Canopy: Publish/Subscribe with Upgraph Combination

Prof. Sasu Tarkoma
University of Helsinki

Markku Antikainen

Aalto University

19.3.2010 Global Internet Symposium

Contents

- Introduction
- Background
- Canopy Overview
- Upgraphs
- Conclusions

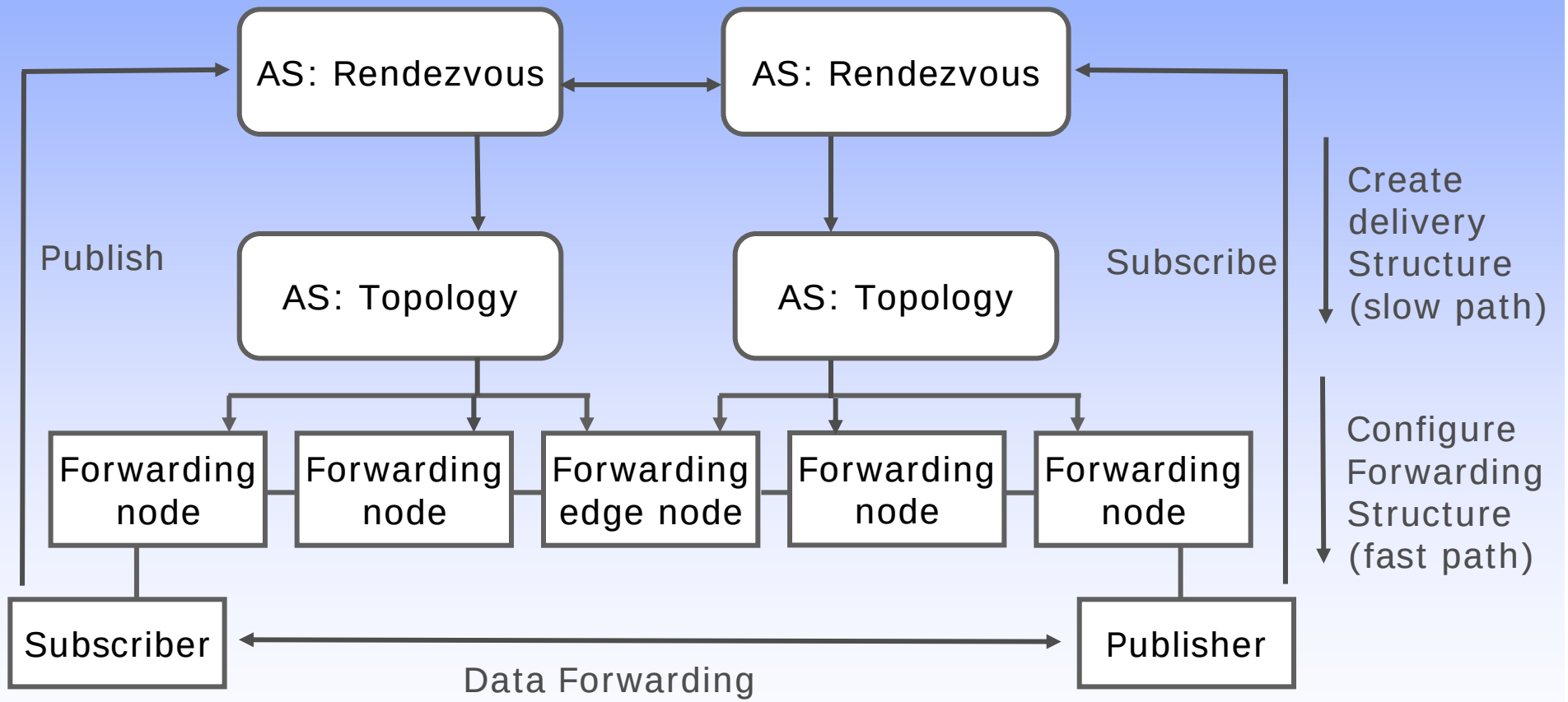
Introduction

- Network endpoints are more and more interested in content
- Publish/subscribe is a candidate paradigm for content distribution
  - Many-to-many, control distributed between subscribers (receivers) and publishers (senders)
- A content distribution system needs to know about packet replication points and policy-compliant paths
- This information can be given to the rendezvous system by the subscribers and publishers by sending their upgraphs to rendezvous entities

