BRAESS-LIKE PARADOXES OF NASH EQUILIBRIA FOR LOAD BALANCING IN DISTRIBUTED COMPUTER SYSTEMS

Hisao Kameda, Eitan Altman, and Takayuki Kozawa

November 30, 1998

ISE-TR-98-157

H. Kameda is with the Institute of Information Sciences and Electronics, University of Tsukuba, Tsukuba Science City, Ibaraki 305-8573, Japan, Tel: +81-298-53-5539, Fax: +81-298-53-5206, Email: kameda@is.tsukuba.ac.jp

E. Altman is with INRIA, BP93, 2004, Route des Lucioles, 06902 Sophia Antipolis Cedex, France. Tel: +33-4-92-38-77-86, Fax: +33-4-93-65-78-58, Email: Eitan.Altman@sophia.inria.fr

T. Kozawa is with Doctoral Program in Engineering, University of Tsukuba, Tsukuba Science City, Ibaraki 305-8573, Japan, Tel: +81-298-53-5156, Fax: +81-298-53-5206, Email: kozawa@osdp.is.tsukuba.ac.jp

This work was supported in part by the Japan Society for the Promotion of Science and in part by the Telecommunications Advancement Organization of Japan.

Braess-like Paradoxes of Nash Equilibria for Load Balancing in Distributed Computer Systems

Hisao Kameda,* Eitan Altman,† and Takayuki Kozawa‡

Abstract

It is well known that inefficiencies occur in the use of resources of computer, telecommunication and traffic networks when routing decisions are taken by the users. The famous Braess paradox which was originally identified and observed in road traffic context, showed that it may happen that by adding capacity to the network, the performance of all users degrades. This paradox has long been known in a framework called the Wardrop equilibrium, in which there are infinitely many individuals (such as car drivers) and in which the decision of one single individual has a negligible effect on the performance of the other indivuduals. Another framework in which such a paradox may occur is that of the Nash equilibrium in which there are a finite number of players, and in which the decision of each player has nonnegligible effect on the other players. It is known that the Nash equilibrium converges to the Wardrop equilibrium when the number of users becomes large. It is thus natural to expect the same type of paradox in the Nash equilibrium context (for a large number of players), whenever it occurs for the Wardrop equilibrium. In this paper, we present cases where a paradox similar to that of the Braess appears in a Nash equilibrium but does not appear in a Wardrop equilibrium in the same environment. We consider the model of load balancing in distributed computer systems. We further establish the uniqueness of the Nash equilibrium for this problem.

Keywords: Braess paradox, Nash equilibrium, Wardrop equilibrium, performance optimization, distributed computer system, load balancing.

1 Introduction

In many systems including communication networks in distributed computer systems, transportation flow networks, etc., we have several distinct objectives for performance optimization. Among them, we have three typical objectives or optima:

*H. Kameda is with the Institute of Information Sciences and Electronics, University of Tsukuba, Tsukuba Science City, Ibaraki 305-8573, Japan. Tel: +81-298-53-5539 Fax: +81-298-53-5206 E-mail: kameda@is.tsukuba.ac.jp .

†E. Altman is with INRIA B.P. 93, 06902 Sophia Antipolis Cedex, France. Tel: +33-4-92-38-77-86 Fax: +33-4-92-38-77-65 E-mail: altman@sophia.inria.fr

‡T. Kozawa is with Doctoral Program in Engineering, University of Tsukuba, Tsukuba Science City, Ibaraki 305-8573, Japan. Tel: +81-298-53-5156 Fax: +81-298-53-5206 E-mail: kozawa@osdp.is.tsukuba.ac.jp .

(1) the system-optimum, overall optimum, or social optimum, where a certain overall and single measure like the total cost or the overall average response time is to be optimized. We call it the *overall optimum* here.
(2) the individual optimum, Wardrop optimum, or user optimum by some people; where each of infinitely many individuals, users, or jobs cannot receive any benefit by changing its own decision. We call it the *individual optimum* here.
(3) the Nash noncooperative optimum, class optimum, or user optimum by some other people, where each of a finite number of users, classes, or players cannot receive any benefit by changing its decision. We call it the *class optimum* here.

Actually, (3) is reduced to (1) when the number of players reduces to 1 and approaches (2) when the players are symmetric and the number of players becomes infinitely many [4].

We can think that the total processing capacity of a system will increase when the capacity of a part of the system increases, and so we expect improvements in performance objectives accordingly in that case. The Braess paradox tells us that this is not always the case; i.e., increased capacity of a part of the system may sometimes lead to the degradation in the benefits of all users in a Wardrop equilibrium [1, 2, 3, 4]. As it is known that the Nash equilibrium converges to the Wardrop equilibrium as the number of users becomes large [4], it is thus natural to expect the same type of paradox in the Nash equilibrium context (for a large number of players), whenever it occurs for the Wardrop equilibrium. Indeed, Korilis et al. found examples wherein the Braess-like paradox appears in a Nash equilibrium with symmetric user classes where all user classes are identical as well as in a Wardrop equilibrium [10, 11].

On the other hand, it has been observed that the increased capacity of a part of a system may lead to somewhat awkward behavior in terms of a system wide measure [5, 6, 17]. In particular, Kameda et al. found a seemingly anomalous case where in a Nash equilibrium each of two processing nodes (servers) forwards the same type of jobs mutually to be processed by the other node, thus incurring additional communication delays [6].

In this paper, we consider a load balancing problem for distributed computer systems and give examples wherein the increased capacity of a part of a system would degrade the benefits of all classes in a Nash equilibrium whereas it should not degrade the benefits of all classes at the same time in a Wardrop equilibrium in the same environment. Our model has asymmetric classes; i.e., classes are not identical. In doing so, we have established the uniqueness of the Nash equilibrium for the new class of problems.

2 The Model and Assumptions

We consider a model consisting of two nodes (hosts) and a communication means that connects both nodes. Nodes are numbered 1 and 2. Each node consists of a single exponential server with service rate μ_i ($i = 1, 2$). Node i has the external Poisson arrival with rate ϕ_i, out of which the rate x_{ii} of jobs are processed at node i. The rate x_{ij} ($i \neq j$) of jobs are forwarded through the communication means to the other node j to be processed there, and the results of those jobs are returned back through the communication means to node i. Then we have $x_{ii} + x_{ij} = \phi_i$ ($i \neq j$), $x_{ij} \geq 0$,

$i, j = 1, 2$. We denote the vector $(x_{11}, x_{12}, x_{21}, x_{22})$ by $\boldsymbol{x}$. We denote the set of $\boldsymbol{x}$'s that satisfy the constraints by $\boldsymbol{C}$ and let $\Phi = \phi_1 + \phi_2$. Within these constraints, a set of values of x_{ij} $(i, j = 1, 2)$ are chosen to achieve optimization. Thus the load on node i is $x_{ii} + x_{ji}$ $(i \neq j)$ and is denoted by β_i. Then, the expected processing (including queueing) time of a job that is processed at node i, is $1/(\mu_i - \beta_i)$ for $\beta_i < \mu_i$ (otherwise it is infinite).

As to the communication means, we consider two alternatives.

(A) The one is a single-channel communication line that is used commonly in forwarding and sending back of jobs that arrive at both nodes. We assume that the expected time length of forwarding and sending back a job is to be

$$G(\lambda) = \frac{t}{1 - (x_{12} + x_{21})t}$$

if $x_{12} + x_{21} < 1$, and is otherwise inifinite, where $\lambda = x_{12} + x_{21}$ is the network traffic. That is, we assume the communication channel is modelled by a processor sharing server with service rate $1/t$; *i.e.*, the mean communication (without queueing) time is t, and, thus, the capacity of the communication line is t.

