Steady State Analysis of Balanced-Allocation Routing

Aris Anagnostopoulos,[1,*] Ioannis Kontoyiannis,[2,†] Eli Upfal[1,*]

[1] *Computer Science Department, Brown University, Box 1910, Providence, Rhode Island 02912-1910; e-mail: {aris, eli}@cs.brown.edu*

[2] *Division of Applied Mathematics and Department of Computer Science, Brown University, Box F, 182 George Street, Providence, Rhode Island 02912; e-mail: yiannis@dam.brown.edu*

Received 27 September 2002; revised 7 November 2003; accepted 24 November 2004
DOI 10.1002/rsa.20071
Published online 7 April 2005 in Wiley InterScience (www.interscience.wiley.com).

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 [Y. Azer, A. Z. Broder, A. R. Karlin, and E. Upfal, SIAM J Comput 29(1) (2000), 180–200; M. Mitzenmacher, Ph.D. Thesis, University of California, Berkeley, August 1996]; we refer to this new algorithm as the *Balanced Dynamic Alternative Routing (BDAR)* algorithm. While DAR checks alternative routes sequentially until available bandwidth is found, the BDAR algorithm compares and chooses the best among a small number of alternatives. 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. © 2005 Wiley Periodicals, Inc. Random Struct. Alg., 26, 446–467, 2005

1. INTRODUCTION

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

Correspondence to: E. Upfal

*Supported in part by NSF Grants CCR-0121154 and DMI-0121495.
†Supported in part by NSF Grant #0073378-CCR and USDA-IFAFS Grant #00-52100-9615.

the network resources while minimizing routing overhead. Typically, between each pair of nodes in the network there is a dedicated bandwidth for communication; namely, no more than a certain fixed number of calls can be simultaneously active between each pair of nodes. This dedicated bandwidth is chosen in order to satisfy the demand for communication between these stations. Only when this bandwidth is exhausted does the admission control protocol try 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 [7], and AT&T's Dynamic Nonhierarchical Routing (DNHR) algorithm [1].

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 is an alternative one examined. The criteria for when a route can or should be used, and the method in which the alternative route is selected have been the subject of extensive research, in particular, in the context of British Telecom's DAR algorithm [6, 7, 8]; see Kelly [9] for an extensive survey.

Dynamic routing can be viewed as a special case of the online load balancing problem, where the load (incoming calls or requests) may be assigned to one or more servers (network links), and jobs (communication requests) can be scheduled only on specific subsets (paths) of the set of servers, as defined by the network topology. In this paper we study the impact of replacing the sequential searches of the routing algorithm by a version of the *balanced-allocation principle*. The basic idea is as follows: Instead of sequentially choosing alternative options (in our case, paths) until a desirable one is found, in the balanced-allocation regime the algorithm randomly chooses and examines a number of possible options, and assigns the job at hand to the option that appears to be the best at the time of the assignment.

A number of papers have demonstrated the advantage of the application of the balanced-allocation principle [2, 3, 4, 16, 17] for standard load balancing problems, where jobs require only one server and can be executed by any server in the system. This research has shown that balanced allocations usually produce a very substantial improvement in performance, at the cost of a small increase in overhead: Since several alternatives are examined even when the first alternative would have been satisfactory, the complexity of the routing algorithm is increased. But, as has been shown before and as we also demonstrate in the present context, examining even a very small number of alternative (thus increasing overhead by a very small amount) can offer great performance improvements.

The idea of employing the balanced-allocation principle to the problem of dynamic network routing as described in this paper was first explored in [11]. In this context the goal is to reduce system congestion and minimize the blocking probability, that is, the probability that a call request is rejected. The main difficulty in applying and analyzing the balanced-allocation principle in a network setting is in handling the dependencies imposed by the topology of the network. The preliminary results in [11] show that the advantage of balanced allocations is so significant that it holds even in the presence of a set of dependencies.

The performance of a routing protocol can be analyzed in a static (finite, discrete time) or in a dynamic (infinite, continuous time) setting. The static case has been extensively studied in [10], extending and strengthening the results in [11]. In this paper we consider the continuous-time case. The analysis of the continuous-time case suggested in [11] was based on applying Kurtz's density-dependent jump Markov chain technique, following the supermarket model analysis in [16, 17]. However, since the argument in [11] is incomplete, we present here a different analysis. Our results concern the long-term behavior of large networks employing a routing protocol based on the balanced-allocation principle. The main tools we employ are a Lyapunov drift criterion used to establish the existence of a stationary distribution for the BDAR routing protocol, and a continuous-time extension of the technique in [3], used to analyze the stationary behavior of a network.

Balanced allocations have also been studied in the context of *queueing* networks, where analogous results (under different asymptotic regimes than the ones in this paper) are obtained in [12, 16, 20, 21], among others.

1.1. Model Description and Main Results

In the types of networks considered in this paper, 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 undirected graph $G = (V, E)$ with $|V| = n$ vertices (stations) and $|E| = N = \binom{n}{2}$ edges (links).

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 λ, 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 tth 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. We assume that each edge has a capacity of $3B$ circuits (one circuit can transmit one call), where $1/3$ of this capacity is reserved for direct calls (namely, it will only be used for call requests between these two nodes), and the rest is reserved for being used as part of an alternative route between two stations.

As in most of our results we consider large networks with a 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 will 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 are no available circuits on the direct path, then the algorithm sequentially chooses alternative routes of length two, without replacement, and assigns the call to the first available path. 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 are available circuits. If not, then the algorithm chooses d length-two alternative paths at random, with replacement, and compares the maximum load among them (in the exact sense that we describe later). Then the call is assigned to the path with the minimum load. As before, if there is no path with free circuits among these d choices, then the call is rejected.

Consider some link e between two stations u and v, with a capacity of $3B$ circuits, from which B are reserved for routing calls between u and v. The rest of the $2B$ circuits, which are reserved for alternative paths, are further split into two. B circuits are reserved for routing calls with u as one of the endpoint station communicating, and B circuits for calls with v as the endpoint.

The model described so far, together with one of the two protocols above, induces a continuous-time stochastic process describing the behavior of the network. As we show below, this system (for fixed n) converges to a stationary regime exponentially fast. For our purposes, the main performance measure is the minimum required bandwidth that ensures that, under the stationary distribution of the network, the blocking probability (i.e., the probability that a new call is rejected) is appropriately small.

In this paper our main goal is to compare the performance of the DAR algorithm with that of BDAR. It is clear that BDAR's performance is dominated by its performance on alternative (length-two) routes. Therefore, in order to simplify the analysis, we consider a variant of BDAR, called BDAR*, which ignores the direct links and services each call only via an alternative route, making use only of the $2B$ alternative connections of each edge. In other words, we assume that each edge has capacity $2B$ and all of it is dedicated to alternative routes. We show that even though the BDAR* policy ignores the direct links, it has superior performance compared to DAR.

The following result illustrates this superiority by exhibiting explicit asymptotic bounds on their bandwidth requirements. It follows from the results in Theorems 5 and 6.

Theorem 1. *Assume that all the edges have a capacity of* $3B$ *circuits. Under the DAR policy, capacity*

$$B = \Omega\left(\sqrt{\frac{\ln n}{d \ln\ln n}}\right), \qquad \textit{as } n \to \infty$$

is necessary to ensure that, under the stationary distribution, a new call is not lost with high probability.

On the other hand if we perform the BDAR policy (thus ignoring the B direct links), capacity*

$$B = \frac{\ln\ln n}{\ln d} + o\left(\frac{\ln\ln n}{\ln d}\right), \qquad \textit{as } n \to \infty$$

suffices to ensure that, under the stationary distribution, a new call is not lost with high probability.

In the above result and throughout the paper, we say that a limiting statement holds "with high probability" (abbreviated "whp") if it holds with probability that is at least $1 - 1/n^c$ for some constant $c > 0$. For example, when we say that a random variable "$X_n = O(\ln n)$ whp," we mean that there are positive constants C and c such that $\mathbf{Pr}(X_n \leq C \ln n) \geq 1 - 1/n^c$ for all n large enough. Similarly, "$X_n = o(\ln n)$ whp" means that there is a $c > 0$ such that, for all $\epsilon > 0$, $\mathbf{Pr}(X_n \leq \epsilon \ln n) \geq 1 - 1/n^c$ for all n large enough.

Note that the result of Theorem 1 is exactly analogous to that obtained in [10] in the discrete-time case.

