UNIVERSITY OF CAMBRIDGE

Data Centric Networking (R202)

paper

An Evaluation of Scalable Application-level Multicast Built Using Peer-to-peer Overlays

authors: M. Castro et. al.

MPhil in ACS

reviewer/presenter: S. Trajanovski (*st508*)

Motivation

Overview of p2p multicast classification

- Lack of IP multicast deployment
- Difference in used overlay
  - Tree building
  - Flooding
- Difference in routing methods
  - Generalized hypercube routing (Pastry)
  - Cartesian hyper space (CAN)
- Paper structure..

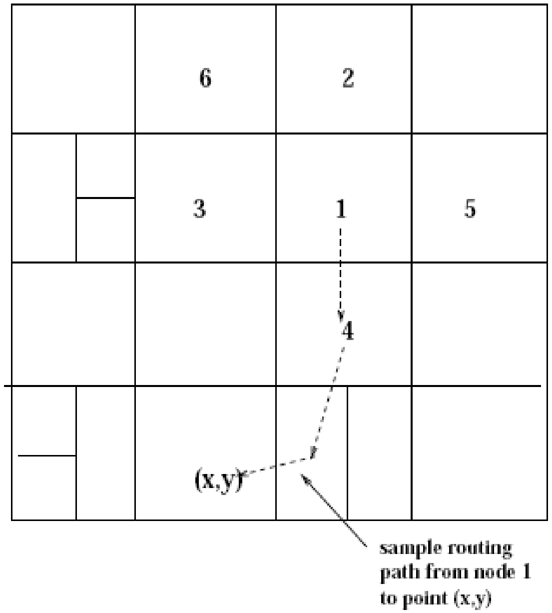
P2P overlay networks

Different routing methods (CAN)

- Topology organization
  - d – dimensional cube
  - Each node owns “its space”
- Node’s addition
  - Selected region is split into 2 parts
- Tunable parameters
  - Node dimension, multiple node/zone
  - Net. Aware routing, uniform partitioning etc.

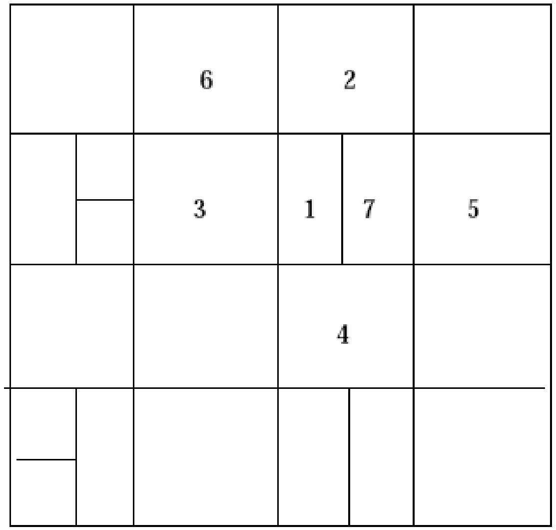
P2P overlay networks

Different routing methods (CAN)



1's coordinate neighbor set = {2,3,4,5}
7's coordinate neighbor set = { }

(before)



1's coordinate neighbor set = {2,3,4,7}
7's coordinate neighbor set = {1,2,4,5}

(after)

2-D CAN topology & node addition

P2P overlay networks

Different routing methods (Pastry)

