REIHE INFORMATIK
23/95

Duration of Asynchronous Operations in Distributed Systems

M. Makhaniok, R. Männer
Universität Mannheim
Seminargebäude A5
D-68131 Mannheim

DURATION OF ASYNCHRONOUS OPERATIONS IN DISTRIBUTED SYSTEMS¶

M. Makhaniok[1], R. Männer[2]

1) Institute of Engineering Cybernetics, Academy of Sciences of Belarus,
Surganov St.. 6, SU-220012 Minsk, Republic of Belarus
2) Lehrstuhl für Informatik V, Universität Mannheim
D-68131, Mannheim, A5/B6, Germany

Abstract. A distributed asynchronous system is investigated. Its processing elements execute common operations concurrently and distributively. They are implemented as combinatorial circuits and exchange data via open collector bus lines. A method is presented to identify and to minimize the duration of an operation and therefore to increase the performance of the system. No hardware modifications are required.

Introduction.

A distributed asynchronous system (DAS) consists of a number of processing elements (PEs) that concurrently execute common operations, e.g., search, comparison or sorting. The PEs are implemented as combinatorial circuits and exchange data via open collector bus lines. Generally every PE can participate or not in an operation, depending on the state of the system, the type of operation, and the actual operands. In this case the problem arises how to determine the duration of the operation, i.e., the moment when all signals on the bus lines of the DAS settled. The problem is caused by possibly different speeds of PEs, the unknown set of participants of the operation, and their unknown input data (operands). Assuming a high number of PEs and a short duration of the operation, it is not possible to exchange such data before the start of an operation. Instead an asynchronous type of operation is advantageous whose duration is determined, e.g., by the PE that terminates its own part of the operation latest. However a PE cannot determine the exact duration of the operation by its own since it depends generally on unknown parameters. It therefore has to assume the worst case. This causes unnecessary waiting time of PEs. An important example is the distributed arbitration operation on multiprocessor buses like Futurebus+ [1] which is executed by a DAS. Here this DAS determines the next bus master, i.e., the PE having the highest priority among all participants of the operation.

1. Duration of an Asynchronous Operation.

To analyze the problem of finding the duration of an asynchronous operation we introduce the following notations. We consider a DAS consisting of a set **Z={1,...,E}** of PEs that execute an operation **W** (see Fig. 1).

¶ This work has been supported by the Deutsche Forschungsgemeinschaft (DFG) under grants Ma 1150/8-1 and 436-WER-113-1-0 (438 113/117/0).

All inputs and outputs of the DAS are binary. The PEs operate asynchronously and can have different speeds. They are interconnected by N buses L_i, $i=1,2,\ldots,N$, and by three synchronization lines [2,3] (in Fig. 1, for clarity, these lines have been omitted). Each j-th PE, $j\in\{1,\ldots,E\}$, is assembled as a pipeline of N units **Unit (j,i)** that are combinatorial circuits. Every two consecutive units of the same PE are connected by a set of individual control lines **Ctrl (j,i)**. All circuits of the same pipeline stage $i\in\{1,2,\ldots,N\}$ execute together one step of the operation. They use common input data from bus L_{i-1} as well as local input data **Data (j,i)**, and assert output data onto bus L_i. This bus consists of open collector lines that are used for exchanging data and for computing a bit-wise global **OR** function of the binary words that are distributed over all PEs.