Background

- Related systems include NIRA, DONA, ROFL, CCN, and PSIRP
- PSIRP includes a 2-tier system where a hierarchical DHT based rendezvous interconnect network joins multiple rendezvous networks together for global reachability
  - Publications are identified by scopes and flat identifiers
  - Typically only scopes are advertised in the interconnect
  - Hierarchical structure guarantees locality for the communication

Overview of PSIRP pub/sub



Interdomain model for pub/sub

- Current Internet structure is the starting point
  - BGP and inter-AS relationships
- PSIRP network model
  - Autonomous domains as in BGP
  - Controlled by different organizations
  - Organizational policies
- The pub/sub inter-AS connections may result in different inter-AS relations than observed today
  - Multicast and caching

Upgraphs

- Definition
  - An upgraph is a dag that contains all the possible paths from a given node to the tier-1 core networks
  - Publisher upgraph roots from the publisher and contains all possible uphill and peer-to-peer paths
  - Subscriber upgraph does not contain peer-to-peer links, but only all possible uphill paths
- P2P paths in publisher upgraph
  - Combination at publisher-side
  - Typically there are more subscribers than publishers
- Discovery
  - Specific discovery process (NIRA)
  - Download BGP table (circa 300 000 entries, feasible for servers and desktops)
  - Receive upgraph from a network element or overlay node