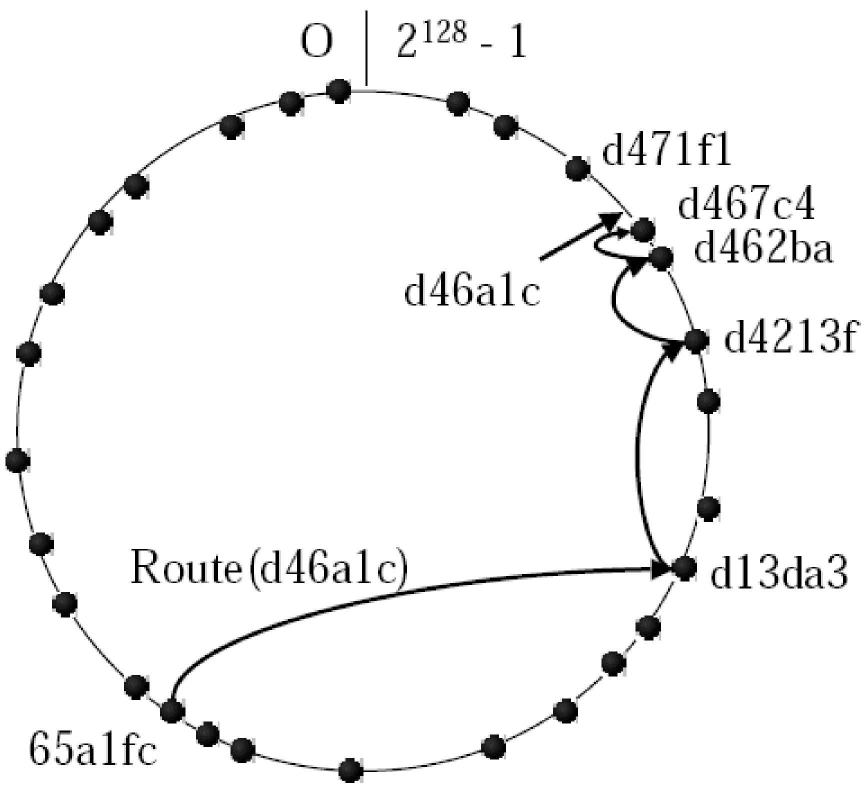
- Pastry namespace (128-bits)
  - nodeId & (message, dest_key)
  - Next node -> closest to the dest_key
- Routing principle
  - 128/b levels and 2^b entries each
  - Next node – sharing max. bits with destination
  - At least b bits “closer to destination”

P2P overlay networks

Different routing methods (Pastry)

0x	1x	2x	3x	4x	5x		7x	8x	9x	ax	bx	cx	dx	ex	fx
60x	61x	62x	63x	64x		66x	67x	68x	69x	6ax	6bx	6cx	6dx	6ex	6fx
650x	651x	652x	653x	654x	655x	656x	657x	658x	659x		65bx	65cx	65dx	65ex	65fx
65a0x		65a2x	65a3x	65a4x	65a5x	65a6x	65a7x	65a8x	65a9x	65aax	65abx	65acx	65adx	65aex	65afx

Routing table



Routing (65a1fc -> d467c4)