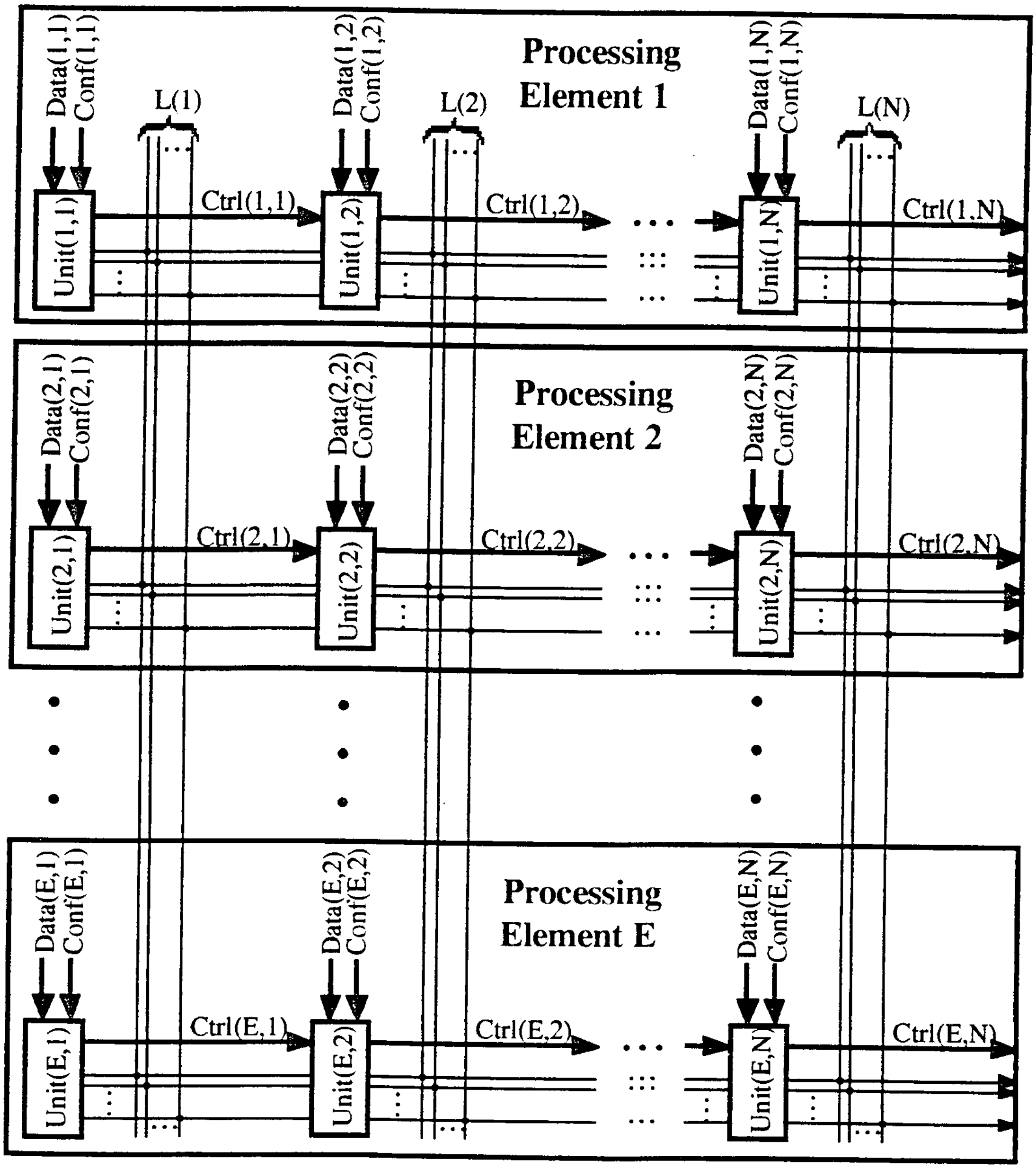

Fig. 1. Interconnection of the units of the PEs of a DAS.

Thus W is a multi-step operation and every unit i of a PE executes a corresponding step. Every step has two parts, a distributed part that is executed distributively by all PEs, and a global part following it (Fig. 2). It is called global since partial results from all PEs are combined here. In the special case considered, all PEs, together with the interconnecting lines used for the global operation of every step, form one combinatorial circuit. This means signals asserted at the input of the PEs are combined – maybe in a complicated way but without feedback loops – to output signals. Since all signals in the circuit propagate forward only, the state of these signals settles in the forward direction. The signals corresponding to the global part of step 1, e.g., will have stabilized at a certain time. Only from then on the combination of these signals in step 2 – although computed all the time – is useful. In this sense a step is started automatically when the output computed in the previous step is available. Thus no synchronization is required during the execution of an operation, only after every operation. Therefore PEs can start the following operation only after the previous one is completed by all of them. Identification of this moment is a complicated problem in asynchronous distributed systems.

