# The Advantage of Balanced-Allocation Routing for ATM Networks

Aris Anagnostopoulos[1]
Computer Science Department
Brown University, Box 1910
Providence, RI 02912-1910, USA
e-mail: aris@cs.brown.edu

Ioannis Kontoyiannis[2]
Division of Applied Mathematics and
Dept. of Computer Science, Brown Univ.
Box F, Providence, RI 02912, USA.
e-mail: yiannis@dam.brown.edu

Eli Upfal[1]
Computer Science Department
Brown University, Box 1910
Providence, RI 02912-1910, USA
e-mail: eli@cs.brown.edu

*Abstract* — **We compare the long-term, steady-state performance of a variant of the standard *Dynamic Alternative Routing (DAR)* technique commonly used in telephone and ATM networks, to the performance of a path-selection algorithm based on the "balanced-allocation" principle; we refer to this new algorithm as the *Balanced Dynamic Alternative Routing (BDAR)* algorithm. We show that, at the expense of a minor increase in routing overhead, the BDAR algorithm gives a substantial improvement in network performance, in terms both of network congestion and of bandwidth requirement.**

## I. Problem Description and Main Result

Fast, high bandwidth, circuit switching telecommunications systems such as ATM and telephone networks often employ a limited path-selection algorithm in order to fully utilize the network resources while minimizing routing overhead. Typically, between each pair of nodes in the network there is a dedicated bandwidth for communication. This dedicated bandwidth is chosen in order to satisfy the demand for communication between these stations. Only when this bandwidth is exhausted the admission control protocol tries to find an alternative route through intermediate nodes. To minimize overhead and routing delays, the protocol checks just a small number of alternative routes; if there are no free connections available on any of these alternatives, then the call or communication request is rejected. Implementations that use this technique include the Dynamic Alternate Routing (DAR) algorithm used by British Telecom, and AT&T's Dynamic Nonhierarchical Routing (DNHR) algorithm.

A common feature in these (and other) currently implemented protocols is the sequential examination of alternative routes. Only when the algorithm examines a route and finds it cannot be used an alternative one is examined. The criteria for when a route can or should be used, and the method by which the alternative route is selected have been the subject of extensive research, in particular, in the context of British Telecom's DAR algorithm.

In the types of networks considered here, a logical link or "bandwidth" is reserved between each pair of stations, and an alternative route is only used when this logical link has already been exhausted. We model such a network as the complete graph $G = (V, E)$ with $|V| = n$ vertices (stations) and $|E| = \binom{n}{2}$ edges (links). Each edge has a capacity of $2B$ calls, half of it dedicated to direct and half to alternative routes.

The input to the system is a sequence of call requests, which are assumed to arrive at Poisson times: New calls onto each link (i.e., between each pair of nodes) arrive according to a Poisson process with rate $\lambda$, all arrival streams being independent. Similarly, the duration of a call is independent of all arrival times, all other call durations, and it is exponentially distributed with mean $1/\mu$.

The routing algorithm has to process the calls on-line, that is, the $t$-th request is either assigned a path or rejected before the algorithm receives the $(t+1)$-th request. Once a call is assigned to a path, that path cannot be changed throughout the duration of the call.

As in most of our results we consider large networks with the number $n$ of nodes growing to infinity, we will also assume that the capacity parameter $B$ may vary with $n$. Specifically, we assume that $B = B_n$ is nondecreasing in $n$, and we also allow the possibility $B = \infty$.

The goal in designing an efficient routing protocol is to assign routes to the maximum possible number of call requests without violating the capacity constraints on the edges. We compare the performance of the following two protocols.

The *d-Dynamic Alternative Routing (DAR) algorithm* works as follows. When a new call request arrives, it tries to route the call through the direct (one-link) path. If there is no available bandwidth on the direct path, then the algorithm sequentially chooses alternative routes of length two and assigns the call to the first available. Up to $d$ such choices are made, and they are made at random. If no possible path is found, then the request is rejected.

The *d-Balanced Dynamic Alternative Routing (BDAR) algorithm* also assigns a new call request to the direct path if there is available bandwidth. If not, then the algorithm chooses $d$ length-two alternative paths at random, and compares the maximum load among them (where the load of such a path is taken to be the maximum load of the two links on that path). Then the call is assigned to the path with the minimum load. As before, if there is no path with free bandwidth among these $d$ choices, then the call is rejected.

For our purposes, the main performance measure is the minimum required bandwidth that ensures that, under the stationary distribution of the network, the probability that a new call is rejected is appropriately small. Our main result shows that the required bandwidth for the BDAR policy is exponentially smaller than that required for the DAR, and is summarized in the following theorem.

**Theorem** *Assume that all the edges have a capacity of $2B_n$ links. Under the DAR policy, edge capacity $B_n = \Omega\left(\sqrt{\frac{\ln n}{d \ln \ln n}}\right)$, as $n \to \infty$, is necessary to ensure that a new call is not lost with high probability.*

*On the other hand, if we perform the BDAR policy, edge capacity $B_n = \frac{\ln \ln n}{\ln d} + o\left(\frac{\ln \ln n}{\ln d}\right)$, as $n \to \infty$, suffices to ensure that a new call is not lost with high probability.*

The full version of the paper can be found at `http://www.cs.brown.edu/people/aris/pubs/balanced.pdf`

[1]Supported in part by NSF grants CCR-0121154, and DMI-0121495.

[2]Supported in part by NSF grant #0073378-CCR.



0-7803-7728-1/03/$17.00 ©2003 IEEE.