Local Routing in Multihop Networks

Dobri Atanassov Batovski

Faculty of Science and Technology, Assumption University, Bangkok, Thailand
<dbatovski@au.edu>

Abstract

The overlapping of multihop paths in decentralized networks may cause congestion at certain nodes and underutilization of other neighbor nodes. The choice of a given path for the delivery of packets from source to destination makes other connected nodes, located one hop away from the said path, prospective candidates for local rerouting in case of substantial increase of the traffic rates. Two-hop local rerouting is considered for general queuing with the use of the method of decomposition for non-product networks. Rate-delay equalization of packet flows among two concurrent local paths is used for splitting a packet flow for a given set of input parameters. The analytic model is solved numerically for both M/M/1 and GI/G/1 queues at the individual nodes. Each node has a single output channel having different service rates for packet delivery to different neighbor nodes .

Keywords: *Rate-delay equalization, local routing, two-hop path, general queuing.*

Introduction

The local routing in decentralized networks becomes an important factor in the establishment of reliable fault-tolerant paths for the delivery of packets over multiple hops. A reason for the occurrence of packet drop is the increased utilization of some nodes which handle more traffic while there are similar local paths along a chosen path, which can be used to share the traffic load. Such local paths exist in variety of topological configurations. In well connected network configurations, the optimal choices from a multitude of local paths are made by solving a combinatorial problem.

A typical case, especially in planar networks, is the rhombic configuration of four connected nodes where two opposite nodes represent local source and destination and the remaining two nodes could be used as intermediate nodes along two concurrent paths since each path consists of two hops as shown in Fig. 1. The said configuration of four nodes seems topologically trivial but there are challenging queuing issues associated with it.

The problem statement of local routing to be considered in this contribution is to find the optimal splitting ratio for a given packet rate so that the two paths could be utilized in the most efficient way for a given set of input traffic and service parameters. The realistic scenario of a single output channel, where the service rate of the channel changes dynamically when packets are sent to different neighbors, is used to interpret the packets from queuing perspective as different types of customers with their specific service rates. Also, each node can receive packets from several input channels simultaneously where the incoming packets are stored in a single first-come first-serve (FCFS) queue. This approach has been used by Batovski (2008) to implement the well known method of decomposition (Pujolle and Wu 1986) for the development of a custom network analyzer in multihop networks.

The said approach can be used to analyze the properties of local routing as well. Local rate-delay optimization model (Inthawadee and Batovski 2008) is to be defined for the chosen rhombic configuration of two paths and used together with the method of decomposition to obtain a suitable analytic model.

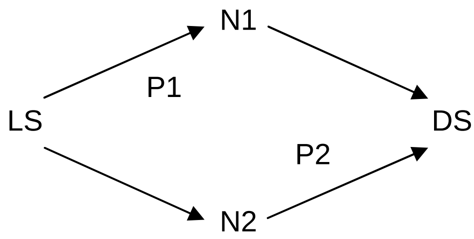

Fig. 1. A sample rhombic configuration of four connected nodes and two alternative paths.

Input Parameters

Denote the four nodes in Fig. 1 as local source (LS), local destination (LD), intermediate node N1 of local path P1 and intermediate node N2 of local path P2, correspondingly.

The description of traffic and service characteristics is considered in terms of the mean value and squared coefficient of variation (scv) of corresponding arbitrary distributions assuming GI/G/1 queuing (Kleinrock 1975) at each node.

Let the traffic to be relayed from LS to LD be described by an arbitrary inter-arrival distribution of mean λ_{LS} and squared coefficient of variation $c^2_{LS,\lambda}$. The remaining traffic to be delivered to all alternative neighbors of LS is considered as a cumulative background traffic being described by a similar arbitrary distribution of mean $\lambda_{LS,B}$ and squared coefficient of variation $c^2_{LS,B,\lambda}$.

The packet delivery from LS to N1 is described by an arbitrary service distribution of mean $\mu_{LS\to N1}$ and squared coefficient of variation $c^2_{LS\to N1,\mu}$. Similarly, the packet delivery from LS to N2 is described by another arbitrary distribution with $\mu_{LS\to N2}$ and $c^2_{LS\to N2,\mu}$. The cumulative service distribution for the background traffic is described by the pair of parameters $\mu_{LS,B}$ and $c^2_{LS,B,\mu}$.

Thus, the set of all input parameters of LS consists of the aforesaid pairs: $(\lambda_{LS}, c^2_{LS,\lambda})$, $(\lambda_{LS,B}, c^2_{LS,B,\lambda})$, $(\mu_{LS\to N1}, c^2_{LS\to N1,\mu})$, $(\mu_{LS\to N2}, c^2_{LS\to N2,\mu})$ and $(\mu_{LS,B}, c^2_{LS,B,\mu})$.

The traffic and service distributions of N1 are described by the following pairs of known parameters related to: the background traffic at N1, $(\lambda_{N1,B}, c^2_{N1,B,\lambda})$; the service distribution for the link N1→LD, $(\mu_{N1\to LD}, c^2_{N1\to LD,\mu})$; and the cumulative background traffic service distribution $(\mu_{N1,B}, c^2_{N1,B,\lambda})$.

Similarly, the pairs of known parameters of N2 are: $(\lambda_{N2,B}, c^2_{N2,B,\lambda})$; $(\mu_{N2\to LD}, c^2_{N2\to LD,\mu})$; and $(\mu_{N2,B}, c^2_{N2,B,\mu})$.