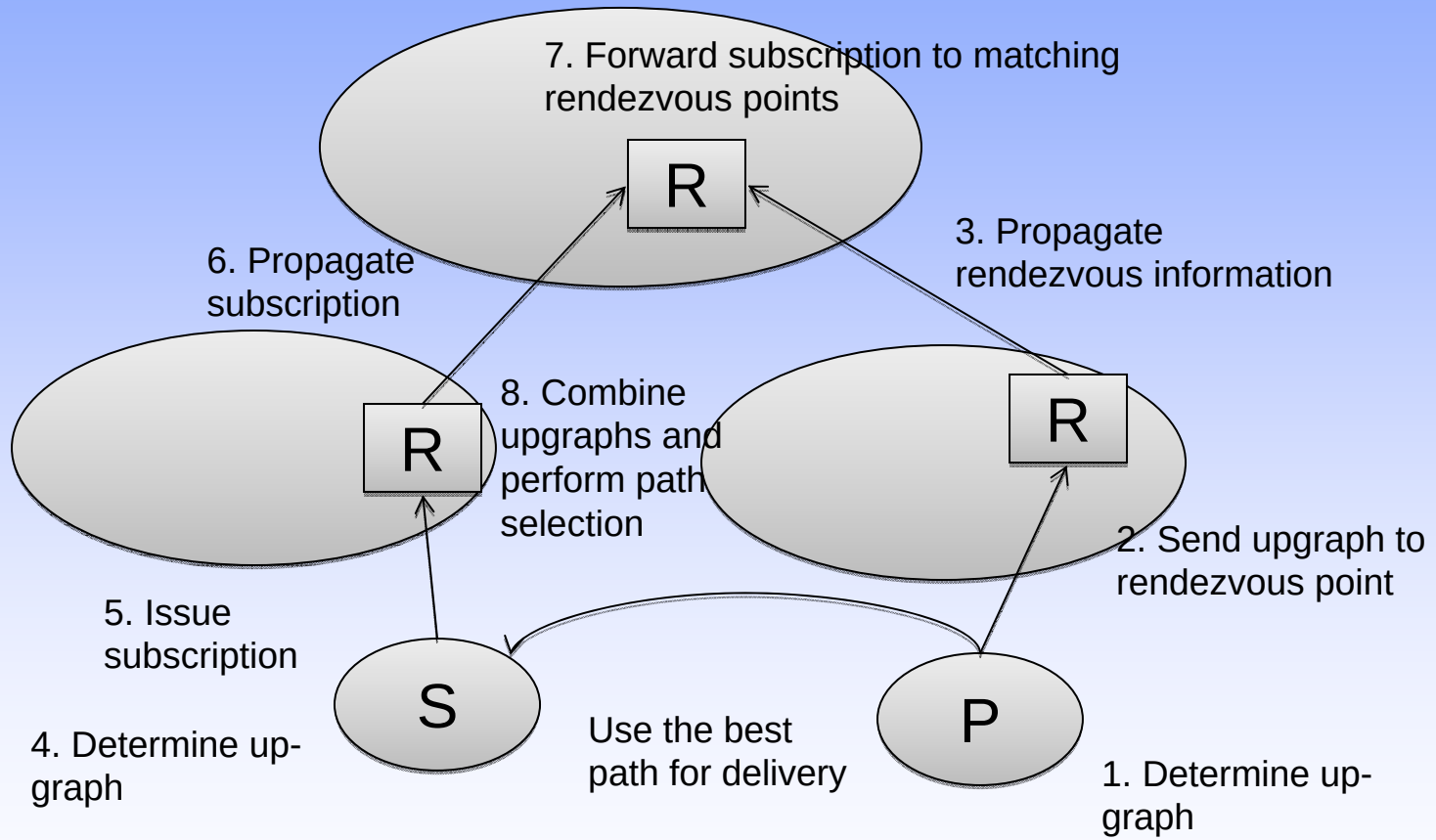
Upgraph combination

- Subscriber and publisher upgraphs can be combined by reversing the subscriber dags
- The upgraph combination gives all possible policy-compliant routes between the publisher and the subscriber
- The rendezvous system has the responsibility of choosing the most optimal route that is going to be used
- In order to increase publishers' and subscribers' possibilities to affect the chosen route, they can set preference values to the upgraph routes

Canopy: Using Upgraphs for Pub/Sub

- The NIRA system used upgraphs for
  - Allowing more control for receiver
  - Finding best paths for unicast
- Canopy uses upgraphs for pub/sub
  - Upgraphs combined at publisher-side rendezvous point
  - Can take both subscriber & publisher policies into account
  - Supports multi-path routing
  - Result is a policy-compliant multicast structure
  - Can be used for both overlays and on the network layer
  - Works with in-packet Bloom filter-based forwarding