2. ANALYSIS OF BALANCED-ALLOCATION ROUTING

This section presents the main contribution of this paper, a steady state analysis of the performance of the BDAR* routing algorithm. The network is a complete graph with n nodes and $N = \binom{n}{2}$ undirected edges. New calls arrive at Poisson times with rate λ and their durations are exponentially distributed with mean $1/\mu$, as described earlier. As it turns out, an important parameter in the analysis of the network load is the ratio $\rho = \lambda/\mu$.

2.1. Unbounded Capacities

We first analyze the maximum load on edges when the algorithm is used on a network with unbounded edge capacity, corresponding to $B = B_n = \infty$. Consider some ordering of the edges, and let

$$\Gamma = \{(e, e') : e, e' \in E,\ e < e',\ e \text{ adjacent to } e'\},$$

be the set of edge pairs that are adjacent to each other. For every pair of adjacent edges $(e, e') \in \Gamma$, let $c_{e,e'}(t)$ denote the number of calls at time t that use edges e and e' (recall that every alternate path consists of two links). Then the above model induces a continuous-time Markov process $\mathbf{\Phi} = \{\Phi(t)\ :\ t \geq 0\}$, evolving on the state space

$$\Sigma = \mathbb{N}^{N(n-2)},$$

where

$$\Phi(t) = (c_{e,e'}(t))_{(e,e')\in\Gamma}.$$

For an edge $e = (u, v)$ we define also $\ell_{e,v}(t)$ to be the number of calls at time t that use edge e and have node v as an endpoint:

$$\ell_{e,v}(t) = \sum_{e' :\ \substack{(e',e)\in\Gamma,\, v \text{ not} \\ \text{adjacent to } e'}} c_{e',e}(t) + \sum_{e' :\ \substack{(e,e')\in\Gamma,\, v \text{ not} \\ \text{adjacent to } e'}} c_{e,e'}(t),$$

and we also define $\ell_e(t)$ to be its combined load at time t, that is,

$$\begin{aligned}\ell_e(t) &= \ell_{e,v}(t) + \ell_{e,u}(t) \\ &= \sum_{e':(e',e)\in\Gamma} c_{e',e}(t) + \sum_{e':(e,e')\in\Gamma} c_{e,e'}(t).\end{aligned}$$

Assume that a call arrives at time t on edge $e = (u, v)$. Algorithm BDAR* selects d nodes uniformly at random with replacement, from $V\backslash\{u, v\}$. Name these nodes $\{w_i\}$ for $i = 1, 2, \ldots, d$, and the corresponding edges $e_i^u = (u, w_i)$ and $e_i^v = (w_i, v)$. The call is then assigned to the path (e_i^u, e_i^v) corresponding to the minimum i satisfying

$$\max\{\ell_{e_i^u,u}(t-), \ell_{e_i^v,v}(t-)\} = \min_{j=1,2,\ldots,d} \max\{\ell_{e_j^u,u}(t-), \ell_{e_j^v,v}(t-)\}.$$

In the above expression, and throughout the entire paper, $f(t-)$ denotes the left-side limit of function f at t, namely, $\lim_{\delta\downarrow 0} f(t - \delta)$. Note that instead of selecting the minimum i satisfying the above expression, we can choose any Markovian rule. Finally, we define

$$M^v_{\geq i}(t) = \sum_{e:e \text{ incident to } v} (\ell_{e,v}(t) - i + 1)^+,$$

$$L^v_{\geq i}(t) = \sum_{e:e \text{ incident to } v} \mathbf{1}_{\{\ell_{e,v}(t) \geq i\}},$$

where $\mathbf{1}_{\mathcal{E}}$ denotes the indicator function of event $\mathcal{E}$, and $x^+ = \max\{x, 0\}$. In words, $L^v_{\geq i}(t)$ counts the number of edges incident to node v with at least i calls with v as an endpoint at time t, and $M^v_{\geq i}(t)$ counts the excess above i at time t on edges incident to v, of calls that have node v as an endpoint. Trivially we have $L^v_{\geq i}(t) \leq M^v_{\geq i}(t)$.

As we show next, this Markov process has a stationary distribution π_n to which it converges exponentially fast, regardless of the initial state of the network. We then prove a high probability bound on the maximum load on any edge in the system under this stationary distribution.

The process $\mathbf{\Phi}$ evolves on Σ according to the model described above. This evolution is formalized by the transition semigroup $\{P^t : t \geq 0\}$ of $\mathbf{\Phi}$, where $P^t(\boldsymbol{c}, \boldsymbol{c}')$ is simply the probability that $\mathbf{\Phi}$ is in state $\boldsymbol{c}'$ at time t given that it was in state $\boldsymbol{c}$ at time zero, $P^t(\boldsymbol{c}, \boldsymbol{c}') = \mathbf{Pr}(\Phi(t) = \boldsymbol{c}' \mid \Phi(0) = \boldsymbol{c})$.

Our first result shows that $\mathbf{\Phi}$ has a stationary (or invariant) distribution to which it converges exponentially fast. It is stated in terms of the "Lyapunov function" $V(x)$, which is defined as 1 + (total number of active calls in state $x \in \Sigma$):

$$V(x) = V(\{c_{e,e'} : (e, e') \in \Gamma\}) = 1 + \sum_{(e,e') \in \Gamma} c_{e,e'} \tag{1}$$

where $c_{e,e'}$ counts the number of calls in state x that use edges e and e'.

Theorem 2. *Assume that the BDAR* algorithm is used on a network with n nodes, each of which has infinite capacity. Then the induced Markov process $\mathbf{\Phi}$ has a unique invariant distribution π_n, and, moreover, for any initial state $x \in \Sigma$, the distribution of $\Phi(t)$ converges to π_n exponentially fast, namely, there is a constant $\gamma < 1$, such that*

$$\sup_y |P^t(x, y) - \pi_n(y)| \leq V(x)\gamma^t, \quad \textit{for all } t \geq 0 \textit{ and all } x \in \Sigma.$$

Proof. Our proof uses the Lyapunov drift criterion for the exponential ergodicity of a continuous time Markov process [13, 5, 14]. To state our main tool, we recall a few definitions, adapted to our case of countable state space.

The *generator* $\mathcal{A}$ of the process $\mathbf{\Phi}$ is a linear operator on functions $F : \Sigma \to \mathbb{R}$ defined by

$$\mathcal{A}F(x) = \lim_{h \downarrow 0} \frac{\mathrm{E}(F(\Phi(h)) \mid \Phi(0) = x) - F(x)}{h}$$

whenever the above limit exists for all $x \in \Sigma$. The *explosion time* of $\mathbf{\Phi}$ is defined as

$$\zeta = \sup_n J_n,$$

where

$$J_0 = 0, \qquad J_{n+1} = \inf\{t \geq J_n : \Phi(t) \neq \Phi(J_n)\}$$

($J_0, J_1, \ldots$ are the jump times of the Markov process). We say $\mathbf{\Phi}$ is *nonexplosive* if $\mathbf{Pr}(\zeta = \infty \mid \Phi(0) = x) = 1$ for any starting state x.

The following theorem follows from the more general results in [14, 5], specialized to the case of a continuous-time Markov process with a countable state space.

Theorem 3. [14, 5] *Suppose a Markov process evolving on a countable state space that is nonexplosive, irreducible (with respect to the counting measure on Σ) and aperiodic. If there exists a finite set $C \subset \Sigma$, constants $b < \infty$, $\beta > 0$ and a function $V : \Sigma \to [1, \infty)$, such that,*

$$\mathcal{A}V(x) \leq -\beta V(x) + b\mathbf{1}_C(x), \qquad x \in \Sigma, \tag{2}$$

then the process is positive recurrent with some invariant probability measure π, and there exist constants $\gamma < 1$, $D < \infty$ such that

$$\sup_y |P^t(x,y) - \pi(y)| \leq D\,V(x)\gamma^t, \quad \textit{for all } t \geq 0 \textit{ and all } x \in \Sigma.$$

It is easy to verify that the process is ψ-irreducible and aperiodic, with the maximal aperiodicity measure ψ being the counting measure on Σ.[1] Also the process is nonexplosive since the number of new calls in a given interval has a Poisson distribution with a finite mean; therefore, the probability of infinite number of transitions in a finite interval is 0.

To show that the drift criterion (2) can be satisfied, we use the Lyapunov function $V(x) = 1 +$ (total number of active calls in state x) defined in Eq. (1) above.