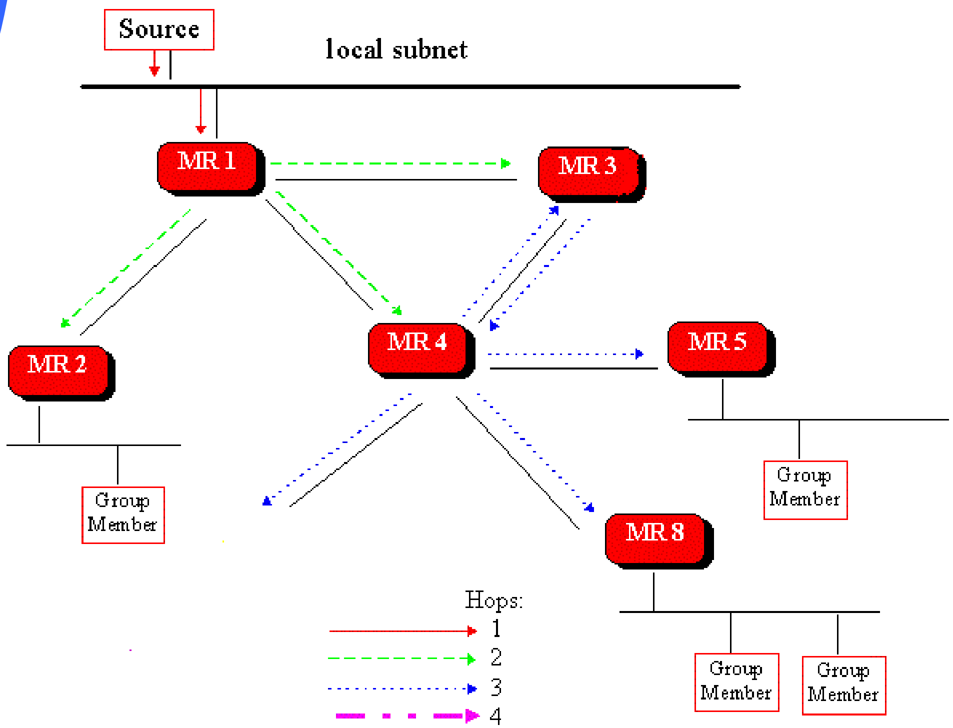
P2P overlay networks

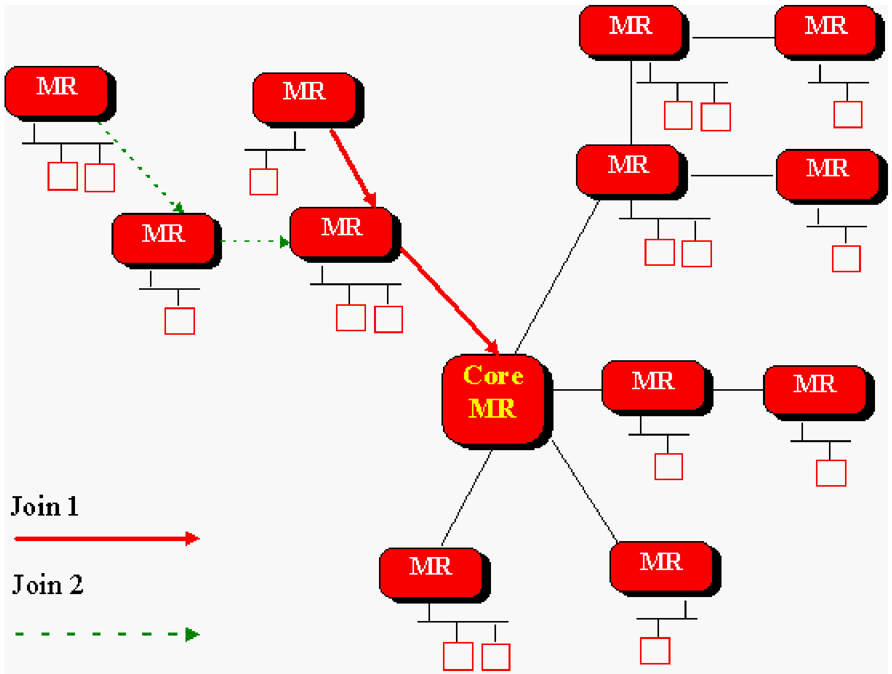
Overlay based application level multicast

- Flooding
  - Main concept: broadcasting
  - Groups: smaller subsets
  - CAN flooding: “naive” & CAN Multicast
  - Pastry flooding: level tagging and forwarding
- Tree – based (Scribe approach)
  - Firstly target group roots
  - Decentralized approach is scalable

P2P overlay networks

Different overlays approaches





flooding

tree-build architectures

Evaluation

Experimental setup

- Experimental environment
  - measured -> number of packets
  - 5 network topologies
  - 1st experimental set: one multicast group
  - 2nd experimental set: multiple multicast groups
  - Experimental phases
    - groups subscription
    - message is multicast to each group