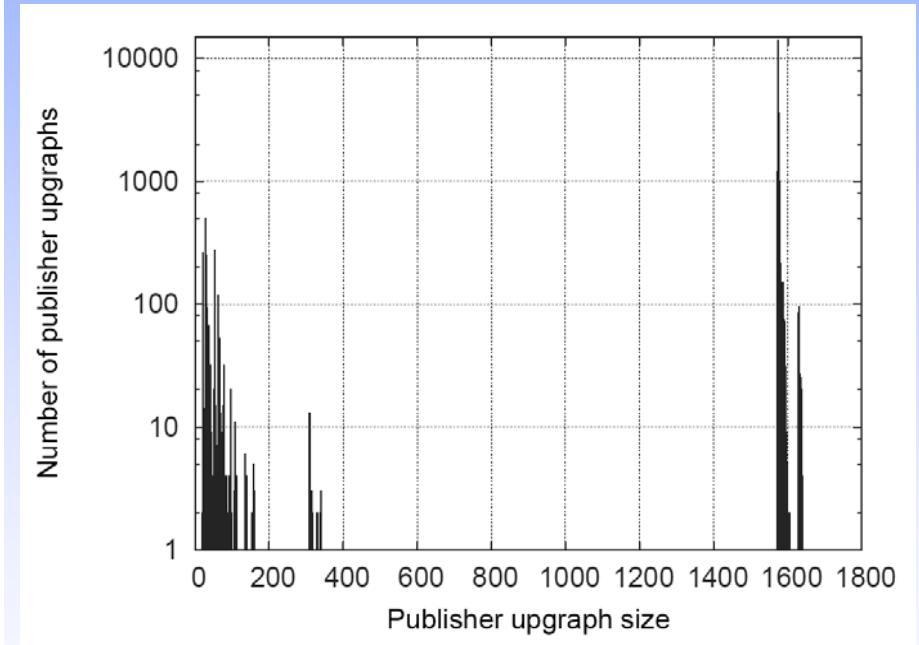
Canopy Overview

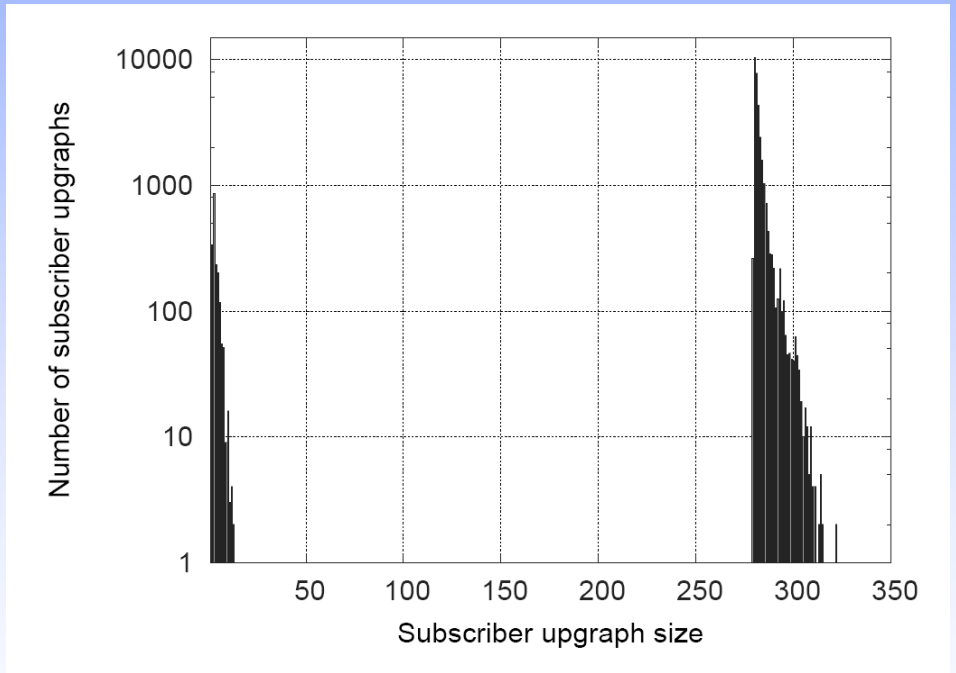


Analysis with CAIDA topology

- 33102 ASes in total, the median size of the publisher upgraph is 1574 nodes (average is 1470) and the median size of the subscriber's upgraph is 282 nodes (average is 264)
- Of the total number of ASes, 31 did not have any provider links making them tier-1 or similar networks.
- 4892 nodes have both provider and client links (tier-2), and 28179 nodes were observed to be content providers (tier-3).

Publisher and subscriber upgraphs





Conclusions

- Canopy uses upgraphs for determining policy-compliant multicast structures for publish/subscribe
  - Representing packet replication points and network paths in upgraphs
  - Combining subscriber and publisher upgraphs at designated rendezvous points in order to determine policy-compliant data distribution strategies
- We present a strategy for combining upgraphs and initial experimental analysis of upgraph histogram sizes based on the current CAIDA topology
- Initial analysis suggests that upgraph combination is a good candidate technology for implementing a global rendezvous system