In order to compute $\mathcal{A}V$ we notice that when a new call enters the system, it increases the loads of two edges by 1, hence the value of V by 1, and when a call terminates the value of V decreases by 1. Therefore, new calls are generated with rate λN and calls are terminated at a rate $\mu(V(x) - 1)$. The probability that in a time interval h there are 2 or more new calls or terminations of calls is $o(h)$.[2] Using these observations we can compute $\mathcal{A}V$:

$$\begin{aligned}\mathcal{A}V(x) &= \lim_{h\downarrow 0} \frac{V(x) + \lambda N \cdot h - \mu \cdot (V(x) - 1) \cdot h + o(h) - V(x)}{h} \\ &= \lambda N - \mu V(x) + \mu.\end{aligned}$$

We define

$$C = \left\{ x \in \Sigma : V(x) < \frac{2\lambda}{\mu} N + 2 \right\},$$

which is clearly finite, and in order to analyze the drift condition we distinguish between the following two cases:

- $x \in C$:

$$\mathcal{A}V(x) = \lambda N - \mu V(x) + \mu \leq -\frac{\mu V(x)}{2} + \lambda N + \mu.$$

[1] This follows along the lines of the arguments in Chapters 4 and 5 of [15]. In particular, note that all sets $\{y\} \in \Sigma$ are ν_1-small and $P^1(x, y) > 0$ for all $x, y \in \Sigma$ so that in fact $\mathbf{\Phi}$ is irreducible and strongly aperiodic.

[2] Here and in the next expression with the notation $o(h)$ we mean that f is $o(h)$ if $\lim_{h\to 0} \frac{f(h)}{h} = 0$. In the rest of the text $o(n)$ has the usual meaning.

- $x \in \Sigma \backslash C$:

$$\mathcal{A}V(x) = \lambda N - \mu V(x) + \mu \leq \frac{\mu V(x)}{2} - \mu V(x) = -\frac{\mu V(x)}{2}.$$

Thus, the drift condition holds for $\beta = \mu/2$ and $b = \lambda N + \mu$. ∎

Having shown the existence of an invariant limiting distribution π_n, we now analyze the maximum load on the edges under this distribution.

Theorem 4. *Consider a network with n nodes, and let π_n be the invariant distribution of the induced Markov process under the BDAR* policy with unbounded edge capacity. Under π_n, the maximum number of calls in any edge is bounded whp. by*

$$\frac{2\ln\ln n}{\ln d} + o\left(\frac{\ln\ln n}{\ln d}\right), \qquad \textit{as } n \to \infty.$$

Proof. In order to compute the maximum edge load under the stationary distribution, we start observing the system at some time point and study its transient behavior; we then use the results to deduce the properties of the invariant distribution. In particular, we show that there exists a $T = O\left(n\frac{\ln\ln n}{\ln d}\right)$, such that for any state of the system at time $\tau - T$ that has sufficiently large probability (we will be more precise later), whp at time τ the maximum number of calls on any edge is

$$\frac{2\ln\ln n}{\ln d} + o\left(\frac{\ln\ln n}{\ln d}\right).$$

The high level idea is the following. We partition the time period T into $\frac{\ln\ln n}{\ln d} + o\left(\frac{\ln\ln n}{\ln d}\right)$ periods of length $O(n)$. Roughly, we argue that at the end of the ith period, whp, for each node, the number of incident edges with load greater than i is at most $2\alpha_i$. The α_i's decrease doubly exponentially, so at the end of the last period we will be able to deduce that there are no edges with more than $\frac{\ln\ln n}{\ln d}$ load towards each direction, whp. The challenge is to handle the dependencies, as the number of calls during some period depends on the number of calls of the previous periods. We now proceed with the details.

We first define the sequence of values $\{\alpha_i\}$, which decrease doubly exponentially:

$$\alpha_\kappa = \frac{(n-2)\rho}{\kappa}, \quad \text{where } \kappa = e\rho \cdot \sqrt[d-1]{2\rho \cdot 4^d},$$

$$\alpha_i = \frac{2\rho \cdot 4^d \cdot \alpha_{i-1}^d}{(n-2)^{d-1}} \quad \text{for } i > \kappa \text{ and } \alpha_{i-1} \geq \frac{1}{4} \cdot \sqrt[d]{\frac{25}{\rho}(n-2)^{d-1} \cdot \ln n},$$

$$\alpha_{i^*} = 50\ln n \quad i^* \text{ is the smallest } i \text{ for which } \alpha_{i-1} < \frac{1}{4} \cdot \sqrt[d]{\frac{25}{\rho}(n-2)^{d-1} \cdot \ln n},$$

$$\alpha_{i^*+1} = 10.$$

Solving the recurrence, we get, for $\kappa \leq i < i^*$,

$$\begin{aligned}\alpha_{i+\kappa} &= (2\rho \cdot 4^d)^{(d^i-1)/(d-1)} \cdot \left(\frac{\rho}{\kappa}\right)^{d^i} (n-2) = \frac{1}{\sqrt[d-1]{2\rho \cdot 4^d}} \cdot \left[\frac{\rho \cdot \sqrt[d-1]{2\rho \cdot 4^d}}{\kappa}\right]^{d^i} (n-2) \\ &= \frac{1}{\sqrt[d-1]{2\rho \cdot 4^d}} \cdot \frac{n-2}{e^{d^i}},\end{aligned} \tag{3}$$

and since i^* is the smallest integer satisfying

$$\alpha_{i^*-1} < \frac{1}{4} \cdot \sqrt[d]{\frac{25}{\rho}(n-2)^{d-1} \cdot \ln n},$$

we get after some calculations

$$i^* = \frac{\ln \ln n}{\ln d} + o\left(\frac{\ln \ln n}{\ln d}\right).$$

Next we define $T = n(i^* - \kappa + 3) = O\left(n \frac{\ln \ln n}{\ln d}\right)$ and an increasing sequence of points in time: Let $t_{\kappa-1} = \tau - T$ and for $i \geq \kappa$, $t_i = t_{i-1} + n$, so that the end of the last period, t_{i^*+2}, is the current time τ.

Let $\mathcal{E}$ denote the event "at time $t_{\kappa-1} = \tau - T$ there are at most $(1+\epsilon)N\rho$ calls in the system," for some constant $\epsilon > 0$, and let

$$\mathcal{C}_i = \{\forall v \in V, t \in [t_i, \tau] : L^v_{\geq i}(t) \leq 2\alpha_i\}.$$

We show by induction that for $i = \kappa, \ldots, i^* + 1$

$$\mathbf{Pr}(\overline{\mathcal{C}_i} \mid \mathcal{E}) \leq \frac{2i}{n^2}. \tag{4}$$

Initially we prove the following lemma, which we use throughout the proof. ■

Lemma 1. *Let $\mathcal{A}$ and $\mathcal{B}$ be events such that* $\mathbf{Pr}(\mathcal{B}) \geq 1 - n^{-c}$ *for some constant c, for n large enough. Then for any constant* $\zeta > 0$ *we have*

$$\mathbf{Pr}(\mathcal{A} \mid \mathcal{B}) \leq (1+\zeta)\,\mathbf{Pr}(\mathcal{A}),$$

for sufficiently large n.

Proof. We have

$$\mathbf{Pr}(\mathcal{A} \mid \mathcal{B}) = \frac{\mathbf{Pr}(\mathcal{A}, \mathcal{B})}{\mathbf{Pr}(\mathcal{B})} \leq \frac{\mathbf{Pr}(\mathcal{A})}{\mathbf{Pr}(\mathcal{B})} \leq \frac{1}{1-n^{-c}}\,\mathbf{Pr}(\mathcal{A}) \leq (1+\zeta)\,\mathbf{Pr}(\mathcal{A}).$$

■

Now we examine the base case of Relation 4. Let $\mathcal{C}^v_i$ be the event

$$\mathcal{C}^v_i = \{\forall t \in [t_i, \tau] : L^v_{\geq i}(t) \leq 2\alpha_i\},$$

and $\mathcal{J}^v$ be the event "no more than $2\lambda(n-1)T$ calls are generated with node v as an endpoint during $[\tau - T, \tau]$." We need to bound the probability of $\overline{\mathcal{J}^v}$, so we prove the following lemma.

