Multisource and Multipath File Transfers through Publish-Subscribe Internetworking

Yannis Thomas*, Christos Tsilopoulos*, George Xylomenos*, George C. Polyzos*†
*Mobile Multimedia Laboratory, Dept. of Informatics, Athens University of Economics and Business
†Dept. of Computer Science and Engineering, University of California, San Diego
{thomasi, tsilochr, xgeorge, polyzos}@aueb.gr

***Abstract*—We present mmFTP, an information-centric and receiver-driven file transfer protocol for the Publish Subscribe Internetworking (PSI) architecture. mmFTP supports both multisource and multipath transfers, while requiring minimal complexity in terms of network operation. We describe the basic design and operation of mmFTP and present preliminary experimental performance results from a prototype implementation deployed in the PlanetLab testbed.**

***Index Terms*—information-centric networking; multipath; multisource; transport**

I. Introduction

Research in *Information Centric Networking* (ICN) seeks to design network architectures and protocols that efficiently utilize in-network resources including *communication*, *data storage* and *computation*, in order to optimize content distribution. Experience from such protocols and applications indicates that content distribution can highly benefit from multisource [1] and multipath [2] transfers, i.e. the use of multiple sources and multiple paths between source-destination respectively. To this end, we present the *Multisource and Multipath File Transfer Protocol* (mmFTP) for the *Publish Subscribe Internetworking* (PSI) ICN architecture [3]. mmFTP is designed to utilize all types of network resources by combining well-known content-distribution techniques into a single protocol. mmFTP is receiver-driven and supports on-path caching, thus it utilizes the network's short-term memory. In addition, mmFTP downloads files from multiple sources, hence it also utilizes the network's long-term memory. Finally, mmFTP supports multipath delivery, i.e. a transfer may use multiple distinct paths among source and destination, thus better utilizing available bandwidth resources and improving network load balancing.

The multisource and multipath features are incorporated in a single protocol without complicating network operation: mmFTP does not require complex signaling for connection establishment with multiple sources, or additional forwarding state in routers for multipath communication. This is due to the decoupling of content resolution, path formation and packet forwarding in PSI [4], along with the adoption of source-routing for data forwarding [5].

II. Multisource and Multipath File Transfers in PSI

mmFTP operation consists of two phases: (i) *slow-path* rendezvous for service establishment and (ii) *fast-path* rendezvous for the immediate host interaction for content delivery (Figure 1).

Slow-path rendezvous. File sources (publishers) advertise available content to the PSI network. An mmFTP receiver (subscriber) requests a file from the network. The network locates the available sources through its *DHT-based resolution* subsystem [3]. The network then selects the paths between the receiver and each source through the decoupled *Topology Management* subsystem [3]. Paths are encoded with LIPSIN [5], the Bloom filter-based source-routing scheme employed by PSI. The Topology Management subsystem may decide to utilize multiple paths per receiver-source pair, if such path diversity exists, thus enabling multipath communication between the receiver and a particular source. In the end, the LIPSIN-coded paths are handed to the receiver.

Fast-path rendezvous. Receivers then interact directly with sources via the path obtained during the slow-path phase. Files are fragmented into MTU-sized packets, each with its own name. The receiver sends *packet requests* to file sources. A packet request contains the packet's name, the forward LIPSIN source-route and the reverse LIPSIN source-route.[1] The request is forwarded to the source over the specified path. File sources, or intermediate routers, that have a cached copy of the requested packet, can extract the reverse LIPSIN source-route from intercepted requests and respond with the corresponding data packet.

Multisource downloads and multipath communications are realized with minimal control signaling, due to the unique characteristics of the PSI architecture. Multiple sources are selected by the PSI resolution system, which has an equivalent signaling cost with a DNS query. Multiple paths are selected by the decoupled Topology Management subsystem and then encoded as source-routes. Hence no additional signaling or forwarding state maintenance in routers is required. Furthermore, sources in mmFTP are stateless. Since the subscriber explicitly requests a packet using its unique identifier and also provides the LIPSIN-coded return path, the publisher can agnostically respond to the requests without performing complex operations. Finally, since LIPSIN allows source routing, the path management that takes place at the subscriber can be rather agile. Depending on the performance experienced, the subscriber can move traffic to different paths and switch from

[1] LIPSIN-coded paths are unidirectional, hence to transmit a data packet in the reverse direction, an explicit reverse path is needed.

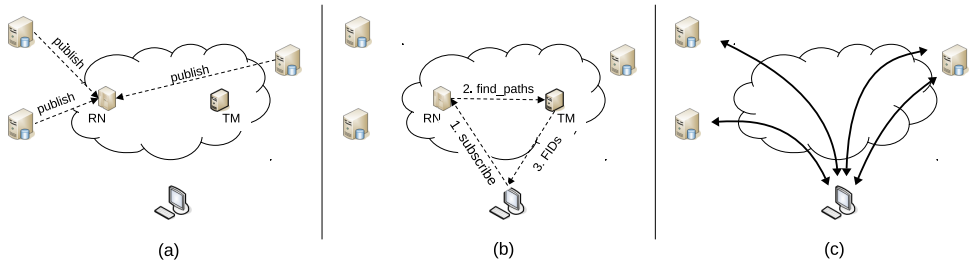

Fig. 1. mmFTP operation. (a) Publication, (b) Slow-path rendezvous, (c) Fast-path rendezvous. RN stands for Rendezvous Node, TM for Topology Management node and FID for LIPSIN Forwarding ID.

