closed path involving each node at least once
  - Routing
    - Publisher forwards message around ring and subscribers read it
    - Upon return to sender message is discarded
  - Routing table of each node only contains id of next node
  - May incur a linear path stretch, i.e. length of cycle

- To lessen stretch additional edges – a.k.a. fibers – are used as shortcuts at cost of increased routing tables

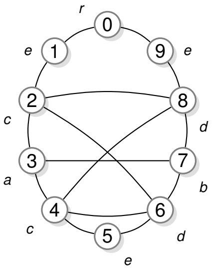
Virtual Ring Routing with Fibers



Communication Graph



Virtual Ring



Virtual Ring with Fibers

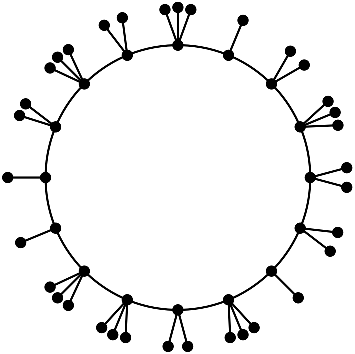
- Fibers are used as short cuts
  - A node skips a ring segment if does not contain a subscriber
  - Concurrent forwarding into disjoint segments

- PSVR is a pub/sub middleware that uses virtual rings

New Routing Structure

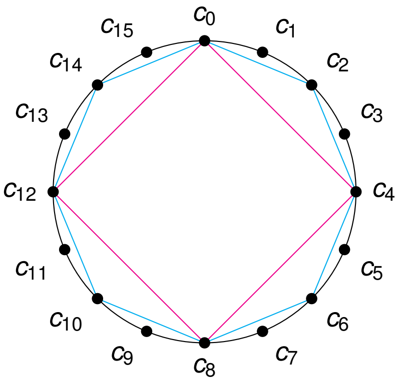
- Observations
  - Shorter rings reduce path stretch
  - More fibers reduce path stretch but may enlarge routing table
- Goal
  - Ring structure with smaller stretch and smaller routing tables such that PSVR can be executed without changes
- New concepts
  - Not all nodes need to be on ring $\longrightarrow$ shorter rings
    - Nodes on ring act as proxies for nodes outside ring (NAT)
  - Systematic approach to select fibers
    - Compromise between path stretch and routing table size

New Routing Structure



Nodes outside ring are homogeneously attached to proxies

Optimal Fiber Structure



Homogeneous structure of fibers inside smaller ring

Contribution

- Efficient distributed algorithm to construct the new routing structure
- Time complexity $O(\log n)$ rounds
- Analysis of algorithm for random graphs
  - Path stretch is w.h.p. in $O(\log n)$
  - Nodes outside ring attach w.h.p. homogeneously to proxies

Assumptions

- Synchronous $\mathcal{CONGEST}$ model of the *distributed message passing model*
  - each message contains at most $O(\log n)$ bits
- Unique identifiers
- Nodes have only limited local memory
- Each node knows size n of network
- A dedicated starting node v_0

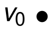
2

Informal Description of Algorithm $\mathcal{A}_{Fiber}$

Algorithm $\mathcal{A}_{Fiber}$

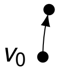
- $\mathcal{A}_{Fiber}$ works in phases
  - Phase 0 ($O(\log n)$ rounds)
    - Starting in v_0 a path P of length $2^r - 1$ is built ($2^r \in O(\log n)$)
  - Phase 1 ($O(\log n)$ rounds)
    - Path P is closed to a ring C of length 2^r
  - Middle phases
    - Concurrently edges of C are replaced by two edges
    - Replaced edges become fibers
    - After k middle phases, C has up to $2^k \log n$ nodes
  - Final phase
    - Each node in $v \in V \setminus C$ randomly selects a neighbor u on C
    - Node u is the proxy for v
- Each middle phase and the final phase lasts $O(1)$ rounds