Lemma 2. *For sufficiently large n, we have*

$$\mathbf{Pr}(\overline{\mathcal{J}^v} \mid \mathcal{E}) < n^{-4}.$$

Proof. Node v has $n-1$ incident links, on each of which new calls are generated according to a Poisson process with rate λ, independently of the other links. Therefore, the number of new calls with v as an endpoint during T steps is distributed according to a Poisson($\lambda(n-1)T$). So by applying a Chernoff bound for the Poisson distribution[3] we get that

$$\begin{aligned}\mathbf{Pr}(\overline{\mathcal{J}^v}) &\le \frac{e^{-\lambda(n-1)T}(e\lambda(n-1)T)^{2\lambda(n-1)T}}{(2\lambda(n-1)T)^{2\lambda(n-1)T}} \\ &= e^{-\lambda(n-1)T+2\lambda(n-1)T+2\lambda(n-1)T\ln(\lambda(n-1)T)-2\lambda(n-1)T\ln(2\lambda(n-1)T)} \\ &= e^{-\lambda(n-1)T(2\ln 2-1)} \\ &< n^{-4},\end{aligned}$$

for sufficiently large n. To complete the proof, we use the fact that the number of new calls during $[\tau - T, \tau]$ is independent of event $\mathcal{E}$. ∎

We now have

$$\begin{aligned}\mathbf{Pr}(\overline{\mathcal{C}_\kappa} \mid \mathcal{E}) &\le n\,\mathbf{Pr}(\overline{\mathcal{C}_\kappa^v} \mid \mathcal{E}) \\ &\le n\,\mathbf{Pr}(\overline{\mathcal{C}_\kappa^v} \mid \mathcal{J}^v, \mathcal{E}) + n\,\mathbf{Pr}(\overline{\mathcal{J}^v} \mid \mathcal{E}).\end{aligned} \tag{5}$$

By Lemma 2, the second term is bounded by $n \cdot n^{-4}$, and we now bound the first term. Conditioning on $\mathcal{J}^v$, we have at most $2\lambda(n-1)T$ new jobs during $[t_{\kappa-1}, \tau]$, say at times $\{\hat{t}_j,\ j = 1, 2, \ldots\}$. Define also $\hat{t}_0 = t_\kappa$. Then

$$\mathbf{Pr}(\overline{\mathcal{C}_\kappa^v} \mid \mathcal{J}^v, \mathcal{E}) \le \sum_{\substack{j=0 \\ \hat{t}_j \ge t_\kappa}}^{2\lambda(n-1)T} \mathbf{Pr}(L^v_{\ge\kappa}(\hat{t}_j) > 2\alpha_\kappa \mid \mathcal{J}^v, \mathcal{E}). \tag{6}$$

Let us compute the number of calls in the system with node v as an endpoint at time $\hat{t}_j$. These calls can be separated to calls that were in the system before time $t_{\kappa-1}$ (let x be their number), and calls that arrived after $t_{\kappa-1}$ (say y).

In order to compute x, we can notice that each of the x calls remains in the system until time $\hat{t}_j$ with probability $e^{-\mu(\hat{t}_j - t_{\kappa-1})}$. Since $\hat{t}_j \ge t_\kappa = t_{\kappa-1} + n$, the probability that a such call survives is bounded by $e^{-n\mu}$. So,

$$\mathbf{Pr}(x > 0 \mid \mathcal{E}) \le (1+\epsilon)N\rho e^{-n\mu} < \frac{1}{n^7},$$

[3] Assume that X is distributed according to a Poisson distribution with rate λ. Then (see, for example, [19, p. 416])

$$\mathbf{Pr}(X \ge i) \le \frac{e^{-\lambda}(e\lambda)^i}{i^i}.$$

and we conclude that conditioning on event $\mathcal{E}$, $x = 0$ with probability at least $1 - n^{-7}$, for sufficiently large n.

In order to bound y, the number of calls arrived after time point $t_{\kappa-1}$, we prove the following lemma.

Lemma 3. *Consider a period Π and a given node v. The number of calls having node v as an endpoint that were generated during Π and are in the system at the end of Π is distributed according to a Poisson distribution with rate bounded by $\rho(n-1)$, independently of $\mathcal{E}$.*

Proof. Let Δ be the duration of the period Π, and let Y be a random variable counting the number of calls that were generated during Π, had v as an endpoint, and are in the system at the end of Π. Node v has $n-1$ incident links on each of which new calls are generated with rate λ, independently of each other. The duration of each call is exponentially distributed with parameter μ. This process is an infinite server Poisson queue [18, p. 18] in which the number of calls at the end of the period is distributed according to a Poisson distribution with rate

$$\lambda(n-1)\Delta p,$$

where

$$p = \int_0^\Delta \frac{e^{-\mu(\Delta - x)}}{\Delta} \mathrm{d}x = \frac{1}{\mu\Delta}\left(1 - e^{-\mu\Delta}\right) \leq \frac{1}{\mu\Delta}.$$

So Y is distributed according to a Poisson distribution with rate at most $\lambda(n-1)/\mu = \rho(n-1)$. Notice also that since Y does not depend on any event prior of Π, the distribution of Y conditioned on $\mathcal{E}$ is still Poisson with the same rate. ∎

By applying this lemma, we have that y is bounded by a Poisson$(\rho(n-1))$. So, from the Chernoff bound, we conclude that $y \leq 2\rho(n-2)$ with probability at least $1 - n^{-7}$, for sufficiently large n.

The probability that at time $\hat{t}_j$ there are more than $2\rho(n-2)$ calls with node v as an endpoint is bounded by

$$\mathbf{Pr}(x > 0 \vee y > 2\rho(n-2) \mid \mathcal{E}),$$

which, using the previous facts, can be bounded by $2n^{-7}$.

Notice now that if node v has fewer than $2\rho(n-2)$ calls at time $\hat{t}_j$, then

$$L^v_{\geq\kappa}(\hat{t}_j) \leq \frac{2\rho(n-2)}{\kappa} = 2\alpha_\kappa.$$

Hence, for all $\hat{t}_j \geq t_\kappa$ we have

$$\mathbf{Pr}(L^v_{\geq\kappa}(\hat{t}_j) > 2\alpha_\kappa \mid \mathcal{E}) \leq 2n^{-7},$$

and by making use of Lemma 1, we get

$$\mathbf{Pr}(L^v_{\geq\kappa}(\hat{t}_j) > 2\alpha_\kappa \mid \mathcal{J}^v, \mathcal{E}) \leq 2 \cdot 2n^{-7} = 4n^{-7}. \tag{7}$$

Combining Relations (5), (6), (7), Lemma 2, and the fact that $T = O(n^2)$, we get that

$$\mathbf{Pr}(\overline{\mathcal{C}_\kappa} \mid \mathcal{E}) \leq n \cdot (2\lambda(n-1)+1) \cdot n^2 \cdot 4n^{-7} + n \cdot n^{-4} \leq n^{-2},$$

for large enough n, which completes the base case ($i = \kappa$) of Relation (4).

For the induction step we assume that

$$\mathbf{Pr}(\overline{\mathcal{C}_{i-1}} \mid \mathcal{E}) \leq \frac{2(i-1)}{n^2}. \tag{8}$$

Assume now that at time t a new call enters the system. Then the call is routed through an edge with (new) load greater or equal to i if in all the d alternative paths at least one of the two edges had load at least $i-1$. More concretely, let $\mathcal{G}$ denote the event "a new call is generated at time t with v as an endpoint," and let u be the other endpoint and $(w_j, j = 1, \ldots, d)$ be the intermediate nodes of the queried alternative paths.

We then have

$$\begin{aligned}
&\mathbf{Pr}(M^v_{\geq i}(t) > M^v_{\geq i}(t-) \mid \Phi(t-), \mathcal{G}) \\
&\qquad \leq \mathbf{Pr}(M^v_{\geq i}(t) > M^v_{\geq i}(t-) \vee M^u_{\geq i}(t) > M^u_{\geq i}(t-) \mid \Phi(t-), \mathcal{G}) \\
&\qquad \leq \mathbf{Pr}(\forall j \in \{1, \ldots, d\} : \ell_{(v,w_j)}(t-) \geq i-1 \vee \ell_{(u,w_j)}(t-) \geq i-1 \mid \Phi(t-), \mathcal{G}) \\
&\qquad \leq \left(\frac{L^v_{\geq i-1}(t-) + L^u_{\geq i-1}(t-)}{n-2} \right)^d,
\end{aligned}$$

therefore,

$$\mathbf{Pr}(M^v_{\geq i}(t) > M^v_{\geq i}(t-) \mid \mathcal{E}, \mathcal{G}, \forall z \in V : L^z_{\geq i-1}(t-) \leq 2\alpha_{i-1}) \leq \left(\frac{2 \cdot 2\alpha_{i-1}}{n-2} \right)^d \stackrel{\triangle}{=} q_i. \tag{9}$$

Notice that for $i = \kappa+1, \ldots, i^*$ we have

$$q_i \leq \frac{\alpha_i}{2\rho(n-2)}. \tag{10}$$

We now define

$$\mathcal{F}_i = \{\forall v \in V : M^v_{\geq i}(t_i) < \alpha_i\}$$

and prove Lemmata 4 and 6, that allow us to conclude that $\mathbf{Pr}(\overline{\mathcal{C}_i} \mid \mathcal{E}) \leq 2i/n^2$, and establish Relation (4).