Analytic Model of Two-Hop Routing

The problem statement is to split the pair $(\lambda_{LS}, c^2_{LS,\lambda})$ into two pairs: $(\lambda_{LS\to N1}, c^2_{LS\to N1,\lambda})$ and $(\lambda_{LS\to N2}, c^2_{LS\to N2,\lambda})$, where:

$$\lambda_{LS} = \lambda_{LS\to N1} + \lambda_{LS\to N2}. \qquad (1)$$

Node LS

The total arrival rate at LS is the sum of the rates of the chosen flows from LS to N1 and N2 as well as the cumulative background flow:

$$\lambda_{LS,T} = \lambda_{LS\to N1} + \lambda_{LS\to N2} + \lambda_{LS,B}. \qquad (2)$$

The total node utilization of LS is

$$\rho_{LS,T} = \rho_{LS\to N1} + \rho_{LS\to N1} + \rho_{LS,B} \qquad (3)$$

where:

$$\rho_{LS\to N1} = \frac{\lambda_{LS\to N1}}{\mu_{LS\to N1}}, \qquad (4)$$

$$\rho_{LS\to N2} = \frac{\lambda_{LS\to N2}}{\mu_{LS\to N2}}, \qquad (5)$$

$$\rho_{LS,B} = \frac{\lambda_{LS,B}}{\mu_{LS,B}}. \qquad (6)$$

Therefore, the mean service rate of LS is derived from the reciprocal relationship for the total node utilization:

$$\mu_{LS,T} = \frac{\lambda_{LS,T}}{\dfrac{\lambda_{LS\to N1}}{\mu_{LS\to N1}} + \dfrac{\lambda_{LS\to N2}}{\mu_{LS\to N2}} + \dfrac{\lambda_{LS,B}}{\mu_{LS,B}}}. \quad (7)$$

The coefficient of variation $c^2_{LS,B}$ of the service time of LS is given by the following expression (Belch *et al*. 1998):

$$c^2_{LS,B} = -1 + \frac{\lambda_{LS\to N1}}{\lambda_{LS,T}}\left(\frac{\mu_{LS,T}}{\mu_{LS\to N1}}\right)^2 (c^2_{LS\to N1,\mu} + 1)$$

$$+ \frac{\lambda_{LS\to N2}}{\lambda_{LS,T}}\left(\frac{\mu_{LS,T}}{\mu_{LS\to N2}}\right)^2 (c^2_{LS\to N2,\mu} + 1)$$

$$+ \frac{\lambda_{LS,B}}{\lambda_{LS,T}}\left(\frac{\mu_{LS,T}}{\mu_{LS,B}}\right)^2 (c^2_{LS,B,\mu} + 1). \quad (8)$$

The method of decomposition consists of three distinct phases: *merging*, *flow*, and *splitting* (Pujolle and Wu 1986, Belch *et al*. 1998).

The *merging phase* is performed on the basis of statistical averaging of arriving flows:

$$c^2_{LS,A} = \frac{1}{\lambda_{LS,T}}(c^2_{LS,\lambda}\lambda_{LS} + c^2_{LS,B,\lambda}\lambda_{LS,B}). \quad (9)$$

In fact, $c^2_{LS,A}$ could directly be measured at LS and its approximate calculation is optional.

Among several known alternatives, the formula of Whitt (1983, 1983b) is used here for the *flow phase*:

$$c^2_{LS,D} = 1 + \rho^2_{LS,T}(c^2_{LS,B} - 1)$$

$$+ (1 - \rho^2_{LS,T})(c^2_{LS,A} - 1). \quad (10)$$

Then $c^2_{LS\to N1,\lambda}$ and $c^2_{LS\to N2,\lambda}$ are obtained during the *splitting* phase:

$$c^2_{LS\to N1,\lambda} = 1 + \frac{\lambda_{LS\to N1}}{\lambda_{LS,T}}(c^2_{LS,D} - 1), \quad (11)$$

$$c^2_{LS\to N2,\lambda} = 1 + \frac{\lambda_{LS\to N2}}{\lambda_{LS,T}}(c^2_{LS,D} - 1). \quad (12)$$

Knowing $c^2_{LS\to N1,\lambda}$ and $c^2_{LS\to N2,\lambda}$, the delay associated with each traffic flow can be obtained:

$$D_{LS\to N1} = \rho_{LS\to N1} + \frac{\rho_{LS\to N1}}{1-\rho_{LS,T}} \frac{c^2_{LS\to N1,\lambda} + c^2_{LS\to N1,\mu}}{2}$$

$$\times \phi(\rho_{LS,T}, c^2_{LS\to N1,\lambda}, c^2_{LS\to N1,\mu}), \quad (13)$$

$$D_{LS\to N2} = \rho_{LS\to N2} + \frac{\rho_{LS\to N2}}{1-\rho_{LS,T}} \frac{c^2_{LS\to N2,\lambda} + c^2_{LS\to N2,\mu}}{2}$$

$$\times \phi(\rho_{LS,T}, c^2_{LS\to N2,\lambda}, c^2_{LS\to N2,\mu}), \quad (14)$$

where the non-linear terms associated with GI/G/1 queues, $\phi(\rho_{LS,T}, c^2_{LS\to N1,\lambda}, c^2_{LS\to N1,\mu})$ and $\phi(\rho_{LS,T}, c^2_{LS\to N2,\lambda}, c^2_{LS\to N2,\mu})$, are obtained using the fit for arbitrary pairs of coefficients of variation introduced by Whitt (1993):

$$\phi = \begin{cases} \dfrac{4(c_A^2 - c_S^2)}{4c_A^2 - 3c_S^2}\phi_1 + \dfrac{c_S^2}{4c_A^2 - 3c_S^2}\Psi; c_A^2 \ge c_S^2 \\ \dfrac{c_S^2 - c_A^2}{2(c_A^2 + c_S^2)}\phi_3 + \dfrac{c_S^2 + 3c_A^2}{2(c_A^2 + c_S^2)}\Psi; c_A^2 \le c_S^2 \end{cases}, \quad (15)$$