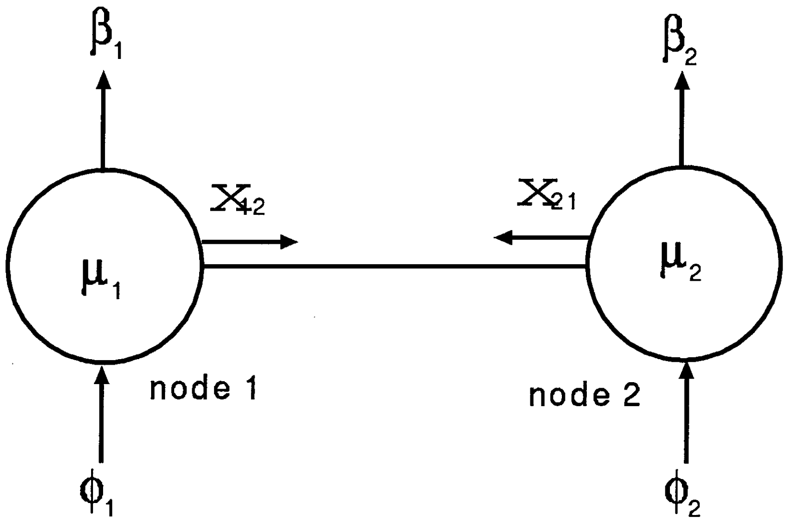

Figure 1: The system model. (case (A))

(B) The other consists of two-way communication lines 1 and 2. One two-way line i is used for forwarding of a job that arrives at node i (and for sending back the processed result of the job). The assumption on the line is the same as (A) except that there are two lines each of which is used only for jobs arriving at one node and is not used in common by two nodes. Thus the expected communication (with queueing) time of a job arriving at node i and being processed at node j ($\neq i$) is expressed as

$$G_i(x_{ij}) = \frac{t}{1 - x_{ij}t}$$

if $x_{ij}t < 1$ (and is otherwise infinite). For case (A), $(1/t) - x_{ij}$ $(i \neq j)$ is the amount of the line capacity left to be used by node i in forwarding the jobs arriving at node i.

Thus, clearly, the communication capacity for case (B) is greater than or equal to that for case (A).

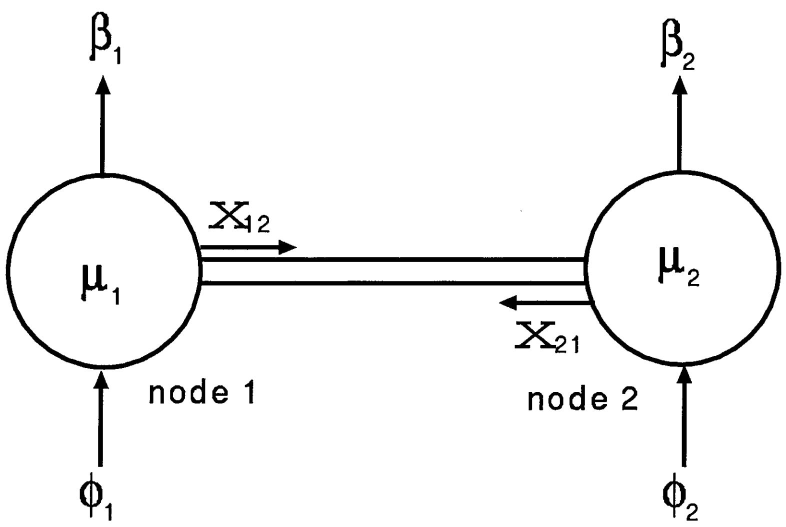

Figure 2: The system model. (case (B))

We refer to the length of time between the instant when a job arrives at a node and the instant when a job leaves the node, where it has arrived, after all processing and communication, if any, are over as *the response time for a job* arriving at the node.

Thus the expected response time of a job that arrives at node i is

$$T_i(\boldsymbol{x}) = \frac{1}{\phi_i} \sum_k x_{ik} T_{ik}(\boldsymbol{x}),$$

where

$$T_{ii}(\boldsymbol{x}) = \frac{1}{\mu_i - \beta_i}$$

if $\beta_i < \mu_i$ (and it is otherwise inifinite), and for $j \neq i$,

$$\begin{aligned} T_{ij}(\boldsymbol{x}) &= \frac{1}{\mu_j - \beta_j} + \frac{t}{1-(x_{ij}+x_{ji})t}, \qquad \text{for case (A)}, \\ &= \frac{1}{\mu_j - \beta_j} + \frac{t}{1-x_{ij}t}, \qquad \text{for case (B)}. \end{aligned}$$

(The above expressions hold, again, only for positive values of denominators, and are otherwise infinite.)

Then, the overall expected response time of a job that arrives at the system is

$$T(\boldsymbol{x}) = \frac{1}{\Phi} \sum_i \phi_i T_i.$$

We have three optima, the overall, the individual, and the node.

(1) The overall optimum is given by such $\bar{\boldsymbol{x}}$ as satisfies the following,

$$T(\bar{\boldsymbol{x}}) = \min T(\boldsymbol{x}) \quad \text{with respect to} \quad \boldsymbol{x} \in \boldsymbol{C}.$$

The solution $\bar{\boldsymbol{x}}$ is characterized by the Kuhn-Tucker condition or the following variational inequality,

$$\boldsymbol{t}(\bar{\boldsymbol{x}})(\boldsymbol{x}-\bar{\boldsymbol{x}}) \geq 0 \qquad \text{for all} \quad \boldsymbol{x} \in \boldsymbol{C}$$

$$\text{where} \quad \boldsymbol{t}(\boldsymbol{x}) = (\frac{\partial}{\partial x_{11}}\Phi T, \frac{\partial}{\partial x_{12}}\Phi T, \frac{\partial}{\partial x_{21}}\Phi T, \frac{\partial}{\partial x_{22}}\Phi T) \quad \text{as seen from [7].}$$

(2) The individual optimum is given by such $\hat{\boldsymbol{x}}$ as satisfies the following for all i,

$$T_i(\hat{\boldsymbol{x}}) = \min_{i \neq j}\{T_{ii}(\hat{\boldsymbol{x}}), T_{ij}(\hat{\boldsymbol{x}})\} \quad \text{such that} \quad \hat{\boldsymbol{x}} \in \boldsymbol{C}$$

The solution $\hat{\boldsymbol{x}}$ is characterized by the following variational inequality,

$$\boldsymbol{T}(\hat{\boldsymbol{x}})(\boldsymbol{x}-\hat{\boldsymbol{x}}) \geq 0 \quad \text{for all} \quad \boldsymbol{x} \in \boldsymbol{C}$$

$$\text{where} \quad \boldsymbol{T}(\boldsymbol{x}) = (T_{11}(\boldsymbol{x}), T_{12}(\boldsymbol{x}), T_{21}(\boldsymbol{x}), T_{22}(\boldsymbol{x})) \quad \text{as seen from [7].}$$

(3) The node optimum is given by such $\tilde{\boldsymbol{x}}$ as satisfies the following for all i,

$$T_i(\tilde{\boldsymbol{x}}) = \min T_i(\boldsymbol{x}), \quad \text{with respect to } x_{ii},\ x_{ij}\ (j \neq i), \quad \text{such that } \boldsymbol{x} \in \boldsymbol{C}.$$

The solution $\tilde{\boldsymbol{x}}$ is characterized by the following variational inequality

$$\tilde{\boldsymbol{t}}(\tilde{\boldsymbol{x}})(\boldsymbol{x}-\tilde{\boldsymbol{x}}) \geq 0 \quad \text{for all} \quad \boldsymbol{x} \in \boldsymbol{C}$$

where

$$\tilde{\boldsymbol{t}}(\boldsymbol{x}) = (\frac{\partial}{\partial x_{11}}\phi_1 T_1, \frac{\partial}{\partial x_{12}}\phi_1 T_1, \frac{\partial}{\partial x_{21}}\phi_2 T_2, \frac{\partial}{\partial x_{22}}\phi_2 T_2) \quad \text{as is shown similary as above.}$$

We are sure of the existence and the uniqueness of the overall, individual, and node (class) optima for the model given here. For the existence and uniqueness of those optima see Section 5.

