An a Priori Estimator for the Delay Distribution in Global Hybrid Multicast*

Matthias Wählisch
Freie Universität Berlin, Inst. für Informatik
Takustr. 9
D–14195 Berlin, Germany
waehlisch@ieee.org

Thomas C. Schmidt
HAW Hamburg, Department Informatik
Berliner Tor 7
D–20099 Hamburg, Germany
t.schmidt@ieee.org

ABSTRACT

Hybrid multicast is regarded a promising technology to overcome the inter-domain deployment problem for group communication. Realistic performance estimators are difficult to obtain due to the diversity of overlay concepts and their complex dependence on the global Internet topology that withstands straightforward simulations or measurements. We contribute a simple analytic model for the expected delay distribution. Parametrized by realistic measurement values, this should serve as a first order estimator for quantifying the delay penalties in a global-scale hybrid multicast packet distribution. We diagnose a strong dependence on hop counts and proximity awareness for the overlay multicast approach in use with promising results for most efficient schemes.

Categories and Subject Descriptors

C.2.2 [**Network Protocols**]: Computer-Communication Networks—*Routing protocols*; C.4 [**Performance of Systems**]: *Modeling techniques, Performance attributes*

General Terms

Performance, Theory

Keywords

Analytical model, IP multicast, structured overlay networks, inter-domain, Scribe, CAN

1. INTRODUCTION

Large-scale IPTV deployments and other multimedia services re-attract interest in globally available, scalable distribution of multicast data within the Internet. From a deployment perspective, hybrid approaches that interconnect isolated multicast-enabled domains by overlay techniques are considered more promising than complex inter-domain multicast routing protocols. Overlay schemes range from dynamic unicast tunneling (e.g., AMT) to structured and unstructured P2P multicast. Evidently, unicast tunnels optimize packet delays at the price of high data redundancy. Structured P2P overlays offer a promising trade-off between network and application efficiency, and their performance characteristics are well understood. However, empirical evaluations of globally distributed hybrid multicast solutions remain hard to obtain.

In the remainder, we briefly introduce the conceptual idea of hybrid multicast in § 2. The simple analytical model is derived in § 3, as well as the evaluations using CAN multicast and Scribe with and without proximity awareness. Our model gives a first order performance estimate, which can reveal the *relative* effects of the different overlay approaches.

2. HYBRID MULTICAST

The core idea of hybrid multicast is the deployment of a gateway function that bridges data between the local underlay and a global overlay as shown in Figure 1. The inter-domain multicast gateway (IMG) receives all IGMP/MLD reports to observe multicast listeners per domain, and intercepts the source stream via the overlay to forward it locally, if required [1]. IMGs can be co-located at existing routers.

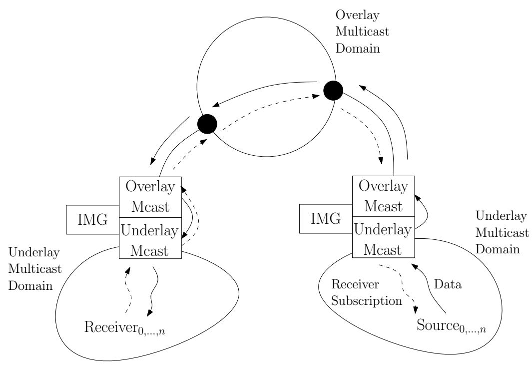

Figure 1: Schematic view of general inter-domain multicast gateway (IMG) scenarios

The overall performance of hybrid multicast is composed of the intra-domain IP-layer distribution, as well as the inter-domain transmission that depends on the overlay scheme in use. The constructed overlay structure may introduce additional hops and a delay stretch, which elongate the forwarding paths in contrast to a global IP-layer multicast solution.

3. PERFORMANCE EVALUATION MODEL

In this section, we present estimates on the expected performance of the hybrid multicast scheme at large-scale deployment. We concentrate on the most relevant measures for

*This work is supported by the German Bundesministerium für Bildung und Forschung within the project H∀Mcast (http://hamcast.realmv6.org).

CoNEXT Student Workshop'09, December 1, 2009, Rome, Italy.
ACM 978-1-60558-751-6/09/12.

Parameter	Value	Source
Inter-AS Delay (β_1)	10.91 ms	[2], Dataset 12/08
Intra-AS Delay (β_2)	14.77 ms	[2], Dataset 12/08
Inter-AS Hopcount (α_1) IP-level	4	[3], [2]
Intra-AS Hopcount (α_2) IP-level	5.5	[4]
Overlay Hopcount (α_1/d) Scribe ($k = 16$)	$\log_{16}(30.000) + 1$	[5]
Overlay Hopcount (α_1/d) CAN ($D = 8$)	$\sqrt[8]{30.000}$	[6]

Table 1: Performance parameters of the Internet and overlay networks

multimedia group services, the distribution of packet delays and in particular the average delay and delay variation (jitter). Therefore we rely on significant previous work about Internet topology, and on a simple model that reflects both, the inter-provider transition between ASs and the distribution therein.

Model: We consider a two-layer distribution system, one layer being situated on the intra-domain, the other operating between domains. For any IP link between routers, we make the common assumption that its delay is exponentially distributed with mean β, and that subsequent links perform independently of each other. When regarding a chain of α equally distributed links, the compound link delay is governed by a Gamma distribution $f_\Gamma(\alpha, \beta, x)$ [7]. Thus link delays in the underlay (intra-AS) and overlay (inter-AS) can individually be modeled by f_Γ.