Phases



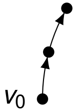
Phase 0

Phases



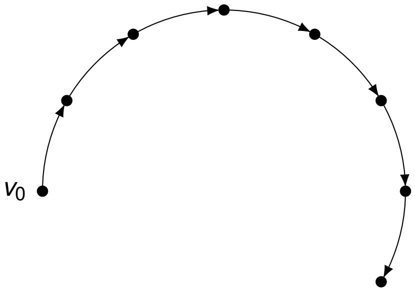
Phase 0

Phases



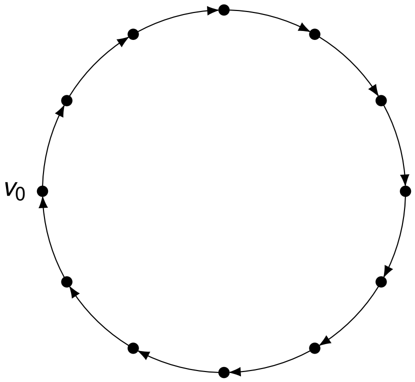
Phase 0

Phases



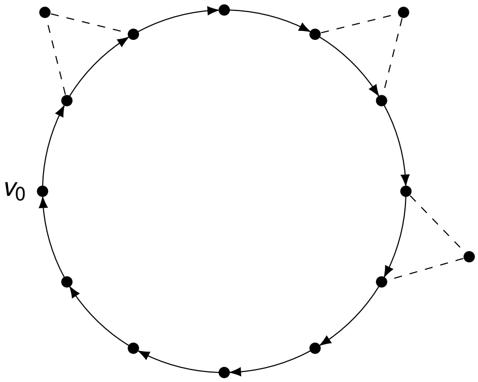
After $O(\log n)$ rounds

Phases



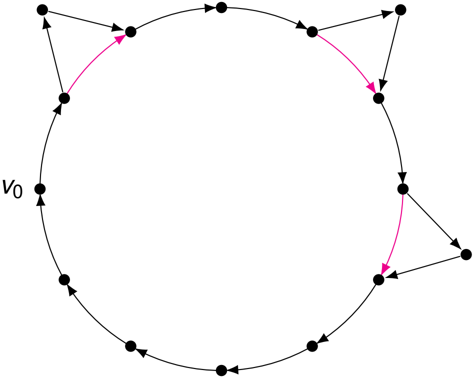
Phase 1, after another $O(\log n)$ rounds