Remark 2.1 Note that, there should be no mutual forwarding in overall and individual optima. That is, in overall and individual optima, either one of x_{ij} $(i \neq j)$ must be zero in case (A) due to [16] or to Section 2.2.2 of [5] and in case (B) due to [12]. And thus, when one of x_{ij} $(i \neq j)$ say x_{ij} is non zero, $T_i(\boldsymbol{x})$ decreases and $T_j(\boldsymbol{x})$ increases with the increase of t as shown in Theorems 2.5 and 2.7 of [5].

3 The Numerical Experiments

We examined the cases of the following parameter values of the model given in Section 2.

$$\begin{array}{ll} \phi_1 = 100(\text{jobs/sec}), & \phi_2 = 20(\text{jobs/sec}) \\ \mu_1 = 120(\text{jobs/sec}), & \mu_2 = 25(\text{jobs/sec}) \end{array}$$

The value of the communication (without queueing) time parameter t is varied from 0 (*sec*) till 1 (*sec*) in steps of 0.005 (*sec*), *i.e.*, t = 0, 0.005, 0.010, 0.015, $\cdots$, 1.

The algorithms used to obtain the overall and individual optima are based on the algorithms given in [8, 9]. The algorithm for the node optimum is obtained similary as above.

4 Results and Discussion

Figures 3 and 4 shows the overall mean response time with various values of the mean communication time parameter t for overall, node (class), and individual optima for cases (A). Note that the overall and individual optima must be identical for cases (A) and (B). Only node optima for cases (A) and (B) may be different from each other. Figures 5 and 6 show the values of node optimal T_1 (Figure 5) and T_2 (Figure 6) with various values of t for the cases (A) and (B). We can observe the Braess-like paradoxes for node optima in three ways.

(I) For case (A) with the decrease of t.

(II) For case (B) with the decrease of t.

(III) With the increase of the line capacity from case (A) to case (B).

That is, there are chances that both T_1 and T_2 increase with the increase of the communication line capacity as in (I), (II) and (III).

Paradox (I) for case (A)

In Figure 6, for $0.08 \leq t \leq 0.12$, we see that there are sets of two values of t for which T_1 and T_2 with the smaller value of t are greater than T_1 and T_2 with the larger value of t, which looks paradoxical. For example, consider the case with $t = 0.08$. Then

$$T_{11} = 0.058..., \quad T_{12} = 0.299..., \quad T_{21} = 0.228..., \quad T_{22} = 0.128...,$$
$$x_{11} = 98.060..., \quad x_{12} = 1.939..., \quad x_{21} = 4.707..., \quad x_{22} = 15.292....$$

Therefore, by noting that $T_1 = (x_{11}T_{11} + x_{12}T_{12})/\phi_1, \quad T_2 = (x_{21}T_{21} + x_{22}T_{22})/\phi_2,$

$$T_1 = 0.0627..., \quad T_2 = 0.1522....$$

Furthermore, consider the case with $t = 0.11$. Then

$$T_{11} = 0.060..., \quad T_{12} = 0.294..., \quad T_{21} = 0.236..., \quad T_{22} = 0.118...,$$
$$x_{11} = 100.0, \quad x_{12} = 0.0, \quad x_{21} = 3.405..., \quad x_{22} = 16.594....$$

Therefore

$$T_1 = 0.0602..., \quad T_2 = 0.1389....$$

Paradox (II) for case (B)

Similary as above, in Figure 6, for $0.08 \leq t \leq 0.12$, we see that there are sets of two values of t for which T_1 and T_2 with the smaller value of t are greater than T_1 and T_2 with the larger value of t, which looks again paradoxical. For example, consider the case with $t = 0.08$. Then

$$T_{11} = 0.056..., \quad T_{12} = 0.249..., \quad T_{21} = 0.214..., \quad T_{22} = 0.135...,$$
$$x_{11} = 96.238..., \quad x_{12} = 3.761..., \quad x_{21} = 6.161..., \quad x_{22} = 13.838....$$

Therefore

$$T_1 = 0.0640..., \quad T_2 = 0.1596..., \quad \text{for} \quad t = 0.08.$$

Furthermore, consider the case with $t = 0.11$. Then

$$T_{11} = 0.060..., \quad T_{12} = 0.228..., \quad T_{21} = 0.236..., \quad T_{22} = 0.118...,$$
$$x_{11} = 100.0, \quad x_{12} = 0.0, \quad x_{21} = 3.405..., \quad x_{22} = 16.594....$$

Therefore

$$T_1 = 0.0602..., \quad T_2 = 0.1389..., \quad \text{for} \quad t = 0.11.$$

Paradox (III) between cases (A) and (B)

We can see for $0.06 \leq t \leq 0.1$, both T_1 and T_2 of case (A) are smaller than T_1 and T_2 of case (B), respectively. Clearly, this looks paradoxical. That is, although the system in case (B) has a larger communication capacity than in case (A), the expected response times in case (B) for both nodes are larger than the corresponding values in case (A) at the same time. We show the details of some case as an example. For example, consider the case where $t = 0.08$. Then as given (I) and (II) in the above. We have

$$\begin{array}{ll} \text{For (A)}: & T_1 = 0.0627..., \quad T_2 = 0.1522..., \\ \text{For (B)}: & T_1 = 0.0640..., \quad T_2 = 0.1596.... \end{array}$$

Note again that the communication line capacity for each node in case (B) is greater than or equal to the communication line capacity for the node in case (A). In the overall and individual optima, there are no differences between cases (A) and (B). Recall what we noted at the end of the previous section. Only in the node optimum, the mean response time of each node (class) in case (A) is worse than that of the node (class) in case (B).

As we saw in the last few statements given in Section 2, for individual optima, such a Braess-like paradox as given here would not occur. Figure 7 shows the values of individual optimal T_1 nad T_2 with various values of t. Those values are identical for both cases (A) and (B). And we can see that T_1 and T_2 for a value of t cannot be greater than for another value of t at the same time, which is not paradoxical.

5 The Uniqueness of the Optima

As described near the end of Section 2, the model given here has the unique values for its overall, individual, and node optima. The uniqueness for overall and individual optima has been given in a more general setting [7], and in Subsection 5.1, we show how the uniqueness for this model can be derived from the general setting. The uniqueness of the node optimum for this model has not been shown yet, and we show it by extending the result of [13]. Thus we present the proof in the setting as general as possible here in Subsection 5.2. We also give some properties of the node optimum that can be derived theoretically in Subsection 5.3.