Lemma 4. *Under the inductive hypothesis*

$$\mathbf{Pr}(\overline{\mathcal{F}_i} \mid \mathcal{C}_{i-1}, \mathcal{E}) \leq n^{-2}.$$

Proof. First we apply Lemma 3 for the interval $\Pi = [t_{\kappa-1}, t_{i-1}]$, and we deduce that the number of calls with v as an endpoint that were generated during Π and remained until time t_{i-1} follows a Poisson distribution with mean bounded by $\rho(n-1)$. Hence, with a Chernoff bound, we get that with probability at least $1-n^{-3}$ there are at most $2\rho(n-1)$ such calls. If we condition on event $\mathcal{E}$, then the total number of calls in the system at time t_{i-1} with node v as an endpoint is at most

$$(1+\epsilon)N\rho + 2\rho(n-1)$$

with probability at least $1 - n^3$. The probability that each of these calls stays in the system until time t_i is bounded by $e^{-n\mu}$ (recall that $t_i - t_{i-1} = n$), so the probability, conditioned on the event $\mathcal{E}$, that some of the calls that were in the system up to time t_{i-1} and had v as an endpoint, stays in the system until time t_i is bounded by

$$n^{-3} + [(1+\epsilon)N\rho + 2\rho(n-1)]e^{-n\mu} < 2n^{-3}$$

for sufficiently large n. By applying Lemma 1 and making use of the induction hypothesis [Eq. (8)] we deduce that the probability that some of those calls stay in the system conditioned on the events $\mathcal{C}_{i-1}$ and $\mathcal{E}$ is bounded by $4n^{-3}$. To analyze the number of the remaining calls that were created during the period $[t_{i-1}, t_i]$, we make use of Lemma 5 which completes the proof of this one. ∎

Lemma 5. *Consider a period Π and a given node v. Conditioning on $\mathcal{C}_{i-1}$ and $\mathcal{E}$, the number of new calls that increased $M^v_{\geq i}$ when they were generated, and remained until the end of Π is less than α_i, with probability at least $1 - n^{-7}$.*

Proof. Let Y be the number of calls that were generated during Π, had v as an endpoint and are in the system at the end of Π. By applying Lemma 3 we get that conditioned on $\mathcal{E}$, Y follows a Poisson distribution with rate bounded by $\rho(n-1)$.

Let now Z be the number of calls in the system at the end of Π whose arrival resulted in the increase of $M^v_{\geq i}$. Denote with $\mathcal{H}_k$ the event $\{Y = k\}$ and let $\{\tilde{t}_j\}_{j=1}^k$ be the time of the arrival of the jth call that exists in the system at the end of Π. We can then write

$$\mathbf{Pr}(Z > r \mid \mathcal{E}, \mathcal{C}_{i-1}) = \sum_k \mathbf{Pr}(Z > r \mid \mathcal{E}, \mathcal{C}_{i-1}, \mathcal{H}_k) \cdot \mathbf{Pr}(\mathcal{H}_k \mid \mathcal{E}, \mathcal{C}_{i-1}).$$

We now fix k and we consider the random variables $\{Z_j\}_{j=1}^k$, where

$$\begin{aligned} Z_j = 1 \qquad & \text{if} \quad M^v_{\geq i}(\tilde{t}_j) > M^v_{\geq i}(\tilde{t}_j -) \\ & \text{and} \quad \forall z \in V : L^z_{\geq i-1}(\tilde{t}_j -) \leq 2\alpha_{i-1}. \end{aligned}$$

From Relation (9) we get that

$$\mathbf{Pr}(Z_j = 1 \mid \mathcal{E}) \leq q_i,$$

so, since (induction hypothesis (4)) $\mathbf{Pr}(\mathcal{C}_{i-1} \mid \mathcal{E}) \geq 1 - 2(i-1)/n^2$, we can apply Lemma 1 and get

$$\mathbf{Pr}(Z_j = 1 \mid \mathcal{E}, \mathcal{C}_{i-1}) \leq (1+\zeta)q_i, \tag{11}$$

for some constant ζ (say 0.05), independently of all the previous Z_j. Notice now that conditioning on events $\mathcal{C}_{i-1}$, and $\mathcal{H}_k$, we have

$$Z = \sum_{j=1}^k Z_j.$$

Hence

$$\mathbf{Pr}(Z > r \mid \mathcal{E}, \mathcal{C}_{i-1}) = \sum_k \mathbf{Pr}\left(\sum_{j=1}^{k} Z_j > r \,\middle|\, \mathcal{E}, \mathcal{C}_{i-1}, \mathcal{H}_k \right) \cdot \mathbf{Pr}(\mathcal{H}_k \mid \mathcal{E}, \mathcal{C}_{i-1}).$$

Again by Lemma 1, we get

$$\mathbf{Pr}(\mathcal{H}_k \mid \mathcal{E}, \mathcal{C}_{i-1}) \le 2\mathbf{Pr}(\mathcal{H}_k \mid \mathcal{E}).$$

So by the fact that the distribution of Y conditioned on $\mathcal{E}$ is Poisson with rate at most $\rho(n-1)$, and by relation (11), we can finally conclude that

$$\begin{aligned}\mathbf{Pr}(Z > r \mid \mathcal{E}, \mathcal{C}_{i-1}) &\le 2 \sum_k \mathbf{Pr}(\text{Binomial}(k, (1+\zeta)q_i) > r) \cdot \mathbf{Pr}(\text{Poisson}(\rho(n-1)) = k) \\ &\le 2\mathbf{Pr}(\text{Poisson}((1+\zeta)\rho q_i(n-1)) > r).\end{aligned}$$

We now distinguish the following two cases:

Case 1: For $i \le i^*$, by using Eq. (10) we get that $(1+\zeta)\rho q_i(n-1) \le 1.1\alpha_i/2$ for $\zeta = 0.05$, and by applying the Chernoff bound, we get that the probability that the number of calls is higher than α_i is bounded by

$$2\frac{e^{-(1.1\alpha_i)/2}\left(e\frac{1.1\alpha_i}{2}\right)^{\alpha_i}}{\alpha_i^{\alpha_i}} \le 2e^{-0.147\alpha_i}.$$

For $i < i^*$ we have from the definition of α_i

$$\begin{aligned}2e^{-0.147\alpha_i} &= 2e^{-0.147\frac{2\rho\cdot 4^d \alpha_{i-1}^d}{(n-2)^{d-1}}} \\ &\le 2e^{-0.147\frac{2\rho\cdot 4^d \frac{1}{4^d}\frac{25}{\rho}(n-2)^{d-1}\ln n}{(n-2)^{d-1}}} \\ &= 2e^{-0.147\cdot 50\ln n} \\ &= o\left(\frac{1}{n^7}\right),\end{aligned}$$

while for $i = i^*$ we get

$$\begin{aligned}e^{-0.147\alpha_i} &= 2e^{-0.147\cdot 50\ln n} \\ &= o\left(\frac{1}{n^7}\right).\end{aligned}$$

Case 2: For $i = i^* + 1$, using Eq. (9) we get that the parameter of the Poisson distribution is

$$(1+\zeta)\rho q_i(n-1) \le (1+\zeta)\frac{4^d \cdot \alpha_{i-1}^d}{(n-2)^d}\rho(n-1) = (1+\zeta)\frac{(4\cdot 50\ln n)^d}{(n-2)^d}\rho(n-1),$$

and with the Chernoff bound we get that the probability that the number of calls is higher than $\alpha_{i^*+1} = 10$ is $o(1/n^7)$. ∎

Lemma 6. *Under the inductive hypothesis*

$$\mathbf{Pr}(\overline{\mathcal{C}_i} \mid \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) \leq n^{-2}.$$

Proof. First we compute

$$\begin{aligned}\mathbf{Pr}(\mathcal{F}_i, \mathcal{C}_{i-1} \mid \mathcal{E}) &= \mathbf{Pr}(\mathcal{C}_{i-1} \mid \mathcal{E}) \cdot \mathbf{Pr}(\mathcal{F}_i \mid \mathcal{C}_{i-1}, \mathcal{E}) \\ &\geq \left(1 - \frac{i-1}{n^2}\right) \cdot \left(1 - \frac{1}{n^2}\right),\end{aligned}$$

by Relation (8) and Lemma 4, so

$$\mathbf{Pr}(\mathcal{F}_i, \mathcal{C}_{i-1} \mid \mathcal{E}) \geq 1 - \frac{1}{n}.$$