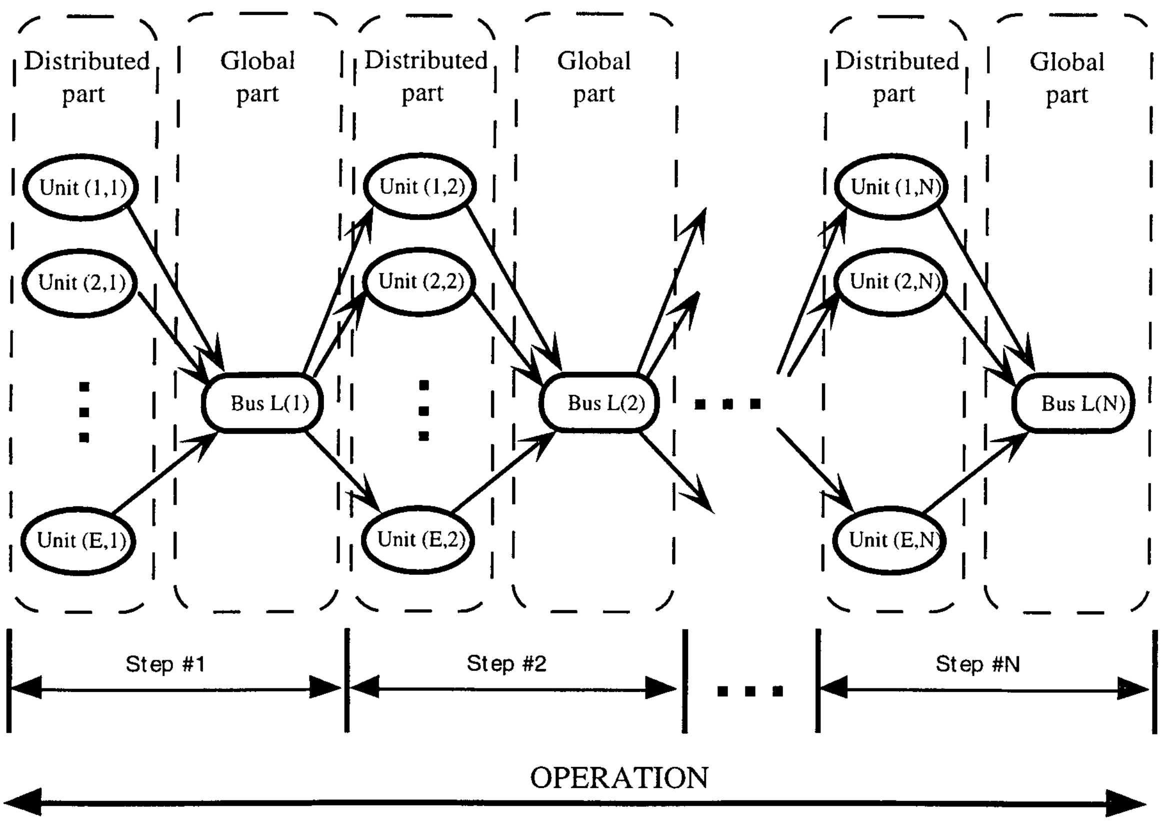

Fig. 2. A multi-step operation.

Let θ_i be the bus-propagation delay, i.e, the time to compute the global part of the **i**-th step of the operation, and τ_j be the individual gate delay to execute any distributed part of each step of the operation by **Unit (j,i)**. We define $\mathbf{d_{j,i} = 1}$, if **Unit (j,i)** takes longest to complete the execution of the distributed part of the **i**-th step[1], $\mathbf{d_{j,i} = 0}$ else, so that $\sum_{j \in Z} \mathbf{d_{j,i}} = \mathbf{1}$. Then the

---

[1] If more than one PE needs the same time to complete the **i**-th step then any but only one PE_j can be chosen for which $\mathbf{d_{j,i} = 1}$.

execution time of a step $i \in \{1,...,N\}$ is the sum of the time θ_i required to compute the i-th global part plus the time $\sum_{j\in Z}\tau_j d_{j,i}$ required to execute the i-th distributed part. The duration of the operation $T(W,Z)$ is the sum of the execution times of all steps, i.e.,

$$T(W,Z)=\sum_{j\in Z}c_j\tau_j+\sum_{i=1,...,N}\theta_i, \qquad (1)$$

where $c_j=\sum_{i=1,...,N}d_{j,i}$ is the number of the individual delays introduced in the operation by the j-th PE.

Two cases are distinguished. In the first case each PE has no information about the speed of the others so that it has to be assumed

$$0<\tau_j<\infty. \qquad (2)$$

In the second case the PEs can exploit the fact that their speed is bounded, i.e., that

$$T_L \le \tau_j \le T_H. \qquad (3)$$

2. Minimal Upper Bound for the Duration of an Asynchronous Operation. Case $0<\tau_j<\infty$.

The synchronization of the operation in such a DAS is done in accordance with [2]. After all PEs have been notified of the beginning of the operation **W**, they assert logic "**1**" signals onto a wired-OR logic synchronization line. During execution of **W** each PE_j waits for a certain individual waiting time $t(j,W)$ before it removes its "**1**" from the line. When the synchronization line changes to the "**0**" state all PEs recognize the end of the operation. Therefore, the total waiting time $t(W)$ for the DAS to identify the end of the operation **W** is

$$t(W)=\max_{j\in Z} t(j,W). \qquad (4)$$

It is obvious that the operation will be finished correctly if this total waiting time is not less than the duration of the operation. This means that the condition

$$t(W)\ge T(W,Z) \qquad (5)$$

must be true for all possible τ_j satisfying either **(2)** or **(3)**. Therefore $t(W)$ is an upper bound for the duration of the operation.

Problem of Investigation. In an asynchronous DAS each PE_j, $j\in Z$, has only information about its own characteristics and the ones specified for the whole system. Thus, a PE knows all the coefficients c_j, i.e., the type of the operation[2], all times θ_i for executing global parts, and its

---

[2] The dependence of the coefficients c_j on a certain type of operation is presented in [3, 4].

individual delay τ_j. However, it generally has no information about the speed of other PEs, i.e., their individual gate delays. Then, according to relations **(2)** and **(3)**, every PE_j has to determine its individual waiting time **t(j,W)** independently of other PEs so that condition **(5)** is always fulfilled for all possible values τ_j. Moreover the individual waiting times of all PEs should be chosen so that the total waiting time is minimized. It seems to be impossible that any PE solves this problem by its own. However if all PEs use the method introduced below, they can determine the moment when **W** will be finished. In accordance with [5] we formulate below the task of minimizing the total waiting time. We fix the type of the operation and introduce the dimensionless coefficient $\mathbf{n_j}$ that allows to measure time locally so that

$$\mathbf{n_j \tau_j = t(j,W) - \sum_{i=1,\ldots,N} \theta_i}, \text{ where } \mathbf{j=1,\ldots,E}. \qquad (6)$$