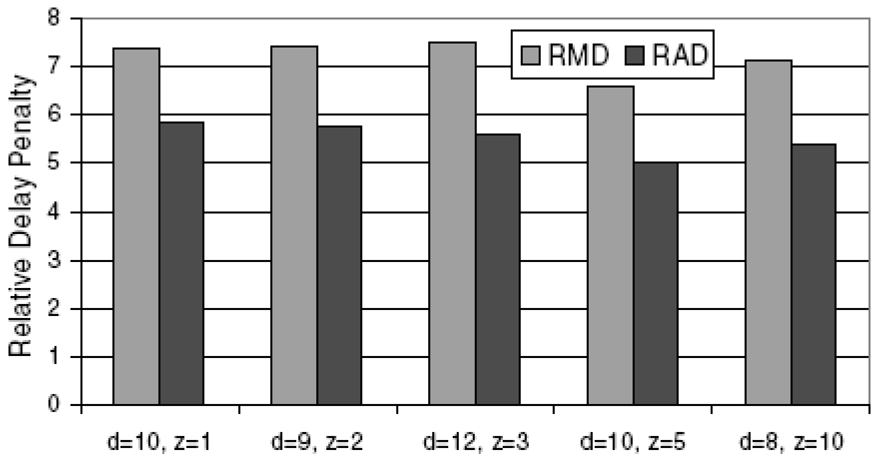
Evaluation

Experimental setup

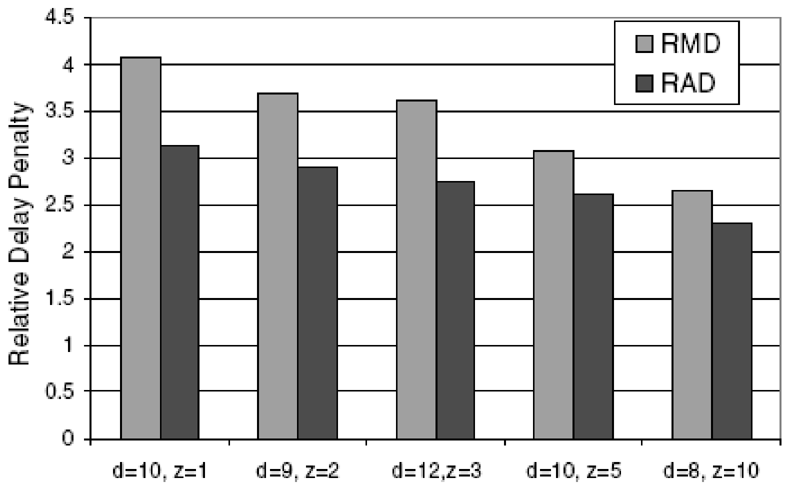
- Experimental criteria
- Relative (ratio app. level/IP multicast values)
  - Relative delay penalty
    - RMD (maximum ratio), RAD (average ratio)
  - Link stress
    - Number of packets over the link
  - Node stress
    - Routing table size (#nodes) & messages received
  - Message duplication

CAN Results

- Parameters
  - Number of dimensions (d); nodes per zone (z)
  - policy: (distance, ratio, NDR), uniform part. (on/off)



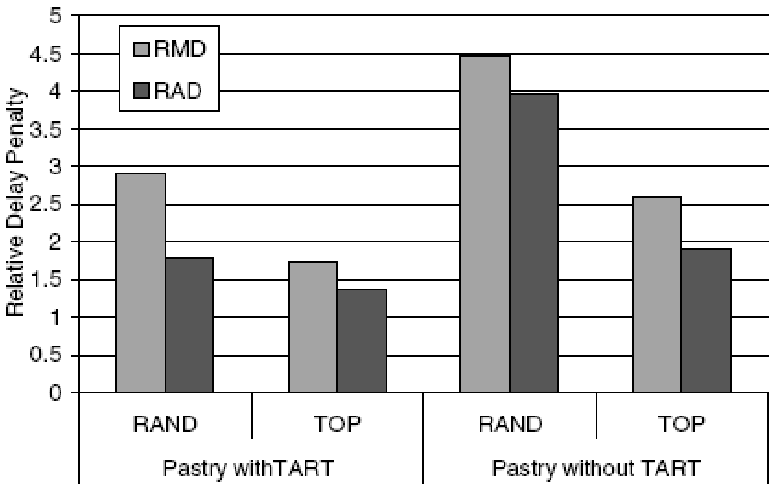
CAN flooding



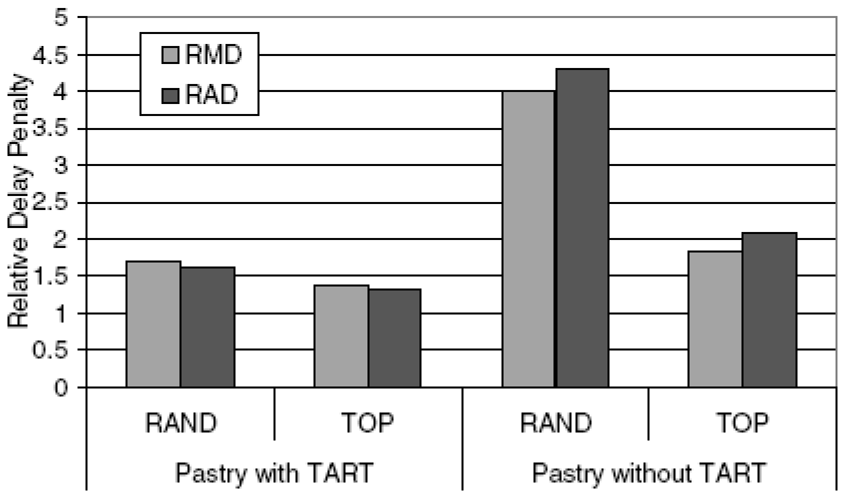
CAN tree-build

Pastry Results

- Parameters
  - b - #number of “matched” dest. bits (b={1,2,3,4})
  - TART & TOP