So, by Lemma 1 we get

$$\mathbf{Pr}(\overline{\mathcal{J}^v} \mid \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) \leq 2\,\mathbf{Pr}(\overline{\mathcal{J}^v} \mid \mathcal{E})$$

and finally, by using Lemma 2, we conclude

$$\mathbf{Pr}(\overline{\mathcal{J}^v} \mid \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) \leq 2\,n^{-4}. \tag{12}$$

Hence, we can get

$$\begin{aligned}\mathbf{Pr}(\overline{\mathcal{C}_i} \mid \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) &\leq n \cdot \mathbf{Pr}(\overline{\mathcal{C}_i^v} \mid \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) \\ &\leq n \cdot \mathbf{Pr}(\overline{\mathcal{C}_i^v} \mid \mathcal{J}^v, \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) + n \cdot \mathbf{Pr}(\overline{\mathcal{J}^v} \mid \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}).\end{aligned} \tag{13}$$

We have a bound for the second term, so we want to bound the first one. For that, we write (recall that $\{\hat{t}_j\}$ are the times of the arrivals of the new calls with node v as an endpoint)

$$\begin{aligned}\mathbf{Pr}(\overline{\mathcal{C}_i^v} \mid \mathcal{J}^v, \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) &\leq \mathbf{Pr}(\exists \tilde{t} \in [t_i, \tau] : L^v_{\geq i}(\tilde{t}) > 2\alpha_i \mid \mathcal{J}^v, \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) \\ &\leq \mathbf{Pr}(\exists \tilde{t} \in [t_i, \tau] : M^v_{\geq i}(\tilde{t}) > 2\alpha_i \mid \mathcal{J}^v, \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) \\ &\leq \sum_{\substack{j=1 \\ \hat{t}_j \geq t_i}}^{2\lambda(n-1)T} \mathbf{Pr}(M^v_{\geq i}(\hat{t}_j) > 2\alpha_i \mid \mathcal{J}^v, \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E})\end{aligned} \tag{14}$$

Conditioning on event $\mathcal{F}_i$, we have $M^v_{\geq i}(\hat{t}_j) > 2\alpha_i$ only if $M^v_{\geq i}$ increased by at least α_i during the interval $[t_i, \hat{t}_j]$. Therefore, by applying Lemmata 1, 4, and 5, we get

$$\mathbf{Pr}(M^v_{\geq i}(\hat{t}_j) > 2\alpha_i \mid \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) < \frac{2}{n^7}.$$

We combine this result with Relation (12) and Lemma 1, and we have

$$\mathbf{Pr}(M^v_{\geq i}(\hat{t}_j) > 2\alpha_i \mid \mathcal{J}^v, \mathcal{F}_i, \mathcal{C}_{i-1}, \mathcal{E}) < \frac{4}{n^7}. \tag{15}$$

If we combine relations (13), (14), and (15), we get the result. ∎

Having proven Lemmata 4 and 6 we can now show that $\mathbf{Pr}(\overline{\mathcal{C}_i} \mid \mathcal{E}) \leq 2i/n^2$:

$$\begin{aligned}
\mathbf{Pr}(\overline{\mathcal{C}_i} \mid \mathcal{E}) &= \mathbf{Pr}(\overline{\mathcal{C}_i} \mid \mathcal{C}_{i-1}, \mathcal{E}) \cdot \mathbf{Pr}(\mathcal{C}_{i-1}, \mathcal{E}) + \mathbf{Pr}(\overline{\mathcal{C}_i} \mid \overline{\mathcal{C}_{i-1}}, \mathcal{E}) \cdot \mathbf{Pr}(\overline{\mathcal{C}_{i-1}}, \mathcal{E}) \\
&\leq \mathbf{Pr}(\overline{\mathcal{C}_i} \mid \mathcal{C}_{i-1}, \mathcal{E}) + \frac{2(i-1)}{n^2} \\
&= \mathbf{Pr}(\overline{\mathcal{C}_i} \mid \mathcal{C}_{i-1}, \mathcal{F}_i, \mathcal{E}) \cdot \mathbf{Pr}(\mathcal{F}_i \mid \mathcal{C}_{i-1}, \mathcal{E}) \\
&\quad + \mathbf{Pr}(\overline{\mathcal{C}_i} \mid \mathcal{C}_{i-1}, \overline{\mathcal{F}_i}, \mathcal{E}) \cdot \mathbf{Pr}(\overline{\mathcal{F}_i} \mid \mathcal{C}_{i-1}, \mathcal{E}) + \frac{2(i-1)}{n^2} \\
&\leq \frac{1}{n^2} + \frac{1}{n^2} + \frac{2(i-1)}{n^2} \\
&= \frac{2i}{n^2}.
\end{aligned}$$

We have therefore shown that the event $\mathcal{C}_{i^*+1}$ holds whp, which implies that for every node v, after the $(i^* + 1)$th period, there will be no more than $2\alpha_{i^*+1} = 20$ incident edges with load more than $i^* + 1$. We will now bound the probability that in the next interval ($[t_{i^*+1}, t_{i^*+2}]$, the last interval of T) there will be an incident edge of v with load more than $i^* + 3$, conditioning on the event $\mathcal{C}_{i^*+1}$. For this to happen, we must have at least 2 new calls to be routed using one of the 20 highly loaded edges. The probability that two specific new calls use these edges is at most

$$\left(\frac{20+20}{n-2}\right)^{2d} = O\left(\frac{1}{n^4}\right), \tag{16}$$

since $d \geq 2$. The expected number of calls with v as an endpoint is $\lambda(n-1)n$, since $(n-1)$ links are connected to v in each of which new calls are generated with rate λ, while the total length of the interval is n. This implies that whp. there will be $O(n^2)$ new calls in the whole period. By combining this fact with Eq. (16), applying Lemma 1, and summing for all the nodes we conclude that at the end of period T there will be no edges with load more than $i^* + 3$ whp.

We now consider the stationary distribution π_n, and show that under it

$$\mathbf{Pr}\left(\ell_{\max} \leq \frac{\ln \ln n}{\ln d} + o\left(\frac{\ln \ln n}{\ln d}\right)\right) = 1 - o\left(\frac{1}{n}\right),$$

where

$$\ell_{\max} = \max_{e=(u,v)\in E} \max\{\ell_{e,u}, \ell_{e,v}\}$$

denotes the maximum number of calls on any edge, in the stationary regime ($\ell_{e,u}$ is the number of calls with u as an endpoint routed through edge e in the stationary regime). Recall that $\Phi(t)$ is the state of the system at time t, and consider the following partitioning of the state space, Σ, of the underlying Markov process:

- $S_1 = \left\{x : V(x) \leq (1+\epsilon)N\rho,\ \ell_{\max} \leq \dfrac{\ln \ln n}{\ln d} + o\left(\dfrac{\ln \ln n}{\ln d}\right)\right\}$,

  that is, states in which the total number of calls in the system is at most $(1+\epsilon)N\rho$, and the maximum load is at most $\frac{\ln \ln n}{\ln d} + o\left(\frac{\ln \ln n}{\ln d}\right)$.

- $S_2 = \left\{x : V(x) \leq (1+\epsilon)N\rho,\ \ell_{\max} > \frac{\ln\ln n}{\ln d} + \Omega\left(\frac{\ln\ln n}{\ln d}\right)\right\}$,

  that is, states in which the total number of calls in the system is at most $(1+\epsilon)N\rho$, and the maximum load is higher than $\frac{\ln\ln n}{\ln d} + \Omega\left(\frac{\ln\ln n}{\ln d}\right)$.
- $S_3 = \{x : V(x) > (1+\epsilon)N\rho\}$,

  that is, states in which the total number of calls in the system is higher than $(1+\epsilon)N\rho$.