Let **E** be the number of PEs participating in the operation **W**. Then, due to the relations **(1)**, **(4)-(6)**, one can guarantee that the operation will always be finished correctly if at least one relation of the system

$$\begin{array}{l} \mathbf{n_{j*}\tau_{j*} \geq \sum_{j\in Z} c_j \tau_j}, \\ \mathbf{j*=1,\ldots,E}, \end{array} \qquad (7)$$

is fulfilled for **(2)** or **(3)** respectively. Thus the solution of the problem is reduced to the determination of the coefficients $\mathbf{n_j}$. It is not easy in general to calculate any specific coefficient $\mathbf{n_{j*}}$ using the system **(7)** because $\mathbf{n_{j*}}$ depends on the unknown values τ_j, $j \in \{Z \backslash j*\}$. However it is possible to simplify **(7)** using the following theorem.

Theorem 1. At least one relation of the system **(7)** is fulfilled for arbitrary values τ_j, $\mathbf{0<\tau_j<\infty}$, $\mathbf{j=1,\ldots,E}$, if and only if

$$\mathbf{\sum_{j\in Z} \frac{c_j}{n_j} \leq 1}. \qquad (8)$$

Proof (sufficient). It has to be shown that, if **(8)** is true, there exists a $\mathbf{j* \in Z}$, so that $\mathbf{n_{j*}\tau_{j*} \geq \sum_{j\in Z} c_j \tau_j}$.

Let us assume the contrary. Then for every $\mathbf{j* \in Z}$ we have $\mathbf{n_{j*}\tau_{j*} < \sum_{j\in Z} c_j \tau_j}$. From this inequality we can derive $\dfrac{\tau_{j*}}{\sum_{j\in Z} c_j \tau_j} < \dfrac{1}{n_{j*}}$ and therefore $\dfrac{c_{j*}\tau_{j*}}{\sum_{j\in Z} c_j \tau_j} < \dfrac{c_{j*}}{n_{j*}}$.

If we sum both sides over $\mathbf{j* \in Z}$, we obtain the contradiction

$$\frac{\sum_{j^*\in Z} c_{j^*}\tau_{j^*}}{\sum_{j\in Z} c_j\tau_j} = 1 < \sum_{j^*\in Z} \frac{c_{j^*}}{n_{j^*}} = \sum_{j\in Z} \frac{c_j}{n_j}.$$

Proof (necessary). We again will disprove the contrary. Assume that expression **(8)** is false and that for every $\tau_j > 0$ at least one of the expressions **(7)** is true. This means that the system

$$n_{j^*}\tau_{j^*} - \sum_{j\in Z} c_j\tau_j < 0 \text{ for every } j^*\in Z,$$

$$-\tau_{j^*} < 0 \text{ for every } j^*\in Z$$

has no solution with respect to τ_{j^*}. It is known from the Voronoi criterion [6] that there always exists a solution $x_{j^*} > 0$, $y_{j^*} > 0$ for the corresponding system of linear equations

$$c_{j^*}\sum_{j\in Z} x_j - n_{j^*}x_{j^*} + y_{j^*} = 0, \text{ for } j^*\in Z.$$

From this system one finds $\left(1 - \sum_{j^*\in Z}\frac{c_{j^*}}{n_{j^*}}\right)\sum_{j\in Z} x_j = \sum_{j^*\in Z}\frac{y_{j^*}}{n_{j^*}}$.

If we take now, as assumed, $\sum_{j^*\in Z}\frac{c_{j^*}}{n_{j^*}} > 1$ then at least one value x_{j^*} or y_{j^*} is less than zero. This contradicts the Voronoi criterion [6].
Q.E.D.

Example. Let us consider a distributed arbitration system that consists of 64 PEs [2-4]. For a binary representation of arbitration priorities, the PEs can therefore use **6**-bit arbitration words. To reduce the arbitration time, one has to minimize the average number of individual gate delays of all PEs which can be expressed by

$$F(n_0,\ldots,n_{63}) = \frac{\sum_{j=0,\ldots,63} n_j}{64}.$$

Here it is assumed that every PE becomes bus master with the same probability. $F(n_0,\ldots,n_{63})=6$ is obtained if we take all coefficients n_j equal to **6** according to the standard setting of Taub [3,7,8]. If, however, **F** is minimized under the boundary conditions **(8)** so that the coefficients are integer, we get **F=4.6** if we set[3]

n_j = **6** for **j** = **16, 17,..., 31, 36, 37, 38, 39, 41, 42,**
= **5** for **j** = **8, 9,..., 15, 40, 43,..., 47, 50, 51, 53,**

[3] Here to each PE_j a binary arbitration word A_j is assigned, e.g., the binary word ‹**101010**› corresponds to PE_{42}. The duration of the operation depends on the number of **1**- and **0**-intervals in the arbitration words of the PEs participating in the operation [3,4].