5.1 The Uniqueness of the Overall and Individual Optima

We show in this subsection the uniqueness of the overall and individual optima for cases (A) and (B). Optimization in the models of this paper is regarded as optimal static routing in open state-independent BCMP queueing networks as defined in [7] in the following way.

Case (A): In the open BCMP network corresponding to the model of this paper, there are three single-server service centers, node 1, 2 and communication channel c, and two sets of an origin and a destination which are service centers with zero service time. Jobs arriving at node i of the model arrive first at origin i of the BCMP network, $i = 1, 2$: From origin i, each job passes through one of two paths: the one is 'origin i

– node i – destination i', and the other is 'origin i – communication channel c – node j $(j \neq i)$ – destination i'. (We can make the latter path to look more realistic by changing its part 'node j $(j \neq i)$ – destination i' to 'node j $(j \neq i)$ – communication channel c – destination i', but this is no essential change in modelling.) The rate of jobs passing through the one and other paths are x_{ii} and x_{ij} $(j \neq i)$, respectively.

Thus we can apply the theorems 3.4 and 4.4 of [7] and show that the utilization factors ρ_1, ρ_2 and ρ_c of nodes 1, 2, and communication channel c are unique in both overall and individual optima. Since we have two independent variables x_{11} and x_{12} and two independent equations.

$$x_{11} + \phi_2 - x_{22} = \rho_1 \mu_1,$$
$$\phi_1 - x_{11} + \phi_2 - x_{22} = \rho_c / t,$$

we see that $\boldsymbol{x}$ is unique in both overall and individual optima in the model.
Case (B): We can show the uniqueness similarly as in Case (A).

5.2 Uniqueness of the Node Optimum

5.2.1 Extending the Model

We show in this subsection the uniqueness of the node optimum or Nash equilibrium for both case (A) and case (B). We shall consider a more general cost, and the more general topology.

More precisely, let there be K sets of classes: $\mathcal{K}_i$ $i = 1, ..., K$ where $\mathcal{K}_i$ contains k_i classes of jobs. The set $\mathcal{K}_i$ corresponds to the classes that send jobs originally to node i. Each of these classes can decide to route part of his flow to node j, $j \neq i$ through a communication line. We consider the problem of uniqueness of the equilibrium in which each one of the classes minimizes its own average delay. Denote $\mathcal{K} = \cup_{i=1}^{K} \mathcal{K}_i$.

We extend the notation from the previous section in the natural way as follows:

$\phi^{(r)}$ the input arrival rate of class r.

$x_{ii}^{(r)}$ the flow originating at node i by class r that is processed at node i.

$x_{ij}^{(r)}$ the flow originating at node i by class r that is processed at node j.

$\mathbf{x}$ the set of all flows, i.e. $\mathbf{x} = (x_{ij}^{(r)}, r \in \mathcal{K}_i, i = 1, ..., K, j = 1, ..., K)$.

$\mathbf{x}_{ij}$ the set of flows through link (i, j), i.e. $\mathbf{x}_{ij} = (x_{ij}^{(r)}, r \in \mathcal{K}_i))$.

β_i the total load on node i, i.e. $\beta_i = \sum_{j=1}^{K} \sum_{r \in \mathcal{K}_j} x_{ji}^{(r)}$.

λ_{ij} the total flow forwarded from node i to node j.

λ the total network flow, i.e. $\lambda = \sum_{i=1}^{K} \sum_{j \neq i} \lambda_{ij}$.

In case (A) we assume that there is a single communication bus through which all network traffic flow, whereas in case (B) we assume that every two processors, say i and j, communicate through two links: ij and ji. ij corresponds to jobs that are forwarded from processor i to be processed at processor j (and are then returned to processor i through the same link once processed).

We define $\mathcal{L}_A$ and $\mathcal{L}_B$ to be the set of system elements for cases (A) and (B) respectively; $\mathcal{L}_A$ contains the set of nodes as well as the communication bus $\mathcal{L}_A = \{1, ..., K, c\}$ (where c stands for the communication bus). $\mathcal{L}_B$ contains the set of nodes as well as the set of links $\mathcal{L}_B = \{1, ..., K, (ij), i = 1, ..., K, j \neq i\}$.

Associated with each system element l and each class r, there is a cost $J_l^{(r)}(\mathbf{x})$.

We allow for a more general cost $J^{(r)}(\mathbf{x})$ for each class r. We make the following assumptions on the cost, which ensure the existence of an equilibrium.

G1 $J^{(r)}$ is the sum of the local processing cost and the communication cost for class r where the latter are only functions of their local flow rate; i.e. in case (A) we have

$$J^{(r)}(\mathbf{x}) = \sum_{i=1}^{K} J_i^{(r)}(\mathbf{x}) + J_c^{(r)}(\mathbf{x})$$

and for case (B) we have for $r \in \mathcal{K}_i$

$$J^{(r)}(\mathbf{x}) = \sum_{j=1}^{K} \left(J_j^{(r)}(\mathbf{x}) + \sum_{i \neq j} J_{ji}^{(r)}(\mathbf{x}_{ji}) \right).$$

G2 $J_l^{(r)}$ are continuous functions whose range is the nonnegative quadrant and their image is $[0, \infty]$.

G3 $J_l^{(r)}$ are convex functions in the rate sent by class r over the system element l. For example, if l is a processor and $r \in \mathcal{K}_l$ then $J_l^{(r)}$ is assumed to be convex in $x_{ll}^{(r)}$.

G4 Whenever finite, $J_l^{(r)}$ is continuously differentiable in the flow sent by class r to system element l. We denote by $K_l^{(r)}(\mathbf{x})$ the partial derivative of $J_l^{(r)}(\mathbf{x})$ with respect to the flow sent by class r to system element l.

G5 If not all classes have finite cost and one of the classes has infinite cost then it can change its own flow to make this cost finite.

G5 ensures that any equilibrium has finite costs for all players.

Lemma 5.1 *Under conditions G1-G5 there exists an equilibrium.*

Proof: The proof follows from Theorem 1 in [14] (see also [13]). ■

We introduce the following further assumptions on the cost that will be used to establish uniqueness.

(Π_1) $K_l^{(r)}$ is a function of two arguments: (i) the total flow on the system element l, and (ii) the flow that class r sends to node element l. $K_l^{(r)}$ is strictly increasing in each of its two arguments.

We note that the costs in the previous section indeed satisfy conditions (Π_1) and G.

5.2.2 Uniqueness for Case (A)

For a fixed assignment of the other classes, class r is faced with a constrained minimization problem. Its associated Lagrangian is given by

$$\Lambda^{(r)}(\mathbf{x}) = J^{(r)}(\mathbf{x}) - \alpha^{(r)}(\sum_{j=1}^{K} x_{ij}^{(r)} - \phi^{(r)}).$$

$\mathbf{x}^*$ is thus an equilibrium if and only if it satisfies the Kuhn Tucker conditions (see [15] pp. 158-165): There exist some real numbers $\alpha^{(r)}, r \in \mathcal{K}$ such that for $r \in \mathcal{K}_i$, $i = 1, 2$, $j \neq i$:

$$K_i^{(r)}(\mathbf{x}^*) \geq \alpha^{(r)} \text{ and } K_i^{(r)}(\mathbf{x}^*) = \alpha^{(r)} \text{ if } x_{ii}^{(r)} > 0, \tag{1}$$

$$K_c^{(r)}(\mathbf{x}^*) + K_j^{(r)}(\mathbf{x}^*) \geq \alpha^{(r)} \text{ and } K_c^{(r)}(\mathbf{x}^*) + K_j^{(r)}(\mathbf{x}^*) = \alpha^{(r)} \text{ if } x_{ij}^{(r)} > 0; \tag{2}$$

$$x_{ii}^{(r)}, x_{ij}^{(r)} \geq 0, \qquad \sum_{j=1}^{K} x_{ij}^{(r)} = \phi^{(r)}. \tag{3}$$

We shall consider in the rest of this subsection only the case of two nodes (i.e. $K = 2$) and a single bidirectional communicating link between them. (We allow for several classes to arrive at each one of the two nodes.)

