Overlay Optimizations for End-host Multicast

Wenjie Wang, David Helder, Sugih Jamin, Lixia Zhang
{wenjiew,dhelder,jamin}@eecs.umich.edu, lixia@cs.ucla.edu

Presented by Wenjie Wang
The University of Michigan

Oct, 2002
End-host Multicast (EM)

- IP multicast needs router and OS support.
- EM is easy to deploy: hosts connect to each other, form a virtual network (overlay).
- However, EM overlays need optimization.
Existing Overlay Schemes

- **Tree Overlay**
  - Simple structure, easy to refine, but has long latencies between members.
  - HMTP and Yoid are protocols that build a tree overlay.

- **Mesh Overlay**
  - Easy to build, complex structure, and hard to refine.
  - Narada and Hypercast are protocols that build a mesh overlay.

A tree overlay

A mesh overlay
Proposed Overlay Building Mechanism

- Build tree with HMTP, an efficient tree protocol.
- TMesh: tree with shortcuts.

- Extra links can provide shortcuts.

TMesh: HMTP with shortcuts.

Question: which shortcuts should be added?
Outline

- Performance of the proposed protocol
- Shortcuts categorization methods
- Performance effect of different kinds of shortcuts
- Heuristics to optimize our protocol
- Evaluation of our optimized protocol
How to Tell Which Overlay is Better?

- **Relative Delay Penalty (RDP):**
  - For a node pair, ratio of distance on the overlay to distance on the network.
  - For example, for B and E, distance on the overlay is 7; distance on the network is 3.

- **Average RDP (ARDP):** average RDP for all node pairs
  - The smaller the ARDP, the more efficient the overlay.

![Network and EM overlay](image1)

![Overlay](image2)
Motivation: Performance Comparison of Overlays

- TMesh can perform better than Narada, even with shortcuts added RANDOMLY.
- Can we make TMesh better? More specifically, what kind of links shall we add to make it better than random?
Keys to Link Selection

What kinds of links we prefer?

- the links that are the most beneficial to the whole group.
- the links that can be selected distributively by members.

How to find these links?

- Categorize links based ONLY on local information.
  - by RDP
  - by Utility
- Evaluate the global benefit of each kind of links.
Categorize Links by RDP

- RDP: ratio of the distance in overlay to the distance in the network.
- For member $F$, Link1 is a link with RDP $8/3$. 
RDP and ARDP

- Adding links with high RDP improves ARDP.
- Heuristic: adding link with RDP above a threshold (hRDP).
Categorize Links by Utility

- Utility: how much closer the link can bring the initiating node to other nodes.
- Utility = \[ \sum_{M} (D_{old}(M) - D_{new}(M)) / D_{old}(M) \]
- For example, for F, utility of Link1 is \((5/8+1/6 = 0.79)\), in which, \((8 - 3)/8\) (for A) + \((6 - 5)/6\) (for B).
Utility and ARDP

- High utility links contribute more to ARDP improvement (hUtil).
- hUtil: requiring knowledge of distance to all nodes, only possible when routing algorithm is available.
- What if the distances to all nodes are not available?
hUtil Based on Limited Distance Information

- Assume only distances to data senders are maintained.
- *Utility*: how much closer the link can bring the initiating node to all the data senders.
- *Average Utility*: $\text{Utility}/M$, where $M$ is the number of data senders.
Heuristic Evaluation

We evaluate the performance effect of different kinds of links and have the following heuristics:

- **hRDP**: add links with RDP above a threshold ($\sqrt{N}$).
- **hUtil**: add links with average utility above a threshold (1/8).

Experiment Setup:

- Network topology of size 4000, 6000, and 8000.
- Group size: varies from 50 to 1000.

*Reminder: ARDP is the average RDP of all node pairs.*
• TMesh achieves lower ARDP than Narada and uses less number of links.
• TMesh w/hRDP shows higher ARDP for small groups because it uses less number of links.
• TMesh w/hUtil (10% Senders) uses more links than other TMesh cases.
The Initial Overlay Matters

- Main difference between TMesh w/hUtil and Narada is the initial overlay.
- TMesh can be built on Yoid tree (YoidTM) and random tree (RandTM).
- HMTP < Yoid < RandTree; TMesh < YoidTM < RandTM
Nodes with high degrees consume more resources.

hUtil and hRDP use more low degree nodes.
Evaluation of Heuristics – Physical Link Stress

- Physical link stress (PLS): number of duplicate packets seen by the same physical link.
- The maximum PLS on TMesh is about half that of Narada.
Conclusion and Future Work

Conclusion:

- Two useful heuristics to optimize overlays: hRDP and hUtil.
- Overlay protocol TMesh:
  - Low latency between members
  - Good node degree distribution
  - Low physical link stress

Future work:

- Implement detailed protocol for TMesh.
- Conduct real Internet experiments.