Pastry flooding



Pastry tree-build

More results

- Link Stress

Configuration	d=10 z=1	d=9 z=2	d=12 z=3	d=10 z=5	d=8 z=10
State size	18	29	38	59	111
Joining phase					
Max	91615	149341	197977	309212	416361
Average	154	183	219	281	431
Flooding phase					
Max	1958	1595	1333	985	631
Average	3.49	3.27	2.93	2.73	2.69

flooding in CAN

Configuration	d=10 z=1	d=9 z=2	d=12 z=3	d=10 z=5	d=8 z=10
State size	18	29	38	59	111
Max	323	220	198	184	225
Average	1.69	1.49	1.42	1.37	1.36

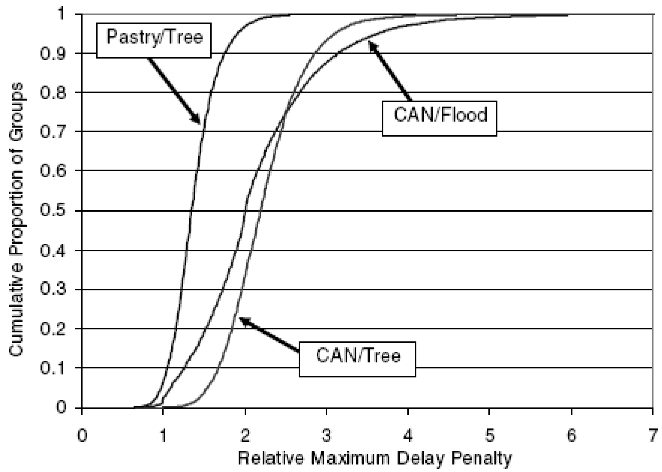
CAN tree-based

	with TART		without TART	
	RAND	TOP	RAND	TOP
Max	6801.4	65.4	2119.0	61.0
Average	4.3	1.4	4.6	1.4

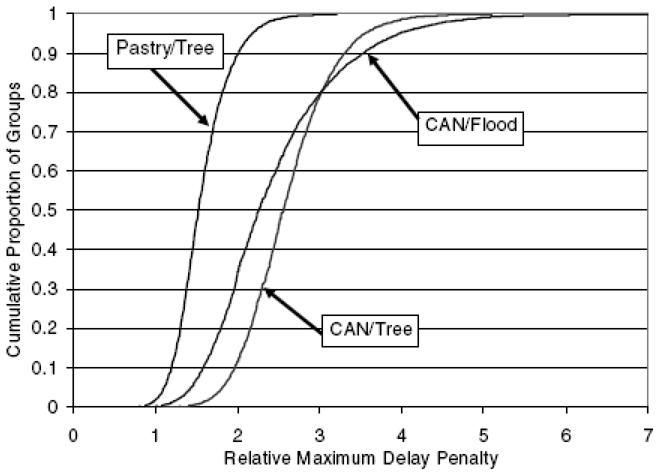
Pastry flooding for b=4

Comparative Results

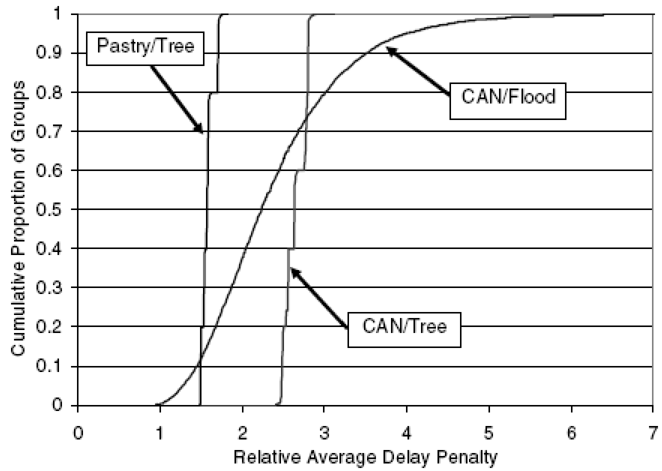
for multiple multicast groups (CDF functions)



RMD with localized members



RMD with distributed members



RMD with both
localized and distributed members

Related Work

- Tapestry and Chord
  - Similar approaches
- Bayeux and Overcast
  - Different concepts
  - Prospective future work
- Non scalable
  - End System Multicast, RONs, ISIS
- IP Multicast

Summary

- First head2head p2p analysis (4 comb.)
  - Flooding & tree building
  - Hypercube & Cartesian metric space
- Tree-based is better than flooding
- Multicast trees-build
  - Pastry better than CAN
- Flooding overlay costs more..
- Related work (further considerations)
  - Overcast, Bayeux, IP Multicast

- Questions??

- Discussion..