Theorem 5.1 *The node optimization has a unique solution under assumptions* (Π_1) *and G1-G5.*

Proof: Let $\hat{\mathbf{x}}$ and $\mathbf{x}$ be two equilibria such that

$$\hat{\lambda} \geq \lambda. \tag{4}$$

Let i be such that $\hat{\beta}_i \geq \beta_i$. We shall show that for all $r \in \mathcal{K}_j$, $j \neq i$,

$$\hat{x}_{jj}^{(r)} \geq x_{jj}^{(r)}, \qquad \text{or equivalently, } \hat{x}_{ji}^{(r)} \leq x_{ji}^{(r)} \tag{5}$$

(the equivalence follows from the constraint (3) on the sum of the flows.) (5) holds trivially if $x_{jj}^{(r)} = 0$, and so we have to check only the case $x_{jj}^{(r)} > 0$. To do so, fix some $r \in \mathcal{K}_j$ and consider the following two subcases. Assume that

(a) $\hat{\alpha}^{(r)} \geq \alpha^{(r)}$. Note that $\hat{\beta}_i \geq \beta_i$ is equivalent to $\hat{\beta}_j \leq \beta_j$ (the equivalence follows from the constraint (3)). Hence

$$K_j^{(r)}(\hat{x}_{jj}^{(r)}, \hat{\beta}_j) \geq \hat{\alpha}^{(r)} \geq \alpha^{(r)} = K_j^{(r)}(x_{jj}^{(r)}, \beta_j) \geq K_j^{(r)}(x_{jj}^{(r)}, \hat{\beta}_j). \tag{6}$$

The equality as well as the first inequality follow from the Kuhn Tucker conditions, whereas the last inequality follows from the monotonicity assumption (Π_1) . Using again the monotonicity assumption (Π_1) , this time for the first argument, we conclude from the fact $K_j^{(r)}(\hat{x}_{jj}^{(r)}, \hat{\beta}_j) \geq K_j^{(r)}(x_{jj}^{(r)}, \hat{\beta}_j)$ that (5) holds.

Thus we try instead of (a):

(b) $\hat{\alpha}^{(r)} \leq \alpha^{(r)}$. (5) holds trivially if $\hat{x}_{ji}^{(r)} = 0$. So it remains to check the case $\hat{x}_{ji}^{(r)} > 0$. We then have for $r \in \mathcal{K}_j$ $(j \neq i)$

$$\begin{aligned} & K_c^{(r)}(x_{ji}^{(r)}, \lambda) + K_i^{(r)}(x_{ji}^{(r)}, \beta_i) \\ \geq\ & \alpha^{(r)} \geq \hat{\alpha}^{(r)} = K_c^{(r)}(\hat{x}_{ji}^{(r)}, \hat{\lambda}) + K_i^{(r)}(\hat{x}_{ji}^{(r)}, \hat{\beta}_i) \\ \geq\ & K_c^{(r)}(\hat{x}_{ji}^{(r)}, \lambda) + K_i^{(r)}(\hat{x}_{ji}^{(r)}, \beta_i). \end{aligned}$$

Here, the first inequality and the equality follow from the Kuhn Tucker conditions, whereas the last inequality follows from the monotonicity of $K_l^{(r)}$ (property (Π_1)). Using again the monotonicity, we conclude that (5) holds in case (b) as well.

We conclude that

$$\sum_{r\in\mathcal{K}_j} \hat{x}_{ji}^{(r)} \leq \sum_{r\in\mathcal{K}_j} x_{ji}^{(r)} \quad (j \neq i). \tag{7}$$

Combining this with $\hat{\beta}_i \geq \beta_i$, we conclude that

$$\sum_{r\in\mathcal{K}_i} \hat{x}_{ii}^{(r)} \geq \sum_{r\in\mathcal{K}_i} x_{ii}^{(r)}.$$

However, since for $i \in \mathcal{K}_i$, $\hat{x}_{ii}^{(r)} + \hat{x}_{ij}^{(r)} = \phi^{(r)}$, it follows that

$$\sum_{r\in\mathcal{K}_i} \hat{x}_{ij}^{(r)} \leq \sum_{r\in\mathcal{K}_i} x_{ij}^{(r)}. \tag{8}$$

Combining this with (7) we conclude that $\hat{\lambda} \leq \lambda$. This contradicts our assumption (4), unless we have equality in (4).

This implies in particular that (5) holds (as we derived above). Now, if for some $r \in \mathcal{K}_j$ (5) holds with strict inequality then (7) and (8) would hold with strict inequality, which would imply that $\hat{\lambda} < \lambda$. But since we established that (4) holds with equality, we conclude that (5) holds with equality. By a symmetric argument we establish (5) also for $r \in \mathcal{K}_i$. We conclude that $\mathbf{x} = \hat{\mathbf{x}}$. ■

An interesting property that can be obtained from the above proof is that not only the Nash equilibrium is unique, but also:

Lemma 5.2 *There are unique Lagrange multipliers for the node optimization (that satisfy the Kuhn-Tucker conditions (1)-(3)) under assumptions* (Π_1) , *G1-G5.*

Proof: Assume that there are two sets of Lagrange multipliers, α and $\hat{\alpha}$ corresponding to the Nash equilibria $\mathbf{x}$ and $\hat{\mathbf{x}}$ (where $\mathbf{x} = \hat{\mathbf{x}}$ due to the uniqueness). Assume that there are some r, j for which $\hat{\alpha}^{(r)} > \alpha^{(r)}$, $r \in \mathcal{K}_j$. It follows that (6) holds with the first inequality being a strict one. Since K_j is strictly monotone in both arguments, it follows from (6) that $\hat{x}_{jj}^{(r)} > x_{jj}^{(r)}$. This contradicts, however, the uniqueness of the Nash equilibrium. ■

5.2.3 Uniqueness of the Node Optimum for Case (B)

We allow in this section for arbitrary numbers of nodes but assume that each class corresponds to a single node.

We impose the following restriction on the model:

(Π_2) : A class i that decides to ship some flow to a node $j \neq i$ should do it using a single hop (a single communication link); it is not possible for it to use two hops (i to k and then k to j), if there is a link that connect nodes i and j directly.

Remark 5.1 Assumption (Π_2) is frequently used in load balancing in distributed computer systems. See [5].

The proof is done by transforming our problem into the following equivalent routing problem for which the uniqueness is known.

Consider a network $\mathcal{G}$ consisting of two nodes: a and b, and of K parallel directed links, all from node a to node b. There are K classes of jobs $i = 1, \dots, K$, all having node a as the source and node b as the destination. A class i of jobs that arrives to node a corresponds to the class of jobs that arrives in our original model to processor i. Hence, the total arrival rate of class r is $\phi^{(r)}$ as in the original problem.