We have shown that

$$\mathbf{Pr}(\Phi(\tau) \in S_2 \mid \Phi(\tau - T) \in S_1 \cup S_2) = o\left(\frac{1}{n}\right),$$

and we can easily show that

$$\mathbf{Pr}(\Phi(\tau) \in S_3 \mid \Phi(\tau - T) \in S_1 \cup S_2) = o\left(\frac{1}{n}\right).$$

Moreover, in the stationary distribution the number of calls in the system has a Poisson distribution with parameter $N\rho$. Hence by using the Chernoff bound

$$\sum_{i \in S_3} (\pi_n)_i = o\left(\frac{1}{n}\right).$$

Then we have

$$\sum_{i \in S_2 \cup S_3} (\pi_n)_i = \sum_{i \in S_2} (\pi_n)_i + \sum_{i \in S_3} (\pi_n)_i.$$

The second term is $o(1/n)$, while for the first one

$$\begin{aligned}
\sum_{i \in S_2} (\pi_n)_i &= \sum_j \mathbf{Pr}(\Phi(\tau) \in S_2 \mid \Phi(\tau - T) = j) \cdot (\pi_n)_j \\
&= \sum_{j \in S_1 \cup S_2} \mathbf{Pr}(\Phi(\tau) \in S_2 \mid \Phi(\tau - T) = j) \cdot (\pi_n)_j \\
&\quad + \sum_{j \in S_3} \mathbf{Pr}(\Phi(\tau) \in S_2 \mid \Phi(\tau - T) = j) \cdot (\pi_n)_j \\
&= \sum_{j \in S_1 \cup S_2} (\pi_n)_j \cdot o\left(\frac{1}{n}\right) + o\left(\frac{1}{n}\right) = o\left(\frac{1}{n}\right).
\end{aligned}$$

Therefore,

$$\sum_{i \in S_2 \cup S_3} (\pi_n)_i = o\left(\frac{1}{n}\right),$$

which implies that

$$\sum_{i \in S_1} (\pi_n)_i = 1 - o\left(\frac{1}{n}\right)$$

and completes the proof of Theorem 4. ■

2.2. Bounded Capacities

In this section we use the analysis of the BDAR* algorithm for unbounded capacities to compute the bandwidth requirement B ($< \infty$) that ensures that a new call is not lost whp.

Theorem 5. *Assume that all the edges have capacity* $3B$ *circuits, which can be a function of* n. *Then, if we perform the BDAR* policy, capacity*

$$B = \frac{\ln \ln n}{\ln d} + o\left(\frac{\ln \ln n}{\ln d}\right), \qquad \text{as } n \to \infty$$

ensures that under the stationary distribution a new call is not lost whp.

Proof. The result for finite B follows from the proof of Theorem 4, which concerns unbounded capacity. Since the Markov process is finite and aperiodic there exists a stationary distribution. Moreover, the analysis for the unbounded case still holds for finite B as long as $B \le i^* + 1$.

A new call between nodes u and v will be rejected if in all the d choices, either the edge incident to node u is used in routing $i^* + 1 = \ln \ln n / \ln d + o(\ln \ln n / \ln d)$ calls with node u as an endpoint, or the edge incident to node v is used in routing $i^* + 1$ calls with node v as an endpoint. With probability at least $1 - o(n^{-1})$, for each node, the number of incident edges with load at least $i^* + 1$ is at most $2\alpha_{i^*+1}$. Therefore, the probability for a call to be rejected is no more than

$$o\left(\frac{1}{n}\right) + \left(\frac{2\alpha_{i^*+1} + 2\alpha_{i^*+1}}{n-2}\right)^d = o\left(\frac{1}{n}\right)$$

since $\alpha_{i^*+1} = 10$. ∎

3. LOWER BOUND ON THE PERFORMANCE OF THE DAR ALGORITHM

To demonstrate the advantage of the balanced-allocation method, we prove here a lower bound on the maximum channel load when requests are routed using the DAR algorithm. This bound shows an exponential gap between the capacity required by the balanced-allocation algorithm and the capacity required by the standard DAR algorithm for the same stream of inputs.

Recall from Section 1.1 that we consider a complete network of n nodes and $N = \binom{n}{2}$ edges. Requests for connections between a given pair arrive according to a Poisson process with rate λ, the duration of a connection has an exponential distribution with expectation $1/\mu$. Edges have capacities of $3B$ circuits, B are used for direct connections, and the remaining $2B$ are used for alternative routes with the capacity reserved for alternative routes furthermore split into two, so that B circuits are used for alternate paths with one node of the edge as an endpoint and B for calls with the other node as an endpoint.

Theorem 6. *Assume that all the edges have capacity* $3B$ *circuits, which can be a function of* n. *Then, if we perform the DAR policy, capacity*

$$B = \Omega\left(\sqrt{\frac{\ln n}{d \ln \ln n}}\right), \qquad \text{as } n \to \infty$$

is necessary to ensure that under the stationary distribution a new call is not lost whp.

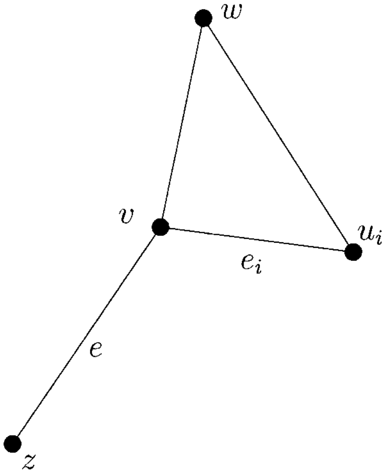

Fig. 1. A call is generated at by edge e at time t.

Proof. We will compute a lower bound on the probability $P = P(B)$, that a request arriving at an arbitrary time t is rejected.

We consider first the probability P_1 that the new call is not routed through the direct link. The process of routing calls through the direct link is an $M/M/B/B$ loss system (Poisson arrival, exponential service time, B servers—corresponding to the B direct links, up to B customers in the system—corresponding to up to B calls that can be routed through the direct links). Applying Erlang's loss formula (e.g., [9]),

$$P_1 = \frac{\left(\frac{\lambda}{\mu}\right)^B}{B!}\left(\sum_{i=0}^{B}\frac{\left(\frac{\lambda}{\mu}\right)^B}{i!}\right)^{-1} \geq e^{-\lambda/\mu}\frac{\left(\frac{\lambda}{\mu}\right)^B}{B!}. \tag{17}$$

Since the arrival is Poisson, it is independent of the state of the queue at the time of arrival, hence the probability that a given pair (v, w) had a request during interval $\Pi = [t-1, t]$ that could not be routed by the direct link is

$$P_{\text{alternate}} = (1 - e^{-\lambda})P_1.$$

Next we lower bound the probability P_2 that a request, generated at time t that failed to use the direct link $e = (v, z)$, fails also to be routed by an alternative path (i.e., all the d attempts to find a nonsaturated alternative path do not succeed). In fact, we will restrict our discussion to the probability that in each of these d routes the first edge (v, u_i) on the alternate route was saturated for alternate paths with endpoint v (Fig. 1).

In order to estimate the probability P_2, we compute a lower bound for the probability $P(e_i, t)$, that an arbitrary edge $e_i = (v, u_i)$ was carrying, at time t, B alternate paths with endpoint v (and thus blocked for any other alternate path starting at v). For this we study the evolution of the system during period $\Pi = [t-1, t]$. We will lower bound the probability $P(e_i, t)$ by the probability that at some point during the interval Π the edge carried B alternate paths with endpoint v, and that none of these paths terminated during this interval.

The second requirement is easy to evaluate. Since the calls have exponential duration with parameter μ, every call that is on edge e_i at time $t-1$, or that is created during Π, will stay in the system until time t with probability at least $e^{-\mu}$, and all the calls do not terminate in that interval with probability at least $e^{-\mu B}$.

Let $\mathcal{C}_i$ be the event "during the interval Π, B different pairs $(v, w_1), \ldots, (v, w_B)$ try to use edge $e_i = (v, u_i)$ as a first choice for alternate path, and for each of these pairs the edge (u_i, w_j) (the second edge in the alternate path) was not blocked." Then,

$$P(e_i, t) \geq \mathbf{Pr}(\mathcal{C}_i) e^{-\mu B}.$$

The difficulty in computing $\mathbf{Pr}(\mathcal{C}_i)$ is bounding the probability that the second edge on the alternate path is not blocked. The following lemma simplifies this computation. ▪

Lemma 7. *Let $\mathcal{D}$ be the event "there is a vertex $u \neq v$ that during the interval Π was the center node for more than $c_1 d\left(\frac{\lambda}{\mu} + \lambda\right)(n-1)$ alternate paths with no endpoint in v." Then,*

$$\mathbf{Pr}(\mathcal{D}) \leq e^{-c_2 n},$$

for some constants $c_1, c_2 > 0$.

Proof. There are $\binom{n-1}{2}$ possible pairs of vertices not containing v. For each pair the number of active calls at time $t-1$ is bounded by a Poisson random variable with parameter λ/μ. The number of new calls between a given pair during the interval is bounded by Poisson random variable with parameter λ.