Phases



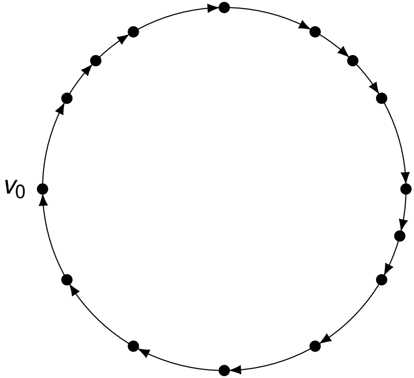
Middle Phase Step 1

Phases



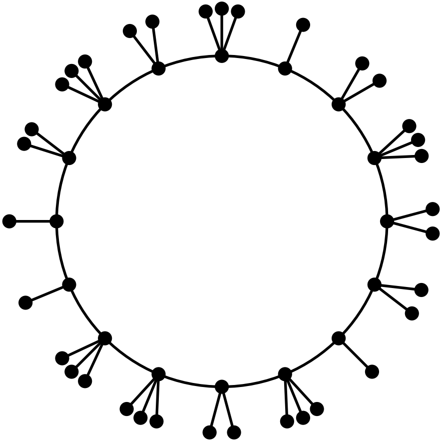
Middle Phase Step 2

Phases



After k Middle Phases: C has at most $2^k \log n$ nodes

Phases

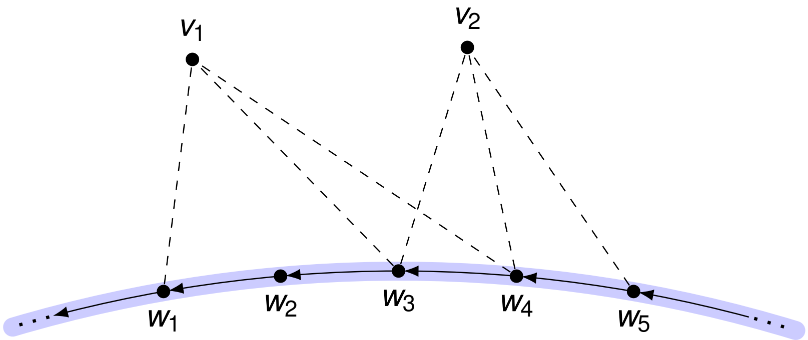


Final Phase

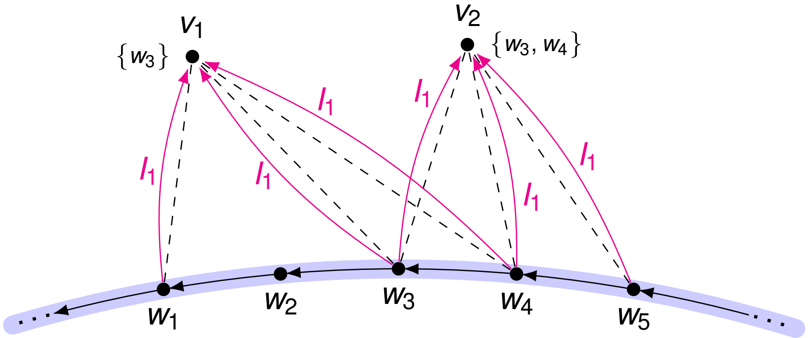
3

Details of $\mathcal{A}_{Fiber}$

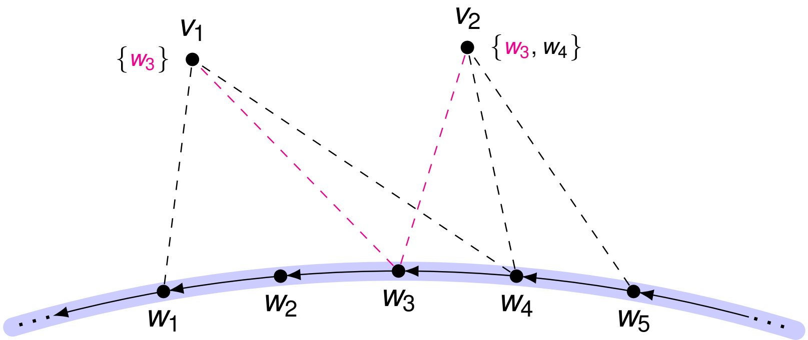
Middle Phases



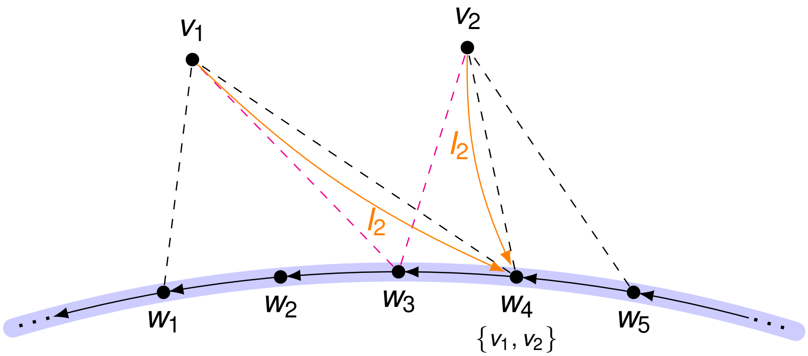
Middle Phases



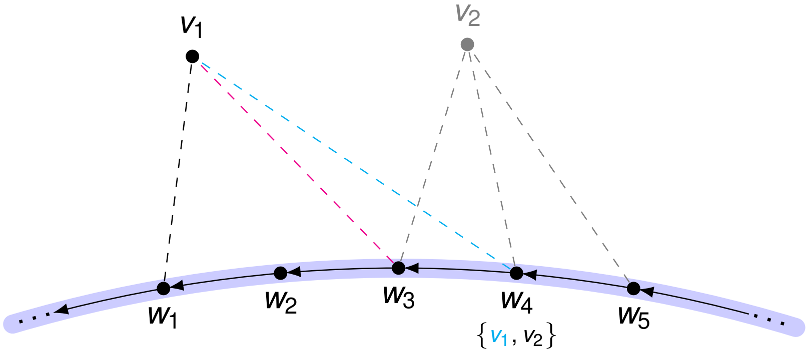
Middle Phases



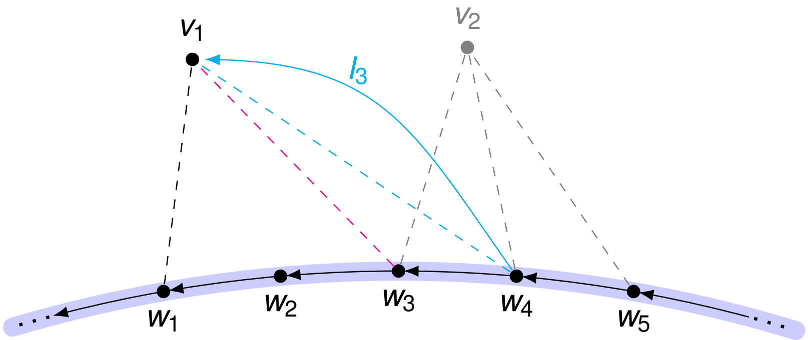
Middle Phases



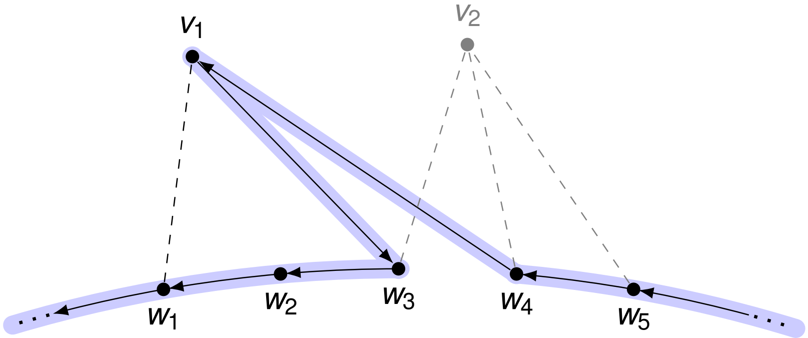
Middle Phases



Middle Phases



Middle Phases



Observations

- Individual extensions do not interfere with each other because
