Brief Announcement: Stretch Between Nearby Peers*

Kirsten Hildrum, John D. Kubiatowicz, Jeremy Stribling
{hildrum, kubitron}@cs.berkeley.edu, strib@csail.mit.edu

Categories and Subject Descriptors: C.2.1 [COMPUTER-COMMUNICATION NETWORKS]: Network Protocols

General Terms: Algorithms, Theory

Keywords: Stretch, object location stretch

A peer-to-peer object location system is an evolving set of computers cooperating to store objects. In this brief announcement, we analyze the *stretch* in such a system. Stretch is the ratio of the distance traveled in the system to find a copy of the object to the direct distance to the closest copy. We give a simple example with low overall stretch (i.e., the average stretch over all pairs) but high stretch nearby pairs, and then show, via simulation, that the simple example is relevant to real networks. More details on this and the related issue of finding nearby neighbors to build systems for low stretch are in the accompanying technical report [1].

For the simple example, consider n overlay nodes at the integers on the number line, from 1 to n, such that adjacent nodes are separated by a distance of one. (We assume that n is a power of 2.) A particular object ID and the routes taken to that ID through the overlay create a logical tree on these nodes. Suppose the tree is "perfect", meaning that exactly every other node is a leaf, and consider the pair of nodes $(2k-1, 2k)$, for some $k > 0$. If $2k$ is a leaf, then $2k$'s distance to its parent $2k-1$ is one, and $2k-1$'s distance to its parent (itself) is zero; the average distance from a node to its parent is $\frac{1}{2}$. To simplify the calculation, imagine a hypothetical average parent located at a distance of $\frac{1}{2}$ from both nodes, at $2k - \frac{1}{2}$. These are not physical nodes; rather, they the "average" over all possible choices. Figure 1 shows the resulting tree.

In Tapestry [2], objects are *published* by placing pointers to them at each node along the path from the publisher to the root, and object location proceeds by checking for pointers along the path from the query source to the root. The stretch is determined by where the publish and search path first intersect, or in other words, the least common ancestor of the publisher and the query source.

We now calculate the average stretch when the publisher and searcher are at distance one from each other. Half of such pairs of nearby nodes share the same parent, and so have location stretch of one, since the request need only travel a distance of one to reach the publisher. Half of the remaining pairs share the same grandparent, for these nodes, the stretch is three (see Figure 1 for an example). In general, for $j \le \log_2 n$ there are $2^{\log_2 n - j}$ pairs with location stretch $2^j - 1$. The average stretch is then $\frac{1}{n}\sum_{j=1}^{\log_2 n} 2^{\log_2 n - j}(2^j - 1)$, or $\log_2 n - \sum_{j=1}^{\log_2 n} \frac{1}{2^j}$, which is less than or equal to $\log n - 1$. Thus, the average stretch between nearby pairs is $O(\log n)$ (and one pair has stretch n), while the average stretch over all pairs is $O(1)$. (The $O(1)$ overall average follows because most pairs are far apart, and the stretch between those pairs is constant.)

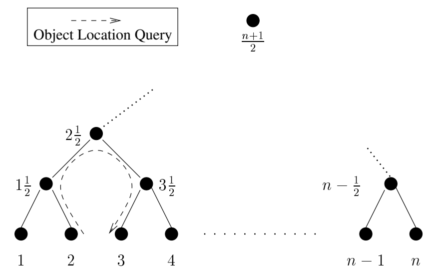

Figure 1: A perfect base-2 Tapestry tree, with a stretch three object location query.

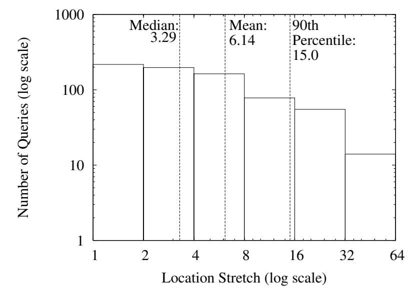

Figure 2: The distribution of stretch for nearby pairs in Tapestry on a simulated network.

Real networks are not lines, and real systems do not produce the perfect tree described above, but the results are relevant. Our thought experiment predicts that the number of pairs with a least common ancestor at level k decreases exponentially in k, while the stretch experienced increases exponentially.

We ran Tapestry [2] on a simulated transit stub network and measured the stretch between nearby pairs. Figure 2 shows the resulting histogram. While most pairs have low stretch, some pairs show very high stretch, matching the predictions of the simple example. In the simulations, the overall mean stretch for queries between any two nodes is just 3.01 (less than half of the mean stretch for close objects), demonstrating that indeed overall measurements can obscure information about the stretch between nearby nodes.

Conclusion A system with low overall object location stretch may still have high stretch for pairs that are nearby. If objects are likely to be placed near the nodes looking for them, overall stretch is probably not the right measure.

REFERENCES

[1] HILDRUM, K., KUBIATOWICZ, J., AND STRIBLING, J. Technical report: Low stretch between nearby peers. Tech. Rep. UCB//CSD-04-1328, UC Berkeley, 2004.

[2] ZHAO, B. Y., HUANG, L., RHEA, S. C., STRIBLING, J., JOSEPH, A. D., AND KUBIATOWICZ, J. Tapestry: A global-scale overlay for rapid service deployment. *IEEE Journal on Selected Areas in Communications* (2003).

*Supported by NSF career award ANI-9985250, NFS ITR award CCR-0085899, and California MICRO award 00-049, and NSF Cooperative Agreement No. ANI-0225660.

PODC'04, July 25–28, 2004, St. John's, Newfoundland, Canada.
ACM 1-58113-802-4/04/0007.