Fix a vertex u. The probability that a given call uses u as a center vertex in an alternate path is bounded by $d/(n-2)$, independently of other calls. Thus, the number of alternating paths through u is stochastically dominated by a Poisson distribution with parameter $\lambda\left(1 + \frac{1}{\mu}\right) d \frac{n-1}{2}$. Applying the Chernoff bound for u and summing over all $n-1$ vertices gives the lemma. ▪

There can be no more than B alternate paths with endpoint v that use a vertex w as a center node. Thus, conditioning on the event $\overline{\mathcal{D}}$, no more than $c_1 d\left(\frac{\lambda}{\mu} + \lambda\right)(n-1) + B$ alternate paths use any vertex $w \neq v$ during the interval Π, and thus, during any time in that interval no more than $\frac{1}{B}\left(c_1 d\left(\frac{\lambda}{\mu} + \lambda\right)(n-1) + B\right)$ edges adjacent to w are blocked for alternating paths using w as a center node.

Focusing back on the edge $e_i = (v, u_i)$, there is a set W_i of vertices such that the edge from u_i to $w \in W_i$ is not blocked for an alternate path with endpoints v and $w \in W_i$ throughout the interval Π. Conditioned on $\overline{\mathcal{D}}$, we have $|W_i| \geq \alpha n$ for some constant $\alpha > 0$.

We can compute

$$\begin{aligned}\mathbf{Pr}(\mathcal{C}_i \mid \overline{\mathcal{D}}) &\geq \binom{\alpha n}{B} \left(P_{\text{alternate}} \cdot \frac{1}{n-2}\right)^B \left(1 - P_{\text{alternate}} \cdot \frac{1}{n-2}\right)^{\alpha n - B} \\ &= e^{-O(B^2 \ln B - B^2 \ln(\lambda/\mu))}. \qquad (18)\end{aligned}$$

The above follows from the fact that there are at least αn edges $(v, w), w \in W_i$, that can create a call during Π with probability $P_{\text{alternate}}$, and select as a first choice for alternate path the path $v - u_i - w$. Note that in the computation we consider no more than one communication request for each pair of vertices $(v, w), w \in W_i$, in order to avoid further dependencies.

Consider now a request that arrives at time t with endpoint v. The probability that the direct link for that request is blocked is P_1.

For simplicity, label the d alternative paths that the call generated at time t (between nodes v and z) as $v - u_i - z$, $i = 1, 2, \dots, d$, and let $\mathcal{E}_i$ be the event "the ith alternative path $(v - u_i - z)$ is blocked." We want to lower bound the probability $P_2 = \mathbf{Pr}(\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_d)$ that the request generated at time t that failed to use the direct link, fails to use all the d alternate paths. Then

$$\begin{aligned} P_2 &\geq \mathbf{Pr}(\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_d) \cdot e^{-d\mu B} \\ &\geq \mathbf{Pr}(\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_d \mid \overline{\mathcal{D}}) \cdot \mathbf{Pr}(\overline{\mathcal{D}}) \cdot e^{-d\mu B} \\ &\geq (1 - e^{-c_2 n}) \cdot e^{-d\mu B} \cdot \prod_{j=1}^{d} \mathbf{Pr}(\mathcal{C}_j \mid \overline{\mathcal{D}}, \mathcal{C}_1, \dots, \mathcal{C}_{j-1}). \end{aligned}$$

Let us try to compute $\mathbf{Pr}(\mathcal{C}_j \mid \overline{\mathcal{D}}, \mathcal{C}_1, \dots, \mathcal{C}_{j-1})$. Let

$$U_i = \{w \in W_i : v - u_i - w \text{ became an active alternate path during } \Pi\}$$

and

$$W_i = W_{i-1} \backslash U_{i-1} = W_1 \Big\backslash \bigcup_{j=1}^{i-1} U_j.$$

Notice that if the calls $\underline{(v - u_i - w)}$ do not terminate during Π, we have $|U_i| = B$, so as long as $dB = o(n)$, conditioned on $\overline{\mathcal{D}}$, there exists a constant α such that $|W_i| \geq \alpha n$, for all $i = 1, \dots, d$. We can repeat the calculation of (18) and get that

$$\mathbf{Pr}(\mathcal{C}_j \mid \overline{\mathcal{D}}, \mathcal{C}_1, \dots, \mathcal{C}_{j-1}) = e^{-O(B^2 \ln B - B^2 \ln(\lambda/\mu))},$$

since a call in W_i is generated, fails to use a direct route, and uses the alternate path $v - u_i - z$, independently of events $\mathcal{C}_1, \dots, \mathcal{C}_{i-1}$. So, finally, we get that

$$P_2 = e^{-O(dB^2 \ln B - dB^2 \ln(\lambda/\mu))}.$$

Putting everything together we conclude that the probability that the call generated at time t is rejected is at least

$$P_1 \cdot P_2 \geq e^{-O(dB^2 \ln B - dB^2 \ln(\lambda/\mu))}.$$

Therefore, in order to guarantee that a new call is not lost whp, the bandwidth must be at least

$$B = \Omega\left(\sqrt{\frac{\ln n}{d \ln \ln n}}\right). \qquad \blacksquare$$

REFERENCES

[1] G. R. Ash, R. H. Cardwell, and R. P. Murray, Design and optimization of networks with dynamic routing, Bell Syst Tech J 60(8) (1981), 1787–1820.

[2] Y. Azar, A. Broder, A. Karlin, and E. Upfal, Balanced allocations, Proc 26th ACM Symp Theory of Computing (STOC ’94), ACM Press, Montreal, Quebec, Canada, 1994, pp. 593–602.

[3] Y. Azar, A. Z. Broder, A. R. Karlin, and E. Upfal, Balanced allocations, SIAM J Comput 29(1) (2002), 180–200.

[4] A. Z. Broder, A. Frieze, C. Lund, S. Phillips, and N. Reingold, Balanced allocations for tree-like inputs, Inform Process Lett 55(6) (1995), 329–332.

[5] D. Down, S. P. Meyn, and R. Tweedie, Exponential and uniform ergodicity of Markov processes, Ann Probab 23(4) (1996), 1671–1691.

[6] R. J. Gibbens, P. J. Hunt, and F. P. Kelly, “Bistability in communication networks,” Disorder in physical systems, Eds., G. R. Grimmet and D. J. A. Welsh, 113–128. Oxford University Press, New York, 1990.

[7] R. J. Gibbens, F. P. Kelly, and P. B. Key, “Dynamic alternative routing,” Routing in Communications Networks, Ed., M. E. Steenstrup, 13–47. Prentice Hall, Englewood Cliffs, NJ, 1995.

[8] P. J. Hunt and C. N. Laws, Asymptotically optimal loss network control, Math Oper Res 18(4) (1993), 880–900.

[9] F. P. Kelly, Loss networks, Ann Appl Probab 1(3) (1991), 319–378.

[10] M. J. Łuczak, C. McDiarmid, and E. Upfal, On-line routing of random calls in networks, Probab Theory Related Fields 125 (2003), 457–482.

[11] M. J. Łuczak and E. Upfal, Reducing network congestion and blocking probability through balanced allocation, IEEE Symp Foundations of Computer Science, 1999, pp. 587–595.

[12] J. Martin and Y. Suhov, Fast Jackson networks, Ann Appl Probab 9(3) (1999), 854–870.

[13] S. P. Meyn and R. Tweedie, Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes, Adv Appl Probab 25 (1993), 518–548.

[14] S. P. Meyn and R. Tweedie, A survey of Foster-Lyapunov conditions for general state space Markov processes, Proc Workshop Stochastic Stability and Stochastic Stabilization, Metz, France, June 1993, Springer, New York, 1994.

[15] S. P. Meyn and R. L. Tweedie, Markov chains and stochastic stability, Communications and Control Engineering Seires, Springer, New York, 1993.

[16] M. Mitzenmacher, The power of two choices in randomized load balancing, Ph.D. Thesis, University of California, Berkeley, August 1996.

[17] M. Mitzenmacher, On the analysis of randomized load balancing schemes, Proc 9th Annu ACM Symp Parallel Algorithms and Architectures (SPAA ’97), ACM Press, Newport, RI, 22–25 June 1997, pp. 292–301.

[18] S. M. Ross, Applied probability models with optimization applications, Dover Publications, New York, 1970.

[19] S. M. Ross, A first course in probability, 5th edition, Macmillan, London, 1998.

[20] Y. Suhov and N. Vvedenskaya, Fast jackson networks with dynamic routing, Probl Inf Transm 38(2) (2002), 136–153.

[21] N. Vvedenskaya, R. Dobrushin, and F. Karpelevich, A queueing system with a choice of the shorter of two queues—an asymptotic approach, Problemy Peredachi Informatsii 32(1) (1996), 20–34.