Brief Announcement: Deterministic SkipNet

Nicholas J. A. Harvey
Microsoft Research
Redmond, WA
nickhar@microsoft.com

J. Ian Munro
School of Computer Science
University of Waterloo
Waterloo, Canada
imunro@cs.uwaterloo.ca

ABSTRACT

We present a deterministic scalable overlay network. In contrast, most previous overlays use randomness or hashing (pseudo-randomness) to achieve a uniform distribution of data and routing traffic.

1. DETERMINISTIC SKIPNET

Many self-organizing peer-to-peer overlay networks have been proposed as building blocks for distributed applications. For most of these overlays, such as Chord [4], their primary purpose is to form a *distributed hash table*. SkipNet [2] is a new overlay based on Skip Lists that organizes nodes and data primarily by their sorted string names, as Skip Lists do, rather than by hashes of those names. Consequently, SkipNet supports useful locality properties in addition to the usual distributed hash table functionality.

This paper describes Deterministic SkipNet, a variant of SkipNet with absolute performance bounds. In contrast, SkipNet is a randomized structure and thus, like most previous overlays, its performance bounds are probabilistic. The only other deterministic scalable overlay design we know of is presented in [1]. Our work herein builds on, but is not an immediate consequence of, Deterministic Skip Lists [3].

In SkipNet, every node has two IDs: a string name ID and a sequence of random digits called a numeric ID. Each numeric ID digit is in $[0, k-1]$, where k is a fixed parameter. As shown in Figure 1, all nodes are connected at level 0 by a "root ring" that is sorted by the nodes' name IDs. Nodes are also members of higher level rings in which the routing pointers skip over many nodes; these rings are also sorted by the nodes' name IDs. All nodes sharing a common numeric ID prefix of length h are members of the same level h ring, which is labeled with their common prefix. Nodes maintain routing pointers to their neighbours in each ring. In a SkipNet with N nodes, the following bounds hold with high probability: node insertion/departure requires $O(\log N)$ messages, searches by name ID or numeric ID require $O(\log N)$ messages, and $O(\log N)$ routing pointers are stored per node.

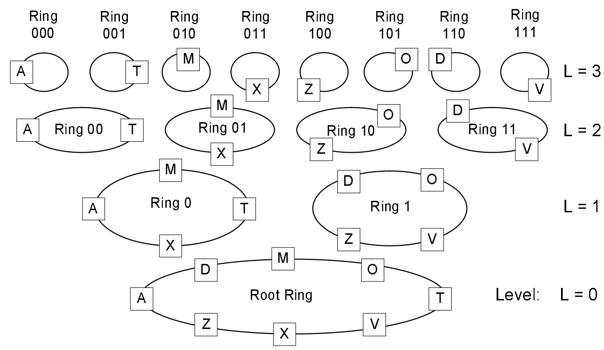

Figure 1: *A SkipNet structure with $k = 2$.*

Deterministic SkipNet nodes also have name IDs and numeric IDs. However, whereas SkipNet's numeric IDs are randomly chosen and are immutable, Deterministic SkipNet's numeric IDs are deterministically chosen and adjusted during node insertion/departure. Also, whereas SkipNet allows k to be arbitrary, Deterministic SkipNet fixes $k = 3$.

To ensure the balance of the ring structure, we maintain an invariant that any two adjacent nodes at height h are between 2 and 5 hops apart at height $h - 1$. Thus, a ring at height h has at most $N/2^h$ nodes, the maximum height of any ring is $O(\log N)$, and $O(\log N)$ routing pointers are stored per node. A simple argument shows that searches by name ID require $O(\log N)$ messages.

The node insertion/departure algorithms must maintain the stated invariant by examining the numeric IDs of neighbouring nodes and performing any necessary *rotations*. At each level, $O(1)$ messages and at most two rotations are required. Summing over all levels, $O(\log N)$ rotations are required. A single rotation at level h consists of two nodes swapping their numeric ID digits and pointers at levels h and above, and requires $O(\log N)$ messages. In total, the insertion/departure algorithms require $O(\log^2 N)$ messages.

2. REFERENCES

[1] I. Abraham et al. A generic scheme for building overlay networks in adversarial scenarios. In *IPDPS '03*.

[2] N. Harvey et al. SkipNet: A Scalable Overlay Network with Practical Locality Properties. In *USITS '03*.

[3] J. I. Munro, T. Papadakis, and R. Sedgewick. Deterministic Skip Lists. In *SODA '92*.

[4] I. Stoica et al. Chord: A scalable Peer-To-Peer lookup service for internet applications. In *SIGCOMM '01*.