To derive a joint probability law, we need to consider the distribution $g(y)$ of the sum of two independent Gamma-distributed random variable with parameters (α_i, β_i). For $\beta_1 \leq \beta_2$, $\beta_1 \sim \beta_2$, this can be well approximated [8] by

$$g(y) = C \cdot \left\{ f_\Gamma(\alpha_1 + \alpha_2, \beta_1, y) + \alpha_2 \left(1 - \frac{\beta_1}{\beta_2}\right) \cdot f_\Gamma(\alpha_1 + \alpha_2 + 1, \beta_1, y) \right\} \quad (1)$$

with $C = (1 + \alpha_2 (1 - \beta_1/\beta_2))^{-1}$.

$g(y)$ serves as the overall, global delay distribution for multicast packets traveling on average α_1 hops on the inter-AS level with an average delay of β_1, and α_2 hops on the intra-AS level with average delay β_2. These parameters can be extracted from different Internet measurements and overlay schemes, where we assume all 30.000 Autonomous Systems to participate in global overlay peering – cf., Table 1. We account for different overlay delay stretches d on the hop level, extracting for Scribe the values $d = 1.6/4.3$ with and without proximity neighbor selection, and for CAN the values $d = 5.5/8.2$. with / without proximity landmarking [9].

Results: In Figure 2, we compare the delay distributions of native multicast and hybrid solutions that deploy Scribe and multicast on CAN on the inter-AS level with and without proximity awareness. Tree-based Scribe with proximity route selection admits only limited delay penalties and a wide overlap in distribution with native multicast, whereas topology-unaware CAN flooding increases the delay scale by almost a factor of 4. Similar results become visible for the delay variation (jitter) that almost doubles at proximity-insensitive overlay schemes.

Discussion: These first order approximations do not claim to reproduce absolute delay values correctly, but should serve as a mean estimator for the relative effects of overlay multicast peering. From that perspective, our analysis reveals a strong dependency on hop distribution and delay stretch of the overlay scheme employed. Our structural estimates do not account for processing delay at forwarding devices,

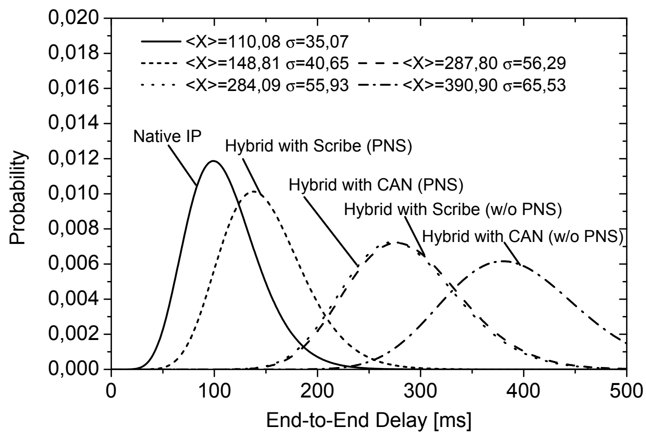

Figure 2: Delay distributions for global IP-layer and hybrid multicast using Scribe and CAN

which may differ between IP-based and overlay routing. Under the assumption of equally efficient implementations, we can conclude that hybrid inter-domain multicast can be deployed with little performance penalty on today's Internet.

4. REFERENCES

[1] M. Wählisch and T. C. Schmidt, "Between Underlay and Overlay: On Deployable, Efficient, Mobility-agnostic Group Communication Services," *Internet Research*, vol. 17, no. 5, pp. 519–534, November 2007.

[2] Y. Shavitt and E. Shir, "DIMES: Let the Internet Measure Itself," *ACM SIGCOMM CCR*, 35 (5), pp. 71–74, 2005.

[3] R. C. Chalmers and K. C. Almeroth, "On the topology of multicast trees," *IEEE/ACM Trans. Netw.*, vol. 11, no. 1, pp. 153–165, 2003.

[4] R. Govindan and P. Radoslavov, "An Analysis of the Internal Structure of Large Autonomous Systems," University of Southern California, CS Dept., Tech. Rep. 02-777, 2002.

[5] A. Rowstron, A.-M. Kermarrec, M. Castro, and P. Druschel, "Scribe: The Design of a Large-Scale and Event Notification Infrastructure," in *3rd COST264 Workshop, NGC 2001.*, ser. LNCS, vol. 2233. London, UK, 2001, pp. 30–43.

[6] S. Ratnasamy, M. Handley, R. M. Karp, and S. Shenker, "Application-Level Multicast Using Content-Addressable Networks," in *3rd COST264 Workshop, NGC 2001.*, ser. LNCS, vol. 2233. London, UK, 2001, pp. 14–29.

[7] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed. New York: Wiley & Sons, 1968, vol. 1.

[8] P. G. Moschopoulos, "The distribution of the sum of independent gamma random variables," *Annals of the Inst. of Stat. Mathematics*, vol. 37, no. 1, pp. 541–544, 1985.

[9] M. Castro, M. B. Jones, A.-M. Kermarrec, A. Rowstron, M. Theimer, H. Wang, and A. Wolman, "An Evaluation of Scalable Application-level Multicast Built Using Peer-to-peer Overlays," in *22nd Infocom 2003*, vol. 2. Washington, DC, USA: IEEE CompSoc, 2003, pp. 1510–1520.