$$\Psi = \begin{cases} 1; & c^2 \ge 1 \\ \phi_4^{2(1-c^2)}; & 0 \le c^2 \le 1 \end{cases} \text{ where } c^2 = \frac{c_A^2 + c_S^2}{2}, \quad (16)$$

$$\phi_1 = 1 + \gamma, \quad (17)$$

$$\phi_2 = 1 - 4\gamma, \quad (18)$$

$$\phi_3 = \phi_2 \exp\left(-\frac{2(1-\rho)}{3\rho}\right), \quad (19)$$

$$\phi_4 = \min\left\{1, \frac{\phi_1 + \phi_3}{2}\right\}, \quad (20)$$

$$\gamma = \min\left\{0.24, \frac{(1-\rho)(m-1)(\sqrt{4+5m} - 2)}{16m\rho}\right\}, \quad (21)$$

where: $\rho = \lambda/\mu$, λ is the mean arrival rate; μ is the mean service rate; m is the number of servers; c_A^2 is the scv of the inter-arrival process; and c_S^2 is the scv of the service-time distribution.

For the chosen communication model involving a single output server per node, assume that $m = 1$ (Batovski 2008). Then, whenever $c_A^2 \ge c_S^2$ and $c^2 = (c_A^2 + c_S^2)/2 \ge 1$, it follows that $\phi = 1$ and Eqs. (13) and (14) reduce to the well-known Allen-Cunneen approximation (Allen 1990):

$$D_{LS\to N1} = \rho_{LS\to N1} + \frac{\rho_{LS\to N1}}{1-\rho_{LS,T}} \frac{c_{LS\to N1,\lambda}^2 + c_{LS\to N1,\lambda}^2}{2}, \quad (22)$$

$$D_{LS\to N2} = \rho_{LS\to N2} + \frac{\rho_{LS\to N2}}{1-\rho_{LS,T}} \frac{c_{LS\to N2,\lambda}^2 + c_{LS\to N2,\lambda}^2}{2}. \quad (23)$$

Similar formulae are used for the calculation of $c_{N1\to LD,\lambda}^2$ of node N1 and $c_{N2\to LS,\lambda}^2$ of node N2. For completeness, the analytical expressions are included below.

Nodes N1 and N2

Due to the equivalence of the expressions for nodes N1 and N2, index *i* is used instead to denote each node, *i* = 1, 2. The total arrival rate at N_i is the sum of the rate of the chosen flow from each node N_i to LD as well as the cumulative background flow. Therefore:

$$\lambda_{Ni,T} = \lambda_{Ni\to LD} + \lambda_{Ni,B}, \quad (24)$$

$$\rho_{Ni,T} = \rho_{Ni\to LD} + \rho_{Ni,B}, \quad (25)$$

$$\rho_{Ni\to LD} = \frac{\lambda_{Ni\to LD}}{\mu_{Ni\to LD}}, \quad (26)$$

$$\rho_{Ni,B} = \frac{\lambda_{Ni,B}}{\mu_{LNiB}}, \quad (27)$$

$$\mu_{Ni,T} = \frac{\lambda_{Ni,T}}{\frac{\lambda_{Ni\to LD}}{\mu_{Ni\to LD}} + \frac{\lambda_{Ni,B}}{\mu_{Ni,B}}}, \quad (28)$$

$$c_{Ni,B}^2 = -1 + \frac{\lambda_{Ni\to LD}}{\lambda_{Ni,T}} \left(\frac{\mu_{Ni,T}}{\mu_{Ni\to LD}}\right)^2 (c_{Ni\to LD,\mu}^2 + 1)$$
$$+ \frac{\lambda_{Ni,B}}{\lambda_{Ni,T}} \left(\frac{\mu_{Ni,T}}{\mu_{Ni,B}}\right)^2 (c_{Ni,B,\mu}^2 + 1), \quad (29)$$

$$c_{Ni,A}^2 = \frac{1}{\lambda_{Ni,T}} (c_{Ni,\lambda}^2 \lambda_{Ni} + c_{Ni,B,\lambda}^2 \lambda_{Ni,B}), \quad (30)$$

$$c_{Ni,D}^2 = 1 + \rho_{Ni,T}^2 (c_{Ni,B}^2 - 1)$$
$$+ (1-\rho_{Ni,T}^2)(c_{Ni,A}^2 - 1), \quad (31)$$

$$c_{Ni\to LD,\lambda}^2 = 1 + \frac{\lambda_{Nii\to LD}}{\lambda_{Ni,T}} (c_{Ni,D}^2 - 1), \quad (32)$$

$$D_{Ni\to LD} = \rho_{Ni\to LD} + \frac{\rho_{Ni\to LD}}{1-\rho_{Ni,T}} \frac{c_{Ni\to LD,\lambda}^2 + c_{Ni\to LD,\mu}^2}{2}$$
$$\times \phi(\rho_{Ni,T}, c_{Ni\to LD,\lambda}^2, c_{Ni\to LD,\mu}^2), \quad (33)$$

where:

$$\lambda_{Ni} = \lambda_{Ni\to LD} = \lambda_{LS\to Ni}, \quad (34)$$

$$c_{Ni,\lambda}^2 = c_{LS\to Ni,\lambda}^2. \quad (35)$$

Rate-Delay Equalization

The rate-delay (λD) product of each path P1 and P2 is obtained as follows:

$$\lambda_{LS\to N1} D_{P1} = \lambda_{LS\to N1}(D_{LS\to N1} + D_{N1\to LD}), \quad (36)$$

$$\lambda_{LS\to N2} D_{P2} = \lambda_{LS\to N2}(D_{LS\to N2} + D_{N2\to LD}). \quad (37)$$