Let $\mathsf{x}_l^{(r)}$ be the rate at which class r sends over link l ($l = 1, 2, \dots, K$). These quantities satisfy the constraints $\mathsf{x}_l^{(r)} \geq 0$, $l = 1, \dots, K$ and $\sum_i \mathsf{x}_i^{(r)} = \phi^{(r)}$. With respect to our original problem, $\mathsf{x}_l^{(r)}$ is interpreted as the rate at which jobs of class r are processed in processor l.

Class r determines $(\mathsf{x}_1^{(r)}, \dots, \mathsf{x}_K^{(r)})$ so as to minimize its cost $\overline{J}^{(r)}(\overline{\mathbf{x}})$, where $\overline{\mathbf{x}} = (\mathsf{x}_l^{(r)}, l \in L, r \in \mathcal{K})$.

Define $\mathsf{x}_l = \sum_{r \in \mathcal{K}} \mathsf{x}_l^{(r)}$. These were denoted in the original model by β_l.

The idea of the equivalent routing model is the following: we eliminate the communicating costs between the processors; this transforms the problem into a standard routing problem between parallel links. In order to compensate for the ij communicating cost that was eliminated we now add it to the cost of class i at processor j. Since for any i and $j \neq i$, only class i used the original communication link ij, the cost of class i in link j in the new problem depends only on the total rate at that link, as well as the rate at which jobs of class i are sent to link j. This is an assumption of type (Π_1) . Hence the results from [13] are applicable.

We now make the above precise. We assume in the new model that $\overline{J}^{(r)}$ is the sum of the link cost functions:

$$\overline{J}^{(r)}(\overline{\mathbf{x}}) = \sum_l \overline{J}_l^{(r)}(\overline{\mathbf{x}}),$$

where $\overline{J}_p^{(r)}(\overline{\mathbf{x}})$ are expressed in terms of the costs (J_l) defined in Subsection 5.2.1 as follows. For any i and any $r \in \mathcal{K}_i$,

$$\begin{aligned} \overline{J}_i^{(r)}(\overline{\mathbf{x}}) &= J_i^{(r)}(\mathsf{x}_i^{(r)}, \mathsf{x}_i), \\ \overline{J}_j^{(r)}(\overline{\mathbf{x}}) &= J_{ij}^{(r)}(\mathsf{x}_j^{(r)}, \mathsf{x}_j) + J_j^{(r)}(\mathsf{x}_j^{(r)}, \mathsf{x}_j) \quad \text{for } j \neq i. \end{aligned}$$

If the costs for the original load balancing problem satisfy assumption (Π_1) , it follows that the costs for the new routing problem also do. The routing problem has a unique Nash equilibrium under assumption (Π_1) . See Theorem 1 in [13] (that theorem states some other assumptions which are not used in its proof). By identifying the decision variables $\overline{\mathbf{x}}$ in the new routing problem with the decision variables $\mathbf{x}$ in the original load balancing problem, we see that the minimization problems faced by each class is the same in both cases, and therefore we conclude that the node optimum in our original problem exists and is unique.

5.3 Some Properties of the Node Equilibrium

We consider node optimization. We use the model, the notation as well as the assumptions G and (Π_1) on the cost, described in the previous subsection 5.2.1. We further assume (Π_2) (Introduced in the previous subsection).

Assume further that the cost $J_l^{(r)}$ of the lth system element is given by the product of the rate of jobs that class r ships through that element and the average delay T_l in this element. For example, for $r \in \mathcal{K}_i$ we have

$$J_i^{(r)}(x_l) = x_{ii}^{(r)} T_i(\beta_i).$$

We further assume that, for each l, the partial derivative of T_l with respect to the total flow through that element is strictly positive. In other words, for every node i,

$$\partial T_i(\beta_i)/\partial \beta_i > 0, \tag{9}$$

with similar expressions for the communication costs.

Theorem 5.2 *Consider either case* (A) *or case* (B). *Let* $\phi^{(r)} > 0$ *for all classes.*
(i) *Assume that at equilibrium, all the traffic of some class* $r \in \mathcal{K}_i$ *arriving to node* i *is routed away from that node (i.e.* $x_{ii}^{(r)} = 0$*).*
Let j *be another node to which some positive flow is routed by class* r.
Consider now any $s \in \mathcal{K}_j$. *If* s *sends some positive amount of flow to node* i *then it also processes some positive amount of flow at node* j.
(ii) *Consider the case of two nodes* $K = 2$, *and assume that* $\phi^{(r)} \geq \phi^{(s)}$, $r \in \mathcal{K}_1, s \in \mathcal{K}_j$.
Assume that class s *sends at equilibrium some flow to node 1.*
Then at equilibrium, class r *sends some strictly positive flow to be processed in node* 1.

Proof: We shall prove for case (B); the proof for case (A) is the same except that the K_{ij} below should be replaced by K_c.
(i) Using the Kuhn Tucker conditions for class r and the fact that $x_{ii}^{(r)} = 0$, we have

$$\begin{aligned} K_i^{(r)}(x_{ii}^{(r)}, \beta_i) &= T_i(\beta_i) \geq \alpha^{(r)} = K_{ij}^{(r)}(x_{ij}^{(r)}, \lambda_{ij}) + K_j^{(r)}(x_{jj}^{(r)}, \beta_j) \\ &= K_{ij}^{(r)}(x_{ij}^{(j)}, \lambda_{ij}) + x_{ij}^{(r)} \partial T_j(\beta_j)/\partial x_{jj}^{(r)} + T_j(\beta_j). \end{aligned}$$

Combining this with Assumption (9) we conclude that

$$T_i(\beta_i) > T_j(\beta_j).$$

Assume that all traffic of source s is routed away from some source s that sends positive flow to node i. By using the Kuhn Tucker conditions for class s we get by the same arguments as above $T_i(\beta_i) < T_j(\beta_j)$, which gives a contradiction. This establishes the proof.
(ii) Assume $\phi^{(r)} \geq \phi^{(s)}$, and assume that class r sends all its traffic to node 2. We have by (10)

$$\begin{aligned} T_1(\beta_1) &\geq \alpha^{(r)} \geq K_{12}^{(r)}(x_{12}^{(r)}, \lambda_{12}) + \phi^{(r)} \partial T_2(\beta_2)/\partial x_{22}^{(r)} + T_2(\beta_2) \\ &> K_{12}^{(r)}(x_{12}^{(r)}, \lambda_{12}) + x_{22}^{(s)} \partial T_2(\beta_2)/\partial x_{22}^{(s)} + T_2(\beta_2) \\ &= K_{12}^{(r)}(x_{12}^{(r)}, \lambda_{12}) + K_2^{(s)}(x_{22}^{(s)}, \beta_2) \geq K_2^{(s)}(x_{22}^{(s)}, \beta_2) \\ &\geq \alpha^{(s)} \geq T_1(\beta_1) \end{aligned} \tag{10}$$

The inequality (10) follows since $\phi^{(r)} \geq \phi^{(s)} > x_{22}^{(s)}$ and since $\partial T_2(\beta_2)/\partial x_{22}^{(r)} = \partial T_2(\beta_2)/\partial x_{22}^{(s)} > 0$. The inequality before the last follows by the Kuhn Tucker conditions. The last inequality is obtained similarly to (10) (and using the fact that class s sends nonzero flow to node 1). We thus obtained a contradiction, which establishes the proof. ■

6 Concluding Remarks

In this paper, we showed the existence of some paradoxical and awkward behaviors of the Nash equilibrium in a model of load balancing for distributed computer systems. We secured uniqueness of the overall and Wardrop optima. We extended the previous results and gave proofs of the uniqueness of the Nash equilibrium for the model. We saw that the Nash noncooperative and independent decision may sometimes lead to the degradation of performance for each member. Please note that this class optimum is with respect to a case of nonsymmetric user classes, whereas symmetric user classes are treated in [10, 11]. Such a paradoxical behavior does not occur for the overall and Wardrop optimum in the same setting of the model. That may imply that the Nash equilibrium may have more complicated characteristics than the overall optimum and further the Wardrop equilibrium.