$= \mathbf{4}$ for $\mathbf{j} = \mathbf{4,\ldots, 7, 33, 34, 49, 52, 54, 55, 57, 58}$,

$= \mathbf{3}$ for $\mathbf{j} = \mathbf{2, 3, 59, 61}$,

$= \mathbf{2}$ for $\mathbf{j} = \mathbf{1, 32, 48, 56, 60, 62}$,

$= \mathbf{1}$ for $\mathbf{j} = \mathbf{63}$,

$= \mathbf{0}$ for $\mathbf{j} = \mathbf{0}$.

This example shows that about 70% of all coefficients $\mathbf{n_j}$ can be less than **6**.

Unexpectedly, the optimal result is achieved if we allow one coefficient to be **higher** than **6**, i.e.,

$\mathbf{n_j}$ $= \mathbf{9}$ for $\mathbf{j} = \mathbf{42}$,

$= \mathbf{6}$ for $\mathbf{j} = \mathbf{21, 37, 41, 43, 45, 53}$,

$= \mathbf{3}$ for $\mathbf{j} = \mathbf{2, 3, 59, 61}$,

$= \mathbf{2}$ for $\mathbf{j} = \mathbf{1, 32, 48, 56, 60, 62}$,

$= \mathbf{1}$ for $\mathbf{j} = \mathbf{63}$,

$= \mathbf{0}$ for $\mathbf{j} = \mathbf{0}$,

$= \mathbf{4}$ else.

In this case the average number of individual gate delays for all PEs will be only **F=3.9**.

By applying Theorem 1 it is therefore possible to reduce the average duration of the arbitration operation by 30% without any change of the hardware of the DAS.

3. Minimal Upper Bound for the Duration of an Asynchronous Operation. Case $T_L \le \tau_j \le T_H$.

Obviously condition **(8)**, obtained under the assumption $0<\tau_j<\infty$, yields an unnecessary high value of the upper bound $\mathbf{t(W)}$ for the duration of operations. This range of τ_j values, introduced in [7,8], is only of theoretical interest, because in all practical cases the τ_j values are bounded. In order to speed up the DAS we consider a new system of boundary conditions **(3)** where $\mathbf{T_L}$ and $\mathbf{T_H}$ have the same lower and upper values for all PEs.

Below we prove a theorem that allows to find all possible $\mathbf{t(j,W)}$ satisfying the condition $\underset{j\in Z}{\mathbf{Max}}\ \mathbf{t(j,W)} \ge \sum_{j\in Z} c_j \tau_j$ and that, therefore, guarantees a correct result of the operation.

Theorem 2. If at least one of the relations **(7)** is true, it is a necessary and sufficient condition for the coefficients $\mathbf{n_j}$ that

$$\sum_{j\in Z}\left(\frac{c_j\beta_j}{n_j} + \frac{c_j T_H(1-\beta_j)}{T_L \underset{j'\in Z}{\mathrm{Max}}\ n_{j'}}\right) \le 1, \qquad (9)$$

where $\beta_j =$ $\mathbf{0}$, if $\mathbf{Max}\ \mathbf{n_{j'}} \ge \dfrac{n_j T_H}{T_L}$,

$\mathbf{1}$, else.

Proof. We fix the operation **W**. Without restriction of arguments we enumerate the elements of the set **Z** such that $n_1 \geq n_2 \geq ... \geq n_E$. Then the conditions **(7)** and **(9)** can be written as

$$\sum_{j \in Z} c_j \tau_j - n_j \tau_j = F_j(\tau_1, ..., \tau_E) \leq 0 \qquad (7')$$

for $j=1,...,E$, $T_L \leq \tau_j \leq T_H$, and

$$\sum_{j \in Z} \left(\frac{c_j \beta_j}{n_j} + \frac{c_j T_H (1 - \beta_j)}{T_L n_1} \right) \leq 1, \qquad (9')$$

where $\beta_j = 0$, if $n_1 T_L \geq n_j T_H$,
$= 1$, else.

We define **N** as the maximal number **j** that satisfies the condition $\beta_j = 1$, and prove the following assertion.

Lemma. It is a necessary and sufficient condition for **(7´)** to be fulfilled at least for one **j**, that the system

$$F_v(\tau_1, ..., \tau_E) > 0, \quad v = 1, 2, ..., N \qquad (7'')$$

is inconsistent.

Proof of Lemma. We consider only the condition $1 \leq N < E$, since the case $N = E$ is trivial. From the inconsistency of the system

$$F_j(\tau_1, ..., \tau_E) > 0, \;\; j = 1, 2, ..., E \qquad (10)$$

follows the inconsistency of the condition of the system **(7´´)**. This is true because each inequality of the system **(7´´)** belongs to the system **(10)**.

On the other hand, suppose that conditions **(10)** are inconsistent and a solution for the system **(7´´)** exists. In this case there also exists an index $j=w$ so that for $N<w\leq E$ the relation $F_w(\tau_1, ..., \tau_E) \leq 0$ is true. But according to the assumptions of our theorem, **N** is the maximal value of the index **j** which satisfies $\beta_j = 1$ and $n_1 T_L \geq n_w T_H$. From these facts and from the obvious relations $n_w \tau_w \leq n_w T_H$ and $n_1 \tau_1 \geq n_1 T_L$, follows that

$$F_1(\tau_1, ..., \tau_E) = F_w(\tau_1, ..., \tau_E) + (n_w \tau_w - n_1 \tau_1) \leq 0.$$

Lemma is proven.