During the rate-delay equalization process, the following system of two equations is solved:

$$\lambda_{LS\to N1} D_{P1}(\lambda_{LS\to N1}) = \lambda_{LS\to N2} D_{P2}(\lambda_{LS\to N2}), \quad (38)$$

$$\lambda_{LS} = \lambda_{LS\to N1} + \lambda_{LS\to N2}, \quad (39)$$

which reduces to a single nonlinear equation after substituting Eq. (39) in Eq. (38):

$$\lambda_{LS\to N1} D_{P1}(\lambda_{LS\to N1}) =$$
$$(\lambda_{LS} - \lambda_{LS\to N1}) D_{P2}(\lambda_{LS} - \lambda_{LS\to N1}). \quad (40)$$

For the most general case of GI/G/1 queuing at the individual nodes, Eq. (40) can be rewritten in a more explicit form:

$$\lambda_{LS\to N1}[\ \rho_{LS\to N1}+\frac{\rho_{LS\to N1}}{1-\rho_{LS,T}}\frac{c^2_{LS\to N1,\lambda}+c^2_{LS\to N1,\mu}}{2}$$
$$\times\phi(\rho_{LS,T},c^2_{LS\to N1,\lambda},c^2_{LS\to N1,\mu})$$

$$+\rho_{N1\to LD}+\frac{\rho_{N1\to LD}}{1-\rho_{N1,T}}\frac{c^2_{N1\to LD,\lambda}+c^2_{N1\to LD,\mu}}{2}$$
$$\times\phi(\rho_{N1,T},c^2_{N1\to LD,\lambda},c^2_{N1\to LD,\mu})\ ]$$

$$=(\lambda_{LS}-\lambda_{LS\to N1})[$$
$$\rho_{LS\to N2}+\frac{\rho_{LS\to N2}}{1-\rho_{LS,T}}\frac{c^2_{LS\to N2,\lambda}+c^2_{LS\to N2,\mu}}{2}$$
$$\times\phi(\rho_{LS,T},c^2_{LS\to N2,\lambda},c^2_{LS\to N2,\mu})$$

$$+\rho_{N2\to LD}+\frac{\rho_{N2\to LD}}{1-\rho_{N2,T}}\frac{c^2_{N2\to LD,\lambda}+c^2_{N2\to LD,\mu}}{2}$$
$$\times\phi(\rho_{N2,T},c^2_{N2\to LD,\lambda},c^2_{N2\to LD,\mu})\ ].\qquad(41)$$

If the Allen-Cunneen approximation (Allen 1990) can be applied for a given set of input parameters, Eq. (41) reduces to:

$$\lambda_{LS\to N1}[\ \rho_{LS\to N1}+\frac{\rho_{LS\to N1}}{1-\rho_{LS,T}}\frac{c^2_{LS\to N1,\lambda}+c^2_{LS\to N1,\mu}}{2}+$$
$$\rho_{N1\to LD}+\frac{\rho_{N1\to LD}}{1-\rho_{N1,T}}\frac{c^2_{N1\to LD,\lambda}+c^2_{N1\to LD,\mu}}{2}\ ]$$

$$=(\lambda_{LS}-\lambda_{LS\to N1})[$$
$$\rho_{LS\to N2}+\frac{\rho_{LS\to N2}}{1-\rho_{LS,T}}\frac{c^2_{LS\to N2,\lambda}+c^2_{LS\to N2,\mu}}{2}$$

$$+\rho_{N2\to LD}+\frac{\rho_{N2\to LD}}{1-\rho_{N2,T}}\frac{c^2_{N2\to LD,\lambda}+c^2_{N2\to LD,\mu}}{2}\ ].$$
$$(42)$$

Equation (42) is a polynomial equation which can be solved for $\lambda_{LS\to N1}$ with standard numerical methods.

For a given set of 22 input parameters consisting of 11 (mean value, scv) pairs for the three nodes LS, N1, and N2:

LS: { λ_{LS}, $\lambda_{LS,B}$, $\mu_{LS\to N1}$, $\mu_{LS\to N2}$, $\mu_{LS,B}$, $c^2_{LS,\lambda}$, $c^2_{LS,B,\lambda}$, $c^2_{LS\to N1,\mu}$, $c^2_{LS\to N2,\mu}$, $c^2_{LS,B,\mu}$ },

N1: { $\lambda_{N1,B}$, $\mu_{N1\to LD}$, $\mu_{N1,B}$, $c^2_{N1,B,\lambda}$, $c^2_{N1\to LD,\mu}$, $c^2_{N1,B,\mu}$ },

N2: { $\lambda_{N2,B}$, $\mu_{N2\to LD}$, $\mu_{N2,B}$, $c^2_{N1,B,\lambda}$, $c^2_{N1\to LD,\mu}$, $c^2_{N1,B,\mu}$ },

the value of $\lambda_{LS\to N1}$ can be obtained and the resultant rate-delay product can be compared with the case when only one of the two paths is utilized.