References

[1] J. E. Cohen and F. P. Kelly, "A PARADOX of Congestion in a Queuing Network," *J. Appl. Prob. 27*, pp.730-734, 1990.

[2] J. E. Cohen and C. Jeffries, "Congestion Resulting from Increased Capacity in Single-Server Queueing Networks," *IEEE/ACM Trans. on Networking 5*, 2, April pp.1220-1225, 1997.

[3] B. Calvert, W. Solomon and I. Ziedins, "Braess's Paradox in a Queueing Network with State-Dependent Routing," *J. Appl. Prob. 34*, pp.134-154, 1997.

[4] A. Haurie and P. Marcottet, "On the Relationship between Nash-Cournot and Wardrop Equilibria," *Networks 15*, 1985, pp.295-308.

[5] H. Kameda, J. Li, C. Kim and Y. Zhang, *Optimal Load Balancing in Distributed Computer Systems*, Springer, 1996.

[6] H. Kameda, T. Kozawa and J. Li, "Anomalous Relations among Various Performance Objectives in Distributed Computer Systems," *Proc. 1st World Congress on Systems Simulation*, IEEE, 1997, pp.459-465.

[7] H. Kameda and Y. Zhang, "Uniqueness of the Solution for Optimal Static Routing in Open BCMP Queueing Networks," *Mathematical and Computer Modelling 22*, 10-12, pp.119-130, 1995.

[8] C. Kim and H. Kameda, "An Algorithm for Optimal Static Load Balancing in Distributed Computer Systems," *IEEE Trans. Compt. 41*, 3, 1990, pp.381-384.

[9] C. Kim and H. Kameda, "Optimal Static Load Balancing of Multi-class Jobs in a Distributed Computer System," *Proc. 10th Intl. Conf. Distributed Comput. Syst.*, IEEE, 1990, pp.562-569.

[10] Y. A. Korilis, A. A. Lazar and A. Orda, "Capacity Allocation under Non-Cooperative Routing", *IEEE Trans. Automatic Control 42*, 3, March 1997, pp.309-871.

[11] Y. A. Korilis, A. A. Lazar and A. Orda, "Avoiding the Braess Paradox in Noncooperative Networks", *Proc. IEEE Conference on Decision & Control (San Diego)*, 1997, pp.864-878.

[12] J. Li and H. Kameda, "Load Balancing Problems for Multiclass Jobs in Distributed/Parallel Computer Systems," *IEEE Trans. Compt. 47*, 3, 1998, pp.322-1998.

[13] A. Orda, R. Rom and N. Shimkin, "Competitive Routing in Multiuser Communication Networks," *IEEE/ACM Trans. on Networking 1*, 1993, pp.614-627.

[14] J. B. Rosen, "Existence and Uniqueness of Equilibrium Points for Concave N-person Games", *Econometrica 33*, pp. 153-163, 1965.

[15] J. F. Shapiro, *Mathematical Programming, Structures and Algorithms*, J. Wiley and Sons, 1979.

[16] A. N. Tantawi and D. Towsley, "Optimal Static Load Balancing in Distributed Computer Systems," *J. ACM 32*, 2, pp.445-465, Apr, 1985.

[17] Y. Zhang, H. Kameda and K. Shimizu, "Parametric Analysis of Optimal Load Balancing in Distributed Computer Systems," *Journal of Information Processing 14*, 4, 1992, pp.433-441.

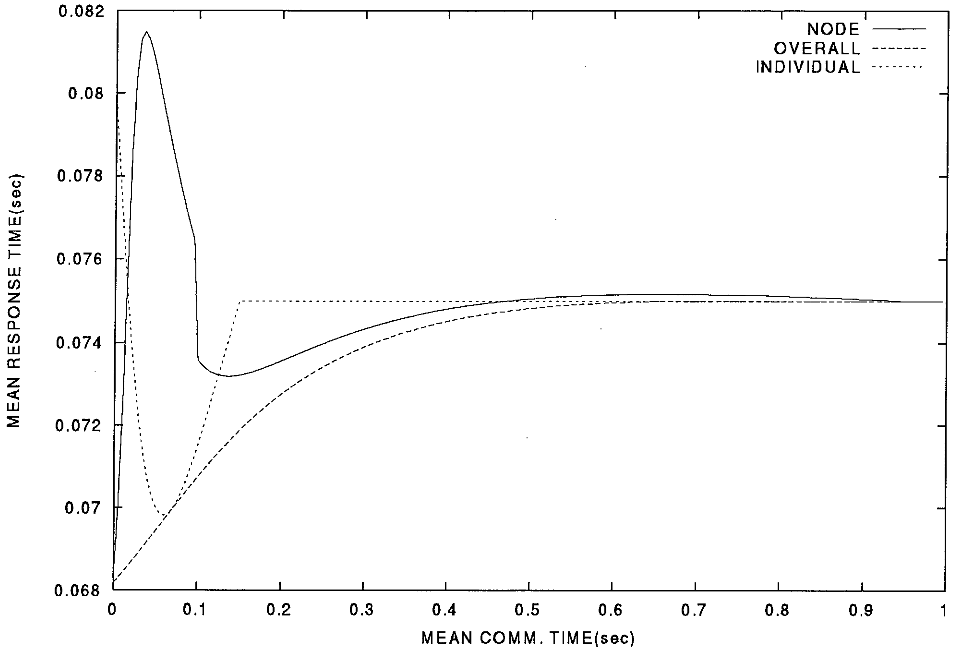

Figure 3: The overall mean response time (case (A)).

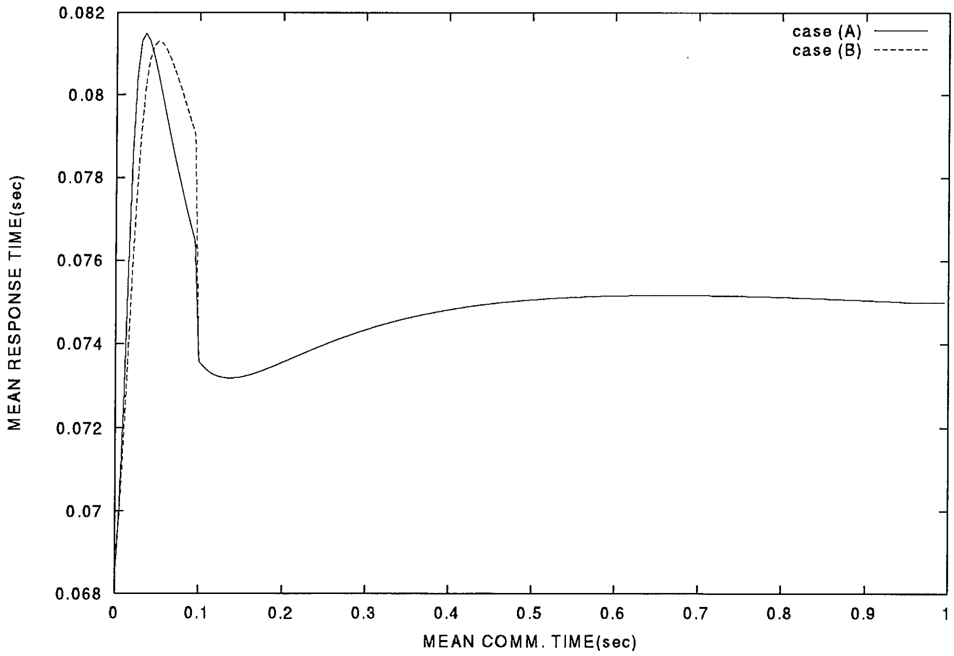

Figure 4: The overall mean response time in node optima (case (A) and (B)).