Continuing the proof of the theorem, we consider for $v=1,2,...,N$ the auxiliary system

$$E_v(\tau_1, ..., \tau_N) = F_v(\tau_1, ..., \tau_N, T_H, ..., T_H) = \sum_{j=1,...,N} c_j \tau_j + \sum_{g=N+1,...,E} c_g T_H - n_v \tau_v > 0. \qquad (7''')$$

Let $\tau_v = b_v$, $v=1,2,...,N$, be a solution of the system (7´´´). Then, obviously,

$$\tau_j = b_j \text{ for } j=1,2,...,N,$$
$$= T_H \text{ for } j=N+1,N+2,...,E$$

is a solution of (7´´). On the other hand, if $\tau_j = b_j$, $j=1,2,...,E$, is a solution of the system (7´´) then the expression

$$E_v(\tau_1,...,\tau_N) = F_v(b_1,...,b_E) + \sum_{g=N+1,...,E} c_g(T_H - b_g) > 0$$

is true for every $v \in \{1,2,...,N\}$. Hence, $\tau_v=b_v$ gives a solution of the system (7´´´). Thus we proved that the system (7´´´) is inconsistent if the conditions of the theorem are true and vice versa. According to that we shall prove the theorem with respect to the system (7´´´).

Proff (necessary). We use **N** as the maximal index **v** and assume that **(9)** is not true. Then, one can conclude that

$$\sum_{v=1,...,N}\left(\frac{c_v}{n_v}\right) + \frac{T_H}{T_L n_1}\sum_{g=N+1,...,p} c_g > 1.$$

However, in this case the values τ_v, which can be defined as $\tau_v = \frac{T_L n_1}{n_v}$ for $v=1,2,...,N$, represent a solution of the auxiliary system (7´´´). This follows from

$$E_v\left(\frac{T_L n_1}{n_1},...,\frac{T_L n_1}{n_N}\right) = \sum_{v=1,...,N}\left(\frac{c_v n_1 T_L}{n_v}\right) + \left(\sum_{g=N+1,...,E} c_g T_H\right) - \frac{n_v n_1 T_L}{n_v} =$$

$$= n_1 T_L\left(\sum_{v=1,...,N}\left(\frac{c_v}{n_v}\right) + \frac{T_H}{T_L n_1}\left(\sum_{g=N+1,...,E} c_g\right) - 1\right) > 0$$

and $T_L \le \frac{T_L n_1}{n_v} \le T_H$ for all $v \in \{1,2,...,N\}$.

From here we get a contradiction to the conditions **(9)** of the theorem.

Q.E.D.

Proof (sufficient). Let the conditions **(9)** be true. We have to prove that (7´´´) is inconsistent. We note that the equations **(9)** and **(8)** are identical, and the conclusion is obvious if **N=E**. For **N<E** we introduce the notation $b_v = \dfrac{T_H \sum_{g=N+1,...,E} c_g}{n_v\left(1 - \sum_{v'=1,...,N}\frac{c_{v'}}{n_{v'}}\right)}$, $v=\{1,2,...,N\}$.

It is simple to check that the equality $E_v(b_1,...,b_N)=0$ is fulfilled. If we choose $\tau_1=b_1+\varepsilon_1$, $\tau_2=b_2+\varepsilon_2$,...,$\tau_N=b_N+\varepsilon_N$, then

$$E_v(b_1,\dots,b_N) = \sum_{v'=1,\dots,N} c_{v'}\varepsilon_v - n_v\varepsilon_v, \; v = 1,2,\dots,N.$$

Now only the following three cases are possible.

1. If $\varepsilon_v < 0$ for every v then there always exists such a value ε_v that $\tau_1=b_1+\varepsilon_1<T_L$ and the system **(7´´´)** is inconsistent. This is true due to the relation $b_1<T_L$ and the inequality **(9´)**.
2. If $\varepsilon_v\le 0$ there exists at least one value v^* of the index v such that $\varepsilon_{v^*}=0$, therefore $E_{v^*}(\tau_1,\dots,\tau_N)\le 0$ and **(7´´´)** is inconsistent again.
3. Finally, if $\varepsilon_r >0$, where $r \in U$, $U \subseteq Z$, then we can create the system

$$\sum_{r\in U} c_r\varepsilon_r - n_s\varepsilon_s > 0, \; s \in U,$$

$$\varepsilon_s > 0, \; s \in U,$$

which is also inconsistent according to the equivalence of the conditions **(7)** and **(8)** (see Theorem 1). Therefore, in all cases there is no solution of the system **(7´´´)**.

Q.E.D.

In formula **(9´)** the parameters defining the total waiting time are

1) The ratio $\frac{T_H}{T_L}$ of the higher to the lower bound of the values τ_j.