M/M/1 Queuing

Eq. (42) can be further simplified for the private case of M/M/1 queuing (Kleinrock 1975) at the individual nodes (assuming that all the squared coefficients of variation of both traffic and service distributions are equal to 1), as follows:

$$\lambda_{LS\to N1}[\ \rho_{LS\to N1}+\frac{\rho_{LS\to N1}}{1-\rho_{LS,T}}+\rho_{N1\to LD}+\frac{\rho_{N1\to LD}}{1-\rho_{N1,T}}$$
$$]$$

$$=(\lambda_{LS}-\lambda_{LS\to N1})[\ \rho_{LS\to N2}+\frac{\rho_{LS\to N2}}{1-\rho_{LS,T}}$$

$$+\rho_{N2\to LD}+\frac{\rho_{N2\to LD}}{1-\rho_{N2,T}}\ ].\qquad(43)$$

Expressing the node utilizations as a function of the single variable, $\lambda_{LS\to N1}$, the following equation is obtained:

$$\lambda_{LS\to N1}[\frac{\lambda_{LS\to N1}}{\mu_{LS\to N1}}$$

$$+\frac{\dfrac{\lambda_{LS\to N1}}{\mu_{LS\to N1}}}{1-(\dfrac{\lambda_{LS\to N1}}{\mu_{LS\to N1}}+\dfrac{(\lambda_{LS}-\lambda_{LS\to N1})}{\mu_{LS\to N2}}+\dfrac{\lambda_{LS,B}}{\mu_{LS,B}})}$$

$$+\frac{\lambda_{LS\to N1}}{\mu_{N1\to LD}}+\frac{\dfrac{\lambda_{LS\to N1}}{\mu_{N1\to LD}}}{1-(\dfrac{\lambda_{LS\to N1}}{\mu_{N1\to LD}}+\dfrac{\lambda_{N1,B}}{\mu_{N1,B}})}\,]$$

$$=(\lambda_{LS}-\lambda_{LS\to N1})\,[\,\frac{(\lambda_{LS}-\lambda_{LS\to N1})}{\mu_{LS\to N2}}+$$

$$\frac{\dfrac{(\lambda_{LS}-\lambda_{LS\to N1})}{\mu_{LS\to N2}}}{1-(\dfrac{\lambda_{LS\to N1}}{\mu_{LS\to N1}}+\dfrac{(\lambda_{LS}-\lambda_{LS\to N1})}{\mu_{LS\to N2}}+\dfrac{\lambda_{LS,B}}{\mu_{LS,B}})}+$$

$$\frac{(\lambda_{LS}-\lambda_{LS\to N1})}{\mu_{N2\to LD}}$$

$$+\frac{\dfrac{(\lambda_{LS}-\lambda_{LS\to N1})}{\mu_{N2\to LD}}}{1-(\dfrac{(\lambda_{LS}-\lambda_{LS\to N1})}{\mu_{N2\to LD}}+\dfrac{\lambda_{N2,B}}{\mu_{N2,B}})}\,]. \qquad (44)$$

For a given set of 11 input parameters for the three nodes LS, N1, and N2:

LS: { λ_{LS}, $\lambda_{LS,B}$, $\mu_{LS\to N1}$, $\mu_{LS\to N2}$, $\mu_{LS,B}$ },

N1: { $\lambda_{N1,B}$, $\mu_{N1\to LD}$, $\mu_{N1,B}$ },

N2: { $\lambda_{N2,B}$, $\mu_{N2\to LD}$, $\mu_{N2,B}$ },

the value of $\lambda_{LS\to N1}$ can be easily obtained.

Computational Examples

The simplest case of M/M/1 queues is considered first. A sample source code for M/M/1 queues is shown in Fig. 2 where the input parameters are presented similarly to their appearance in Eq. (44). Two examples of solving Eq. (44) with Mathematica version 5.1 (2004) are provided for two different sets of input parameters.

The source code for GI/G/1 queues is not shown here due to space limitations but it has a similar structure and contains much more lines to include the node delay formulae and the formulae of the method of decomposition.

```
Clear[λLSN1, λLS, λLSB, μLSN1, μLSN2,
μLSB, λN1B, μN1LD, μN1B, λN2B, μN2LD,
μN2B];

λLS=0.2; λLSB=0.5; μLSN1=1.0; μLSN2=1.0;
μLSB=1.0; λN1B=0.6; μN1LD=0.9; μN1B=0.8;
λN2B=0.4; μN2LD=0.7; μN2B=0.9;

Plot[(λLSN1*((λLSN1/μLSN1)+((
λLSN1/μLSN1)/(1.0-((λLSN1/μLSN1)+(( λLS-
λLSN1)/μLSN2)+(λLSB/μLSB))))+(
λLSN1/μN1LD)+((λLSN1/μN1LD)/(1.0-
((λLSN1/μN1LD)+(λN1B/μN1B)))))),

((λLS-λLSN1)*(((λLS-λLSN1)/μLSN2)+((( λLS-
λLSN1)/μLSN2)/(1.0-((λLSN1/μLSN1)+((λLS-
λLSN1)/μLSN2)+( λLSB/μLSB))))+((λLS-
λLSN1)/μN2LD)+((( λLS-λLSN1)/μN2LD)/(1.0-
(((λLS-λLSN1)/μN2LD)+( λN2B/μN2B)))))),

{λLSN1, 0.0, λLS}, TextStyle -> {FontFamily ->
"Arial", FontSize -> 12}, Frame -> True,
FrameLabel -> {"λLS -> N1", "Rate-Delay
Product"}]
```

Fig. 2. A sample source code for obtaining solutions of Eq. (44).

M/M/1 queues

The ideal case of equal background traffic rates and equal service rates is shown as an example of equal splitting of the initial rate λ_{LS} among two equivalent paths P1 and P2. The second example deals with different service rates and the unequal splitting of the initial rate λ_{LS} between paths P1 and P2.