  - Each node outside C sends in each middle phase at most one request to integrate and
  - each edge of C accepts at most one integration request

4

Analysis of $\mathcal{A}_{Fiber}$

Analysis of $\mathcal{A}_{Fiber}$ for Random Graphs

- We analyze algorithm $\mathcal{A}_{Fiber}$ for the class of random graphs $G(n, p)$ with $p \in O(\log n / \sqrt{n})$
- Claim: In each middle phase w.h.p. size of C is increased by constant factor

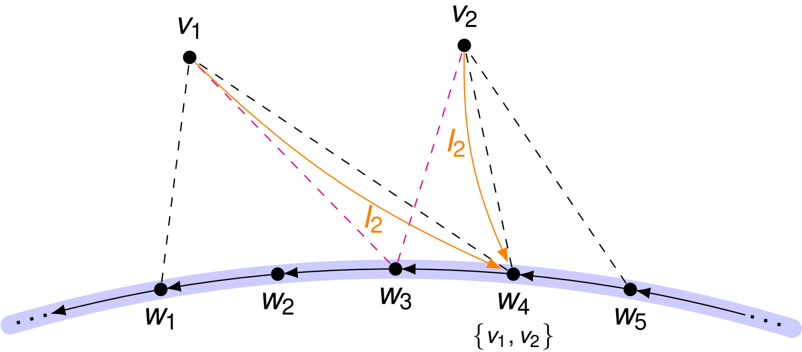
Random variable Y

Number of nodes that are integrated into C in a middle phase.

- Goal: Lower bound for Y that holds w.h.p.
- First we compute number of nodes outside C that send an invitation I_2?

Analysis of Middle Phases

- How many nodes outside C send an invitation I_2?