It is important to note that there is no dependence on time directly, but only on the time ratio. If PEs with similar characteristics are used, conditions **(9)** permit to take into account the inevitable spread of parameters which determine their speed. If, however, we consider the ideal system where $T_H = T_L$, then the individual waiting time can be identified by each participant as $n_j = \sum_{j\in Z} c_j$ in order to terminate the operation correctly.

2) The system used for coding data.

It is very important to find an optimal coding system for the data since this allows to increase the performance of a DAS without hardware modifications. The optimal system defines minimal values for the coefficients c_j in respect to the given operation or a group of operations. As an example, the algorithm to construct the optimal codes for the arbitration operation has been discussed in [4]. It allows to reduce the total waiting time by, e.g., 30-50% without modifications of the hardware.

3) The set **Z** of PEs participating in the operation is characterized by coefficients c_j.

One example for this dependency is the increase of the total waiting time $t(W)$ with an increasing number of PEs participating in the operation. It converges to the value given by the formula $\max_{j\in Z} t(j,W)$.

Therefore for a high number of PEs it might be reasonable to change the synchronization procedure [2], i.e., to allow any PE to force the end of an operation as soon as it determines that all other participants cannot longer affect the result of the operation. Conditions for such type of identification will be discussed in the next section.

4. Minimization of the Average Value of the Maximal Upper Bound of the Duration of an Asynchronous Operation.

Below the optimization task will be formulated for the case in which the DAS executes the operation frequently. Here the function **F(Z)** to be minimized is proportional to the average number of time units in the operation. In accordance with the definitions **(1,4-7)** it can be written as $F(Z)=\sum_{j\in Z} a_j n_j$, where a_j are the constants which depend on the type and frequency of the operation. If all τ_j satisfy condition **(2)** then a_{j*} is equal to the probability that PE_{j*} participates in **W**. Let p_k, $k=\{1,\ldots,E\}$, be the probability that a group of **k** PEs participate in **W**. If the coefficients n_j are ordered as $n_1\geq n_2\geq\ldots\geq n_E$, then $a_j=\sum_{d=1,\ldots,E-j+1}\frac{p_d C^{d-1}_{E-j}}{C^d_E}$, $j\in Z$, where $\sum_{d=1,\ldots,E} p_d=1$, $C^d_E=\frac{d!(E-d)!}{E!}$.

Theorem 3. At least one relation of the system **(7)** is fulfilled for arbitrary values τ_j, $T_L\leq\tau_j\leq T_H$, $j=1,\ldots,E$, if

$$\sum_{j'\in Z'}\left(\frac{c_{j'}}{n_{j'}}\right)+\frac{T_H}{T_L n_{j*}}\sum_{j''\in Z''} c_{j''}\leq 1, \tag{11}$$

where $Z'\cup Z''=Z$, $Z'\cap Z''=\emptyset$; $j^*, j'\in Z'$; $j''\in Z''$; $n_{j*}T_L\geq n_{j''}T_H$.

Proof. Let us suppose the contrary. Then

$$n_j\tau_j<\sum_{j'\in Z'} c_{j'}\tau_{j'}+\sum_{j''\in Z''} c_{j''}\tau_{j''} \tag{12}$$

for all $j\in Z$. This implies

$$\frac{c_{j'}\tau_{j'}}{n_{j'}\tau_{j'}}>\frac{c_{j'}\tau_{j'}}{\sum_{j'\in Z'} c_{j'}\tau_{j'}+\sum_{j''\in Z''} c_{j''}\tau_{j*}\frac{T_H}{T_L}}, \tag{13}$$

for $j'\in Z'$ and

$$\frac{c_{j''}\tau_{j*}T_H}{n_{j*}\tau_{j*}T_L}>\frac{c_{j''}\tau_{j*}T_H}{T_L\left(\sum_{j'\in Z'} c_{j'}\tau_{j'}+\sum_{j''\in Z''} c_{j''}\tau_{j*}\frac{T_H}{T_L}\right)}, \tag{14}$$

for $j''\in Z''$

If we take the sum of (13) over $j'\in Z'$ and of (14) over $j''\in Z''$, and add up the corresponding left and right parts, we obtain the contradiction to the assumption **(7)**.

$$\sum_{j'\in Z'}\left(\frac{c_{j'}}{n_{j'}}\right) + \frac{T_H}{n_{j*}T_L}\sum_{j''\in Z''} c_{j''} > \frac{\sum_{j'\in Z'} c_{j'}\tau_{j'} + \sum_{j''\in Z''}\frac{c_{j''}\tau_{j*}T_H}{T_L}}{\sum_{j'\in Z'} c_{j'}\tau_{j'} + \sum_{j''\in Z''}\frac{c_{j''}\tau_{j*}T_H}{T_L}}.$$

Q.E.D.[4]

Solution of the optimization task of finding the minimal average value for the duration of the operation. To speed up the operation one has to minimize the function **F(Z)** under either condition **(8)** or **(9)** and to determine the values n_j, $j\in Z$. For that we will formulate the two theorems.