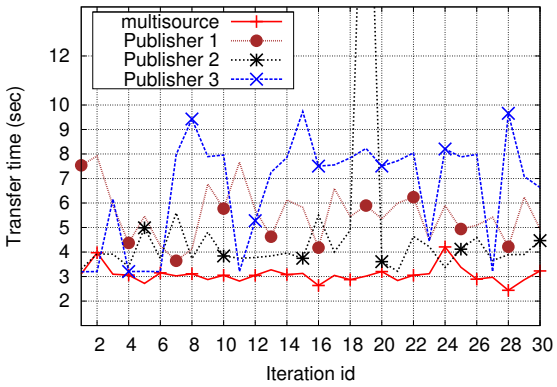

Fig. 2. Download times of single- and multi-source transfers.

multisource to multipath mode on-the-fly, applying congestion and flow control policies to better adjust to the network's state.

III. Implementation and Experimentation

A. Implementation

We implemented mmFTP over Blackadder, the PSI prototype implementation [6]. Our implementation includes mmFTP sender and receiver applications, as well as an extended Topology Management subsystem that computes multiple paths between a receiver and a source using the k-shortest paths algorithm by Yen [7], with hop count as the metric. The subscriber deploys a TCP-like congestion management scheme to each supplied path. The amount of requests transmitted over a particular path is decided through an AIMD sliding-window mechanism using the *slow start* and *congestion avoidance* algorithms [8]. Each path has its own private window management, thus constituting a separate flow.

We deployed Blackadder with mmFTP on the PlanetLab testbed to evaluate our design in a realistic environment. The deployment is realized as an overlay network: four Blackadder nodes scattered across Europe (Greece, Italy, Finland and UK) in a star topology.

B. Performance gains with multisource

The subscriber (located in Greece) downloads a 12 MB file from the 3 publishers. First, the file is downloaded in single-source mode from each publisher (three separate downloads) and then from all three publishers concurrently. Due to unpredictable congestion in the PlanetLab testbed, we performed 30 iterations of the experiment. Figure 2 shows the transfer time for each iteration. The best performance corresponds to the multisource case, with an average download time of 3.07 s (equivalent to 3.9 MB/s), which was 37% less than the fastest single-source (Publisher 2 with 4.8 s).

As evidenced by the spikes in Fig. 2, mmFTP is also much more stable in multisource mode: the variance of the download times for our 30 iterations was only 0.1 in multisource mode, while the smallest variance in single-source mode was 1.13 (Publisher 1). This shows that mmFTP dynamically avoids congested paths or sources, a common situation in the PlanetLab testbed.

IV. Conclusions and Future Work

mmFTP is a receiver-driven file transfer protocol for the PSI architecture. It supports multisource and multipath transfers. mmFTP combines well-known content distribution techniques into a single protocol, without requiring complicated network signalling or adding state to routers, due to the advantages of the PSI architecture. We implemented a prototype version of mmFTP and evaluated its performance on PlanetLab. Our preliminary results verify the effectiveness of multi-source delivery.

References

[1] S. Androutsellis-Theotokis and D. Spinellis, "A survey of peer-to-peer content distribution technologies," *ACM Comput. Surv.*, vol. 36, no. 4, pp. 335–371, 2004.

[2] D. Wischik, C. Raiciu, A. Greenhalgh, and M. Handley, "Design, implementation and evaluation of congestion control for multipath TCP," in *Proc. of the USENIX NSDI Conference*, 2011.

[3] G. Xylomenos, X. Vasilakos, C. Tsilopoulos, V. Siris, and G. Polyzos, "Caching and mobility support in a publish-subscribe Internet architecture," *IEEE Communications Magazine*, vol. 50, no. 7, pp. 52–58, 2012.

[4] D. Trossen, M. Särelä, and K. Sollins, "Arguments for an information-centric internetworking architecture," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 40, no. 2, pp. 26–33, 2010.

[5] P. Jokela, A. Zahemszky, C. Esteve-Rothenberg, S. Arianfar, and P. Nikander, "LIPSIN: line speed publish/subscribe inter-networking," in *Proc. of the ACM SIGCOMM Conference*, 2009.

[6] D. Trossen and G. Parisis, "Designing and realizing an information-centric internet," *IEEE Communications Magazine*, vol. 50, no. 7, pp. 60–67, 2012.

[7] J. Yen, "Finding the k shortest loopless paths in a network," *Science management*, vol. 17, no. 11, pp. 712–716, 1971.

[8] V. Jacobson, "Congestion avoidance and control," in *Proc. of the ACM SIGCOMM Conference*, 1988.