Fast Network Synchronization

Sachin K Agarwal, Avi Yaar, David Starobinski, Ari Trachtenberg
{ska, ayaar, staro, trachten}@bu.edu
Department of Electrical and Computer Engineering, Boston University

Much of the popularity of mobile computing and handheld devices can be attributed to their ability to deliver information to users on a seamless basis. Periodic synchronization of data between various devices is an enabling technology for this computing paradigm and an essential feature of any effective mobile and heterogeneous network architecture. We propose a peer-to-peer synchronization algorithm based on recent information-theoretic advances [3] and demonstrate an implementation of that algorithm through an improved PC-PDA synchronizer.

PalmPilot PDAs (Personal Digital Assistants) synchronize their data with that on a host PC either locally over a USB or serial connection, or remotely through a TCP/IP network - using a protocol known as HotSync [1, 2]. This protocol has two modes of operation: *FastSync*, in which stored modification flags identify the exact records that need to be exchanged; and *SlowSync*, in which all of the PDA s records are transferred to the PC for an item-by-item comparison to assess their modification status. FastSync represents the ideal case for network synchronization, where the differences between the synchronizing computers are known (through the modification flags) *a priori*. FastSync is limited to a two-host system, however, because each host maintains only one set of flags, so only differences between two specific hosts can be detected. The more generic SlowSync is representative of present-day network synchronization technology, which sacrifices network bandwidth and efficiency for robustness and simplicity.

The major problem of the SlowSync algorithm is that it does not scale with the size of the synchronizing data sets. This vulnerability will become more and more costly as data storage capacity outpaces the increase in network bandwidth. To address this scalability issue, we have developed and analyzed an algorithm termed CPISync (Characteristic Polynomial Interpolation Synchronization), based on a recent solution to the set reconciliation problem given in [3]. The CPISync algorithm uses interpolation of a rational function whose components can be factored to identify the differences between data sets stored on any two synchronizing computers. Since the number of sample points needed to correctly perform the interpolation depends only on the number of *differences* between the two synchronizing data sets, the communication complexity of CPISync is proportional to the differences between the sets. The algorithm's main drawback, however, is that the interpolation and factorization steps demand a cubic computation time (in the number of differences). The impact of this is lessened by the fact that the computation can be distributed asymmetrically according to the relative computing power of the synchronizing machines [1, 4].

We have implemented and deployed the CPISync algorithm in a PC-PalmPilot synchronization system, which we feel accurately models a peer-to-peer link in a heterogeneous network. Our PDA application, a memo entry program called *SyncMemo*, uses CPISync to synchronize data with its PC counterpart *PCMemo*. In tests against the PalmPilot s built-in *MemoPad* application, which uses the SlowSync algorithm, *SyncMemo* consistently synchronized faster, except in cases where there was a large number of differences between the synchronizing data sets. We generated results that helped to determine the communication and computation tradeoffs between CPISync and SlowSync in a variety of potential applications. The results from our prototype can be applied to reducing communication bottlenecks for many network protocols in large, heterogeneous systems, including the Domain Name System (DNS), rsync, heterogeneous mobile device synchronization, resource discovery and link state routing.

Website: *http://ipsit.bu.edu/nislab*

This work was supported in part by NSF Career Grant No. CCR-0133521

REFERENCES

[1] S. Agarwal, *Data Synchronization in Mobile and Distributed Networks,* Masters Thesis, July 2002.

[2] S. Agarwal, D. Starobinski and A. Trachtenberg, *On the Scalability of Data Synchronization Protocols for PDAs and Mobile Devices*, IEEE Network (Special Issue on Scalability in Communication Networks), Vol. 16, No.4, pp.22-28, July/August 2002.

[3] Y. Minsky, A. Trachtenberg and R. Zippel, *Set Reconciliation with Nearly Optimal Communication Complexity*, International Symposium on Information Theory, Washington D.C., June 2001.

[4] A. Trachtenberg, D. Starobinski, and S. Agarwal, *Fast PDA Synchronization using Characteristic Polynomial Interpolation*, Proceedings of the IEEE INFOCOM 2002, pp. 1510-1519, New York.