Equal Background Traffic Rates and Equal Service Rates of Paths P1 and P2

Consider the following set of input parameters (the highest mean service rate is normalized to 1):

LS: { λ_{LS} = 0.2, $\lambda_{LS,B}$ = 0.5, $\mu_{LS\to N1}$ = 1, $\mu_{LS\to N2}$ = 1, $\mu_{LS,B}$ = 1},

N1: { $\lambda_{N1,B}$ = 0.5, $\mu_{N1\to LD}$ = 1, $\mu_{N1,B}$ = 1},

N2: { $\lambda_{N2,B}$ = 0.5, $\mu_{N2\to LD}$ = 1, $\mu_{N2,B}$ = 1}.

This example is provided as a demonstration that if traffic and service conditions for both paths remain the same, the traffic is equally split among paths P1 and P2.

This is also a typical case of background traffic when the congestion builds up and the node utilization exceeds 40%. It is assumed that multiple flows of packets contribute to the background traffic.

A visual representation of the optimal point of the rate-delay equalization problem is shown in Fig. 3. The two curves in Fig. 3 represent the rate-delay product of paths P1 vs. $\lambda_{LS\rightarrow N1}$ and the rate-delay product of path P2 vs. $\lambda_{LS\rightarrow N2}$, where $\lambda_{LS\rightarrow N2}$ = λ_{LS} - $\lambda_{LS\rightarrow N1}$. The crossing point between the two curves is the optimal point of $\lambda_{LS\rightarrow N1}$ for which the balance of the two traffic flows is achieved.

Equal Background Traffic Rates and Different Service Rates of Paths P1 and P2

Consider the following set of input parameters:

LS: {λ_{LS} = 0.2, $\lambda_{LS,B}$ = 0.5, $\mu_{LS\rightarrow N1}$= 1, $\mu_{LS\rightarrow N2}$ = 1, $\mu_{LS,B}$ = 1},

N1: {$\lambda_{N1,B}$ = 0.1, $\mu_{N1\rightarrow LD}$= 1, $\mu_{N1,B}$ = 1},

N2: {$\lambda_{N2,B}$ = 0.1, $\mu_{N2\rightarrow LD}$ = 0.35, $\mu_{N2,B}$ = 0.35}.

In this example, the service rates of path P1 are greater than the service rates of path P2. The result of rate-delay equalization is shown in Fig. 4. The optimal point is shifted to the right indicating that the traffic rate $\lambda_{LS\rightarrow N1}$ of path P1 is greater than the traffic rate $\lambda_{LS\rightarrow N2}$ of path P2.

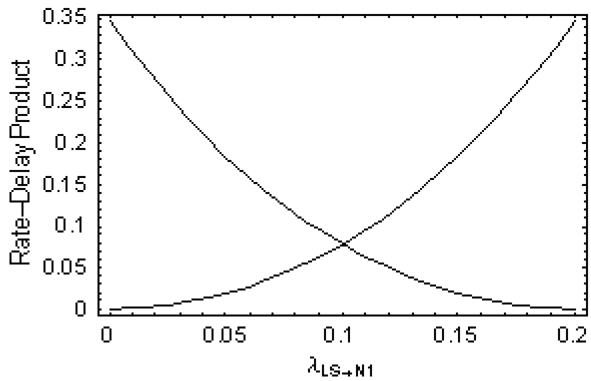

Fig. 3. Sample graphical solution of rate-delay equalization problem for equal background traffic rates and equal service rates of paths P1 and P2.

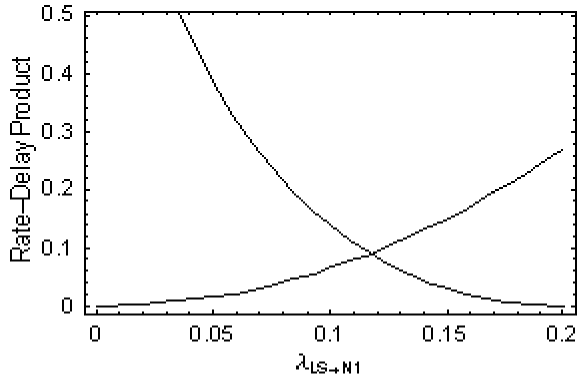

Fig. 4. Sample graphical solution of rate-delay equalization problem for equal traffic rates and different service rates of paths P1 and P2.

GI/G/1 Queues

The use of the method of decomposition for GI/G/1 queues is illustrated for three sets of input parameters of the three nodes LS, N1, and N2. Figures 5-7 shown below include the solutions obtained for M/M/1 queues, GI/G/1 queues with the Allen-Cunneen approximation, and GI/G/1 queues with the Whitt formula.

Equal Background Traffic Rates, Equal Service Rates and Different SCVs of Paths P1 and P2

Consider the following set of input parameters:

LS: {λ_{LS} = 0.2, $\lambda_{LS,B}$ = 0.5, $\mu_{LS\rightarrow N1}$ = 1, $\mu_{LS\rightarrow N2}$ = 1, $\mu_{LS,B}$ = 1, $c^2_{LS,\lambda}$ = 1, $c^2_{LS,B,\lambda}$ = 1, $c^2_{LS\rightarrow N1,\mu}$ = 1, $c^2_{LS\rightarrow N2,\mu}$ = 1, $c^2_{LS,B,\mu}$ = 1},

N1: {$\lambda_{N1,B}$ = 0.5, $\mu_{N1\rightarrow LD}$ = 1, $\mu_{N1,B}$ = 1, $c^2_{N1,B,\lambda}$ = 4, $c^2_{N1\rightarrow LD,\mu}$ = 4, $c^2_{N1,B,\mu}$ = 4},

N2: {$\lambda_{N2,B}$ = 0.5, $\mu_{N2\rightarrow LD}$ = 1, $\mu_{N2,B}$ = 1, $c^2_{N1,B,\lambda}$ = 0.5, $c^2_{N1\rightarrow LD,\mu}$ = 0.5, $c^2_{N1,B,\mu}$ = 0.5}.