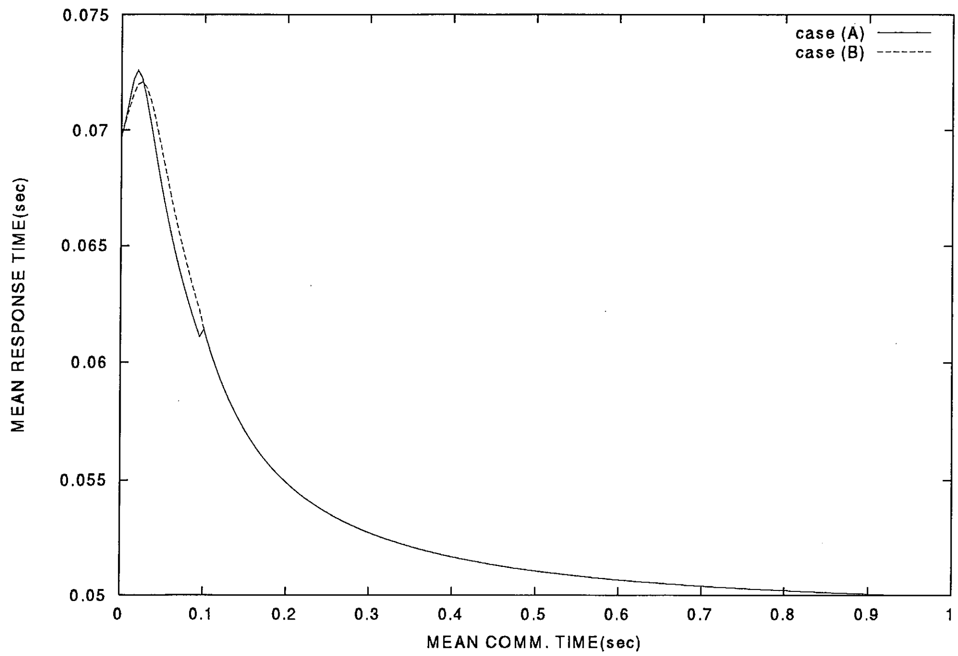

Figure 5: The mean response time of a job that arrives at node 1 in node optima (cases (A) and (B)).

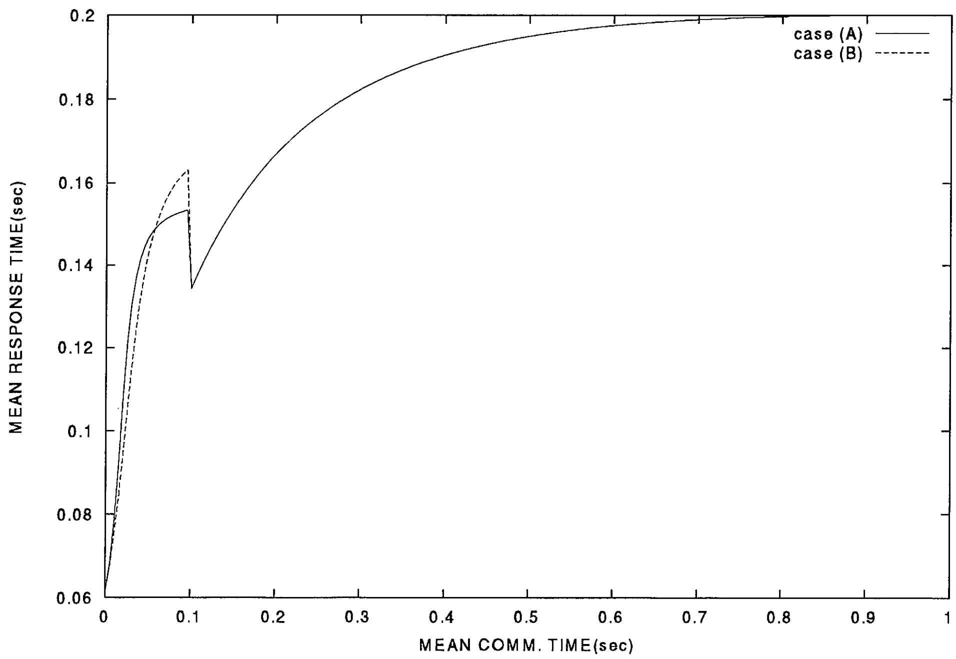

Figure 6: The mean response time of a job that arrives at node 2 in node optima (cases (A) and (B)).

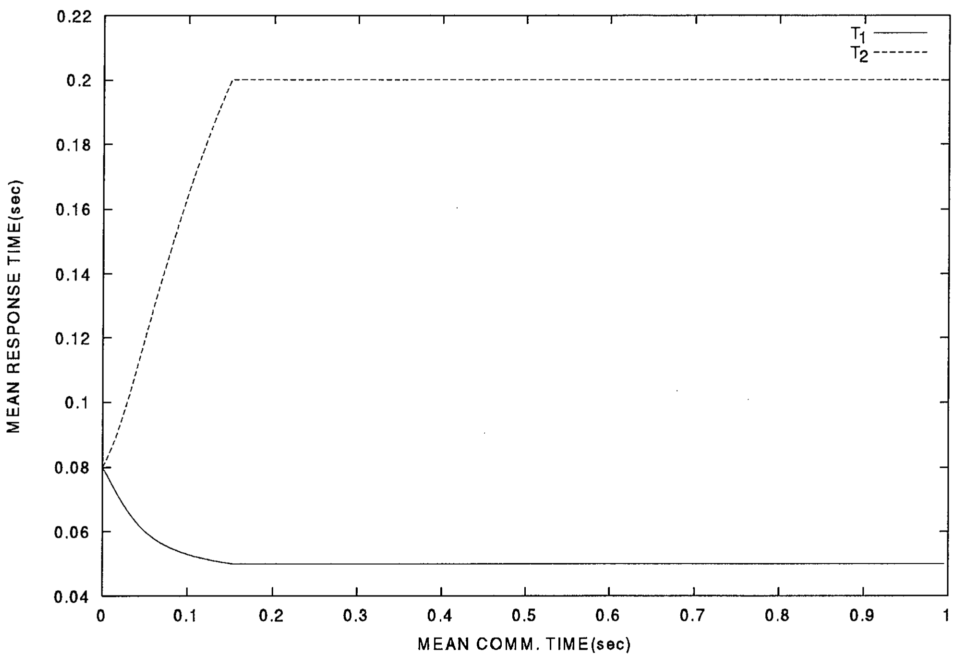

Figure 7: The mean response times, T_1 and T_2, of a job that arrives at nodes 1 and 2 in individual optima, respectively. They have the same values for cases (A) and (B).