On Designing Overlay Networks for Peer–to–Peer Systems

Gennaro Cordasco

December 2005
Abstract

Over the last five years, Peer-to-Peer (P2P) networking has generated a tremendous amount of interest world-wide among both Internet surfers and computer networking professionals. As a matter of fact since May 1999 when Shawn Fanning launched the first P2P file-sharing application, Napster [Nap], the phenomenon of file-sharing has continued to sensationaHTMLay grow in popularity and appears set to remain an important feature of the Internet for the immediate future. However, as this dissertation will argue, P2P is much more than just a way of exchanging MP3s over the Internet.

Peer-to-Peer systems rely on the computing power and bandwidth of the participants in the network (peers) rather than concentrating it in a relatively few servers.

One of the most important benefits provided by P2P systems is the “Scalability”. In a P2P systems, each consumer of the resources also donates resources. Nevertheless, “Scalability” has been recognized as the central challenge in designing such systems.

Unfortunately, the initial designs for P2P systems had significant scaling problems. Thus, the chaotic, ad hoc topologies of the first-generation Peer-to-Peer architectures have been replaced by a set of topologies with an emergent order, provable properties as well as excellent performance. Indeed, several research groups have – independently – proposed the second generation of P2P systems (DHT systems) that support a Distributed Hash Table (DHT) functionality [DR01, ZHS+04, SML+03, RFH+01].

In a DHT system, files and peers are associated with a key and each peer in the system is responsible for storing a certain range of keys. Peers set up direct connections among themselves based on their key and form an overlay network with each node having several other peers as neighbors. These connection are used for routing messages between the peers.
Chord [SML+03] is one of the most popular DHT system proposed in the literature. It implements the idea of “consistent hashing” [KLL+97] to build the Distributed Hash Table (DHT) based on several independent nodes and exhibits an overlay network with logarithmic node degree and diameter.

In this dissertation the design and implementation of overlay networks based on DHTs are discussed with the main features of this contribution being:

- The proposal of a family of novel Peer-to-Peer overlay networks, based on the Fibonacci number system, retaining all positive aspects that made Chord [SML+03] a popular topology for routing in P2P networks. The schemes proposed simultaneously improve on the maximum/average number of hops for lookups and the routing table size per node.

- The proposal of routing schemes that optimize the average number of hops for lookup requests in Peer-to-Peer overlay networks. This work is inspired by the recently introduced variation of greedy routing, called neighbor-of-neighbor (NoN) [MNW04], which allows an optimal average path length with respect to the degree to be obtained. This strategy does not make use of randomization and as a consequence, the NoN technique can be implemented within these schemes without adding any overhead. Analyzed networks include several popular topologies: Chord, F-Chord(α), Hypercube based networks, Symphony, Skip-graphs. The improvement is obtained with no harm to the operational efficiency (e.g. stability, ease of programming, scalability, fault-tolerance) of the considered systems.

- The proposal of a family of Distributed Hash Table systems with the aim of combining routing efficiency of the randomized networks – average path length $O(\log n / \log \log n)$ vs. the $O(\log n)$ average path length of the deterministic system – with the programmability and start-up efficiency of a uniform system – i.e., a system in which the overlay network is transitive and greedy routing optimal. The proposed family is parameterized with a positive integer $c$ which measures the amount of randomness that is used. Varying the value $c$, the system goes from the deterministic case ($c = 1$) to an “almost uniform” system. Increasing $c$ to relatively low values allows routing with optimal average path length while retaining most of the advantages of a uniform system, such as easy programmability and quick boot-
strap of the nodes entering the system. A matching lower bound for the average path length of the family of routing schemes for any $c$ is also given.