Hybrid Multicast Implementation

Matthias Wählish, Thomas C. Schmidt
Stig Venaas

{waehlisch, t.schmidt}@ieee.org,
stig@cisco.com
Agenda

- Motivation
- Analytical Performance Evaluation
- Integration into Real-world Mcast Protocols
- Conclusion
Motivation – No Inter-domain Multicast
Hybrid Multicast

Idea:
- Connect different multicast islands
- Combine different technologies to provide group communication

Challenges:
- Allow for a self-organizing, unified distribution
- Find 'natural' way for the interplay of mcast protocols
- Performance of hybrid scenarios
Ingredients for Hybrid Multicast

- Multicast routing protocols in overlay and underlay
  - Definition for the interplay
- Appropriate multicast naming and mapping scheme
  - Common Multicast API
- Gateways
  - Easy to use and extendable system architecture
  - Discovery and configuration of gateways
  - HAMcast middleware
Scribe – RP-based Overlay Multicast (Castro et al 2002)

- Large-scale distribution service based on Pastry
- Rendezvous Point chosen from Pastry nodes
  - Choice according to group key ownership
  - RP roots shared distribution tree (analogue PIM-SM)
- Shared tree created according to reverse path forwarding
  - Nodes hold *children tables* for forwarding
  - New receiver routes a *SUBSCRIBE* towards the RP
  - *Subscribe* intercepted by intermediate nodes to update children table, reverse forwarding done, if node not already in tree
Multicast on CAN (Ratnasamy et al. 2001)

- Within a previously established CAN overlay members of a Group form a “mini” CAN
  - Group-ID is hashed into the original CAN
  - Owner of the Group key used as bootstrap node
- Multicasting is achieved by flooding messages over this mini CAN
- Number of multicast states is limited by $2d$ neighbours
  - independent of multicast source number!
- Can Multicast scales well up to very large group sizes
  - Replication load limited to neighbours ($2d$)
  - But tends to generate packet duplicates
CAN Forwarding

Ratnansamy et al. 2001
Performance Evaluation

Objectives:

- First order performance estimate, which can reveal the relative effects of different overlay approaches
- Derive a simple analytical model for the expected delay distribution in global hybrid multicast
Performance Evaluation Model

Observation:
- Performance of hybrid multicast is composed of
  - Inter-domain IP-layer distribution
  - Intra-domain transmission, which depends on overlay scheme in use
    - Two-layered distribution system
- Measurements for delay distributions are available
  - For example, Chalmers and Almeroth, TON, 2003

How do we derive a delay distribution for hybrid mcast?
Building the Performance Model

Common Assumptions:
- Delay of any IP link between routers is exponentially distributed
- Subsequent links perform independent of each other

Model: Details: See CoNEXT’09 student workshop paper
- Single link delay: $\beta$, and path length: $\alpha$
- Compound link delay of equally distributed links: $f_\Gamma(\alpha,\beta,\chi)$
Overall, Global Delay Distribution

Two-layered distribution (from Gammas):

\[ g(y) = C \cdot \left\{ f_\Gamma(\alpha_1 + \alpha_2, \beta_1, y) + \alpha_2 \left(1 - \frac{\beta_1}{\beta_2}\right) \cdot f_1^\cdot(\alpha_1 + \alpha_2 + 1, \beta_1, y) \right\} \]

with \( C = \left(1 + \alpha_2 \left(1 - \frac{\beta_1}{\beta_2}\right)\right)^{-1} \).

with parameters taken from external measurements:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter-AS Delay ((\beta_1))</td>
<td>10.91 ms</td>
</tr>
<tr>
<td>Intra-AS Delay ((\beta_2))</td>
<td>14.77 ms</td>
</tr>
<tr>
<td>Inter-AS Hopcount ((\alpha_1)) IP-level</td>
<td>4</td>
</tr>
<tr>
<td>Intra-AS Hopcount ((\alpha_2)) IP-level</td>
<td>5.5</td>
</tr>
<tr>
<td>Overlay Hopcount ((\alpha_1/d)) Scribe ((k = 16))</td>
<td>(\log_{16}(30.000) + 1)</td>
</tr>
<tr>
<td>Overlay Hopcount ((\alpha_1/d)) CAN ((D = 8))</td>
<td>(\frac{5}{30.000})</td>
</tr>
</tbody>
</table>
What Can We Expect?

A priori performance estimator:

![Graph showing probability distributions for different configurations.](image)
Protocol Engineering: Bringing native IP and OLM together

How do we couple native IP multicast routing protocols with overlay multicast?

Here, we concentrate on

1. DVMRP
2. PIM-SM
DVMRP

- Arbitrary router will not be informed about new receivers
- Immediately knows new sources
  - Prune/graft approach
- Source-specific trees + no central multicast instance

Relay Agent Operations:
- Receives all multicast underlay data automatically + joins stream
  - enableEvents(); new_source_event + join();
- Send all data to overlay + forward data to underlay
  - Initiate all-group join based on namespace extension in API
  - Underlying DVMRP will limit unwanted traffic automatically
PIM-SM

- Rendezvous Points receives multicast and listener states
  - Simplifies source and receiver awareness

- Designated routers of a PIM domain send receiver subscriptions towards RP

Relay Agent Operation:

- Place agent close to Rendezvous Point

- PIM register messages initiate new_source_event
  - Join the multicast group in underlay

- Join multicast group in overlay based on new_receiver_event in underlay
Conclusion

- Hybrid multicast schemes can be implemented by common multicast API
- Under the assumption of equally efficient implementations, hybrid inter-domain multicast can be deployed with little performance penalty on today's Internet
- Real-world measurements on the way