The SCVs of path P1 are set to 4 and the SCVs of path P2 are set to 0.5. The optimal point of rate-delay equalization is shown in Fig. 5. The optimal $\lambda_{LS\rightarrow N1}$ is shifted to the left so that $\lambda_{LS\rightarrow N1} < \lambda_{LS\rightarrow N2}$. The solution obtained for M/M/1 queues is identical with the one shown in Fig. 3 and is provided here for comparison. No significant difference is observed for graphs obtained with the Allen-

Cunneen approximation and with the Whitt formula and the said graphs almost coincide with each other. The differences in the SCVs have an effect similar to the one observed for different service rates of paths P1 and P2. Therefore, the knowledge of the second moment of traffic and service distributions could allow certain adjustment of the optimal $\lambda_{LS\to N1}$ if the distributions are non-exponential.

Different Background Traffic Rates and Service Rates and Different SCVs of Paths P1 and P2

Consider the following set of input parameters:

LS: {λ_{LS} = 0.2, $\lambda_{LS,B}$ = 0.5, $\mu_{LS\to N1}$ = 1, $\mu_{LS\to N2}$ = 1, $\mu_{LS,B}$ = 1, $c^2_{LS,\lambda}$ = 1, $c^2_{LS,B,\lambda}$ = 1, $c^2_{LS\to N1,\mu}$ = 1, $c^2_{LS\to N2,\mu}$ = 1, $c^2_{LS,B,\mu}$ = 1},

N1: {$\lambda_{N1,B}$ = 0.5, $\mu_{N1\to LD}$ = 1, $\mu_{N1,B}$ = 1, $c^2_{N1,B,\lambda}$ = 4, $c^2_{N1\to LD,\mu}$ = 4, $c^2_{N1,B,\mu}$ = 4},

N2: {$\lambda_{N2,B}$ = 0.5, $\mu_{N2\to LD}$ = 0.75, $\mu_{N2,B}$ =0.75, $c^2_{N1,B,\lambda}$ = 0.5, $c^2_{N1\to LD,\mu}$ = 0.5, $c^2_{N1,B,\mu}$ = 0.5}.

The SCVs of path P1 are set to 4 and the SCVs of path P2 are set to 0.5. The service rates of path P2 are reduced to 0.75. The optimal point of rate-delay equalization is shown in Fig. 6. The optimal solution, $\lambda_{LS\to N1} \approx \lambda_{LS\to N2}$, is located at the center of the graph and can be interpreted in terms of the mutual compensation of two opposing tendencies, a shift to the left in favor of path P2 due to the higher SCVs (uncertainties) of path P1 and another shift to the right in favor of path P1 due to the lower service rates of path P2. The solution obtained for M/M/1 queues is shifted to the right, $\lambda_{LS\to N1} > \lambda_{LS\to N2}$, due to the reduced service rates of path P2.

This example shows that both first and second moments of traffic and service distributions play an important role in shifting the optimal point of $\lambda_{LS\to N1}$. The most significant shifts are to be expected when service rates increase and SCVs decrease for one of the paths as compared to the second one.

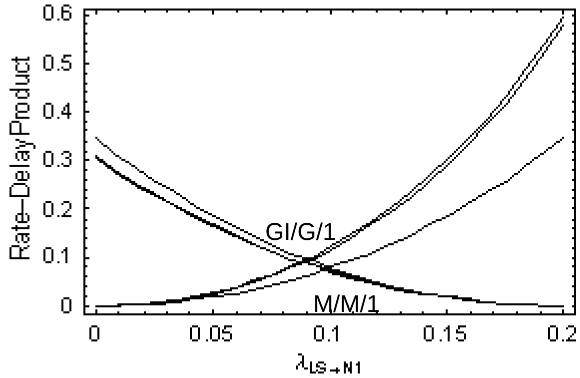

Fig. 5. Sample graphical solution of rate-delay equalization problem for equal background traffic rates, equal service rates and different SCVs of paths P1 and P2.

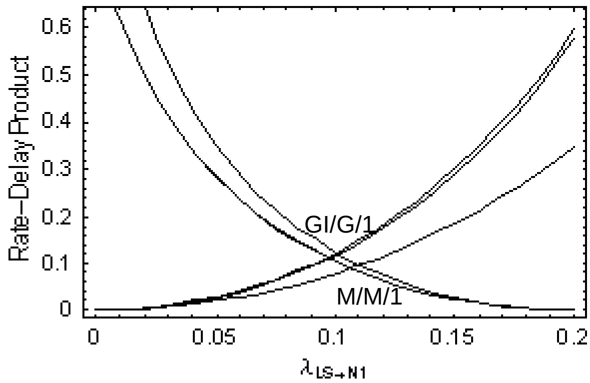

Fig. 6. Sample graphical solution of rate-delay equalization problem for different background traffic rates and service rates and different SCVs of paths P1 and P2.

Congested Single Paths

Consider the following set of input parameters:

LS: {λ_{LS} = 0.2, $\lambda_{LS,B}$ = 0.5, $\mu_{LS\to N1}$ = 1, $\mu_{LS\to N2}$ = 0.4, $\mu_{LS,B}$ = 1, $c^2_{LS,\lambda}$ = 1.5, $c^2_{LS,B,\lambda}$ = 1.1, $c^2_{LS\to N1,\mu}$ = 1.1, $c^2_{LS\to N2,\mu}$ = 1.1, $c^2_{LS,B,\mu}$ = 1},

N1: {$\lambda_{N1,B}$ = 0.2, $\mu_{N1\to LD}$ = 0.4, $\mu_{N1,B}$ = 0.4, $c^2_{N1,B,\lambda}$ = 4, $c^2_{N1\to LD,\mu}$ = 4, $c^2_{N1,B,\mu}$ = 4},

N2: {$\lambda_{N2,B}$ = 0.5, $\mu_{N2\to LD}$ = 1, $\mu_{N2,B}$ = 1, $c^2_{N1,B,\lambda}$ = 1.1, $c^2_{N1\to LD,\mu}$ = 1.1, $c^2_{N1,B,\mu}$ = 1.1}.