Analysis of Middle Phases

- A node $v \in V \setminus C$ sends an invitation if it is connected to at least one pair of consecutive nodes on C
- This event has probability p^2, but these events are not independent

- Event π_v: $v \in V \setminus C$ forms a triangle with at least one evenly numbered edge of C
- π_v has probability $1 - (1 - p^2)^{c/2}$ and π_v's are independent

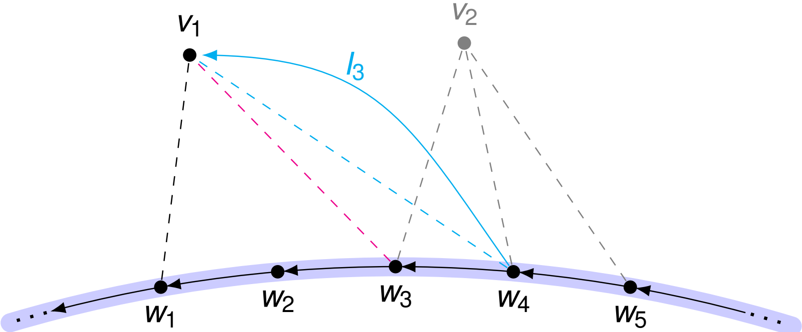
Random Variable *X*

X is the number of nodes v where π_v occurs. X is a lower bound for number of nodes that send invitation I_2.

$$E[X] = (n - c)(1 - (1 - p^2)^{c/2}) \quad (1)$$

Analysis of Middle Phases

- How many nodes on C send an accept message l_3?



Analysis of Middle Phases

- Computation of Y can be reduced to bins and balls model
  - X number of balls; c number of bins
  - Each ball is thrown randomly in any of the c bins
  - Probability that $v \in C$ is connected to w in $V \setminus C$ is independently of v and w equal to p.
- Y is equal to number of nonempty bins

$$E[Y|X = x] = c\left(1 - \left(1 - \frac{1}{c}\right)^{x}\right) \tag{2}$$

- Use Chernoff bound to find upper bound for Y that holds w.h.p.

Result

Theorem

For a random graph $G(n, p)$ with $p \in O(\log n/\sqrt{n})$ the following holds w.h.p.

1. *After phase 1 ring C consist of $O(\log n)$ nodes*
2. *After $O(\log n)$ middle phases C has size $O(\log n\sqrt{n})$ and the fiber graph has diameter $O(\log n)$*
3. *After final phase the maximum number of nodes attached to a single proxy is less than $e(n/c - 1)$ with probability $1 - 1/c$*

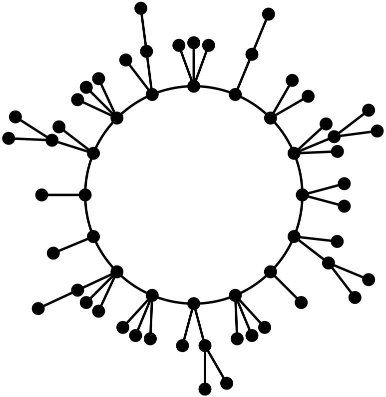
5

Conclusion & Extensions

Conclusion

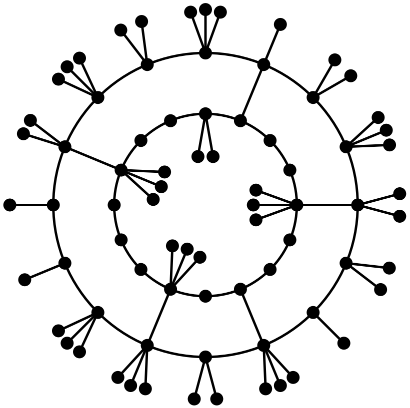
- Proposal for a new routing structure for pub/sub systems in wireless networks
- $O(\log n)$ distributed algorithm to build this structure
- Analysis for random graphs $G(n, p)$ with $p \in O(\log n/\sqrt{n})$

Extension I



Nodes outside ring need not to be connected directly to proxies. This allows to have even shorter rings

Extension II



The double ring structure allows PSVR to send more messages concurrently

A $O(\log n)$ Distributed Algorithm to Construct Routing Structures for Pub/Sub Systems

20^{th} Int. Symposium on Stabili ... ystems

Volker Turau

Professor

Phone +49 / (0)40 428 78 3530

e-Mail turau@tuhh.de

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

Institute of Telematics
Hamburg University of Technology

TUHH