Theorem 4. If the condition $\sum_{j\in Z}\left(\frac{c_j}{n_j}\right)=1$ holds true then **Min** $F(Z)=(\sum_{j\in Z}\sqrt{c_j a_j})^2$.

Proof. We will use polar coordinates in E-dimensional space $(\sqrt{a_1 n_1},\dots,\sqrt{a_E n_E})$, where **E** is the number of PEs participating in the operation. Then using the system

$$a_1 n_1 = F(\varphi)\prod_{l=1,\dots,E-1}\sin^2\varphi_l,$$

$$a_d n_d = F(\varphi)\cos^2\varphi_{d-1}\prod_{l=d,\dots,E-1}\sin^2\varphi_l,\ \text{where}\ d=2,\dots,E-1, \quad (15)$$

$$a_E n_E = F(\varphi)\cos^2\varphi_{E-1}$$

one can represent the function to be minimized as

$$F(\varphi) = \frac{c_1 a_1}{\prod_{l=1,\dots,E-1}\sin^2\varphi_l} + \frac{\sum_{d=2,\dots,E-1} c_d a_d}{\cos^2\varphi_{d-1}\prod_{l=d,\dots,E-1}\sin^2\varphi_l} + \frac{c_E a_E}{\cos^2\varphi_{E-1}}.$$

According to the definition of the extremum we have

$$\frac{\partial F(\varphi)}{\partial\varphi_l} = 0,\ \text{where}\ l=1,2,\dots,E-1. \quad (16)$$

Let us consider **l=1**. From **(15)** and **(16)** we have

$$\frac{n_1 a_1}{n_2 a_2} = \frac{\sin^2\varphi_1}{\cos^2\varphi_1}$$

and **(17)**

---

4 From (11) we can derive very important fact that the end of the operation can be determined, in principle, by any group of PEs or even by only one PE. The only requirement is to select the coefficient n_{j*} so that the condition $n_{j*}T_L \geq n_{j''}T_H$ is fulfilled.

$$\frac{c_1a_1}{\sin^4\varphi_1}=\frac{c_2a_2}{\cos^4\varphi_1}.$$

Eliminating the parameter φ_1 we get $\frac{n_1}{n_2}=\frac{\sqrt{c_1a_2}}{\sqrt{c_2a_1}}$. In the same way we obtain $\frac{n_k}{n_j}=\frac{\sqrt{c_ka_j}}{\sqrt{c_ja_k}}$ for any **k** and **j**. Now using condition **(8)** for all $j\in Z$, we find $n_k=\frac{\sqrt{c_k}}{\sqrt{a_k}}\sum_{j\in Z}\sqrt{c_ja_j}$.

Taking the sum of a_kn_k over all $k\in Z$ yields $\text{Min } F(Z)=\left(\sum_{j\in Z}\sqrt{c_ja_j}\right)^2$.

Q.E.D.

Applying this result to the investigated problem we obtain its solution, i.e., the optimal individual waiting time for each PE in the DAS is

$$t(j^*,W)=\tau_{j^*}\frac{\sqrt{c_{j^*}}}{\sqrt{a_{j^*}}}\sum_{j\in Z}\sqrt{c_ja_j}\text{ for any } j^*\in Z.$$

Theorem 5. If the condition $\sum_{j'\in Z'}\left(\frac{c_{j'}}{n_{j'}}\right)+\frac{T_H}{n_{j^*}T_L}\frac{\sum_{j''\in Z''}c_{j''}}{n_{j''}}=1,\ j^*\in Z$, holds true then

$$\text{Min } F=\left(\sqrt{a_{j^*}\left(c_{j^*}+\frac{T_H}{T_L}\sum_{j''\in Z''}c_{j''}\right)}+\sum_{j\in Z'\backslash j^*}\sqrt{c_{j'}a_{j'}}\right)^2.$$

Consider the two following examples.

Example 1. Let three PEs participate in the operation with $T(1,Z)=16\tau_1+\tau_2+\tau_3$, and let τ_j satisfy relation **(2)**. Then according to [9] one solution is $n_1=n_2=n_3=18$, i.e., $F=54$. If we use Theorem 3 then $n_1=24$, $n_2=n_3=12$. Therefore, $F=36$ and the duration of the operation can be reduced by a factor of **1.5**.

Example 2. Let three PEs with the arbitration words **101, 010, 001** participate in the arbitration operation under condition **(2)**. Arbitration of the first two takes time $T(1,Z)=2\tau_1+\tau_2$, and of the last two requires time $T(2,Z)=\tau_2+\tau_3$. Thus, there are two different possibilities for the operation. From **(15)** we find

$$\frac{2}{\sin^2\varphi_1\sin^2\varphi_2}+\frac{1}{\cos^2\varphi_1\sin^2\varphi_2}=F(\varphi)$$

and

$$\frac{1}{\cos^2\varphi_1\sin^2\varphi_2}+\frac{1}{\cos^2\varphi_2}=F(\varphi).$$

This causes $\cos^2\varphi_2=\frac{2}{2+\sin^2\varphi_1}$, and from $\left(\partial F(\varphi_1,\varphi_2(\varphi_1