The optimal point of $\lambda_{LS\to N1}$ for rate-delay equalization is shown in Fig. 7. The result indicates that the reduced service rates of both paths could follow to congestion if only one of the paths is utilized. The splitting of λ_{LS} greatly improves the performance.

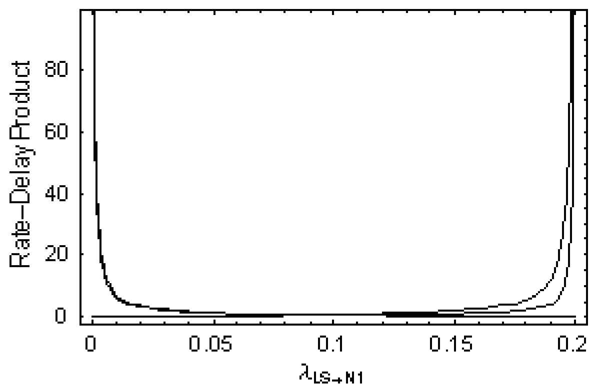

Fig. 7. Sample graphical solution of rate-delay equalization problem for congested single paths.

Analysis and Discussion

The obtained graphs show that it is not mandatory to obtain the exact optimization point of rate-delay equalization in order to have a significant improvement of the performance. The slow increment of intersecting curves around the optimal point allows one to use sub-optimal solutions with the same success. This is quite obvious especially for the case of strong congestion as shown in Fig. 7 where a wide plateau is formed around the optimal point and any choice of $\lambda_{LS\to N1}$ within the plateau could significantly improve the performance of local packet relaying as compared to the long delays at the edges of the displayed graphs.

The provided numerical examples demonstrate that the optimization problem for the rhombic configuration of four nodes is easily solvable numerically and fast sub-optimal algorithms could possibly be implemented in decentralized networks whenever a loop of four nodes is formed locally for a certain period of time.

The analysis of different traffic and service scenarios follows to the conclusion that the performance in terms of delay optimization can be improved up to one order of magnitude in cases of uncongested and slightly congested local paths and up to two orders of magnitude in extreme cases of strongly congested local paths. The throughput is also improved and in most scenarios remains optimal if the traffic rates after rate-delay equalization do not exceed the service rates or specific node utilization thresholds above which packet traffic drop is initiated. The exact level of performance depends on the specific set of parameters for a given scenario.

It should be noted that more precise results with the method of decomposition for small or big SCVs can be obtained with a complex modification proposed by Whitt (1994). Small SCVs in wireless networks being related to quasi-deterministic traffic and service distributions are to be expected in some private cases with low congestion levels in efficiently configured wireless network topologies and high service rates for channels with high signal-to-noise ratios.

Conclusion

The local routing along a rhombic configuration of four nodes is considered for a given packet flow assuming that the information about all other flows is interpreted in terms of known background traffic. The proper splitting of a traffic flow among two-hop local paths P1 and P2 requires the exchange of control packets only one hop away from each node. Thus the local topology forming a loop of four connected nodes is estimated to be most reliable among other alternatives because the local rerouting involving more than two hops is unstable in decentralized networks in mobile environment. Also, it is more difficult to provide real-time optimization with the substantial increase of the number of parameters to be shared more than one hop away from each node.

Note that a loop of three nodes is to be considered as a private case and a simplification of the four-node loop. In the triangular case, one local path consists of a single hop and the second path has two hops. Such scenario could be beneficial in some cases when the single-hop path has substantially lower service rates and higher SCVs (uncertainties) than the two-hop alternative.

References

Allen, A.O. 1990 Probability, statistics and queueing theory with computer science applications. Academic Press Professional, Inc., San Diego, CA, USA.

Batovski, D.A. 2008 Semi-analytic evaluation of quality of service parameters in multihop networks. Assumption University Journal of Technology (AU J.T.) 11(4): 215-224, April.

Belch, G.; Greiner, S.; de Meer, H.; and Trivedi, K.S. 1998. Queueing networks and Markov chains: Modeling and performance evaluation with computer science applications. Wiley-Interscience, John Wiley & Sons, Inc., New York, NY, USA.

Inthawadee, S.; and Batovski, D.A. 2008. Flow control in distributed gateways. Annals of Telecommunications 63(9-10): 523-527, September-October.

Kleinrock, L. 1975. Queueing systems, Volume 1: Theory, Wiley Interscience, New York, NY, USA.

Mathematica. 2004. Version 5.1. Wolfram Research, Inc., Champaign, IL, USA.

Pujolle, G., and Wu, A. 1986. A solution for multiserver and multiclass open queueing networks. Information Systems and Operations Research 24(3): 221-30.

Whitt, W. 1983. The queueing network analyzer. Bell System Technical Journal 62(9): 2779-815, November.

Whitt, W. 1983b. Performance of the queueing network analyzer. Bell System Technical Journal 62(9): 2817-43, November.

Whitt, W. 1993. Approximations for the GI/G/m queue. Production and Operations Management 2(2): 114-61, Spring.

Whitt, W. 1994. Towards better multi-class parametric-decomposition approximations for open queueing networks. Annals of Operations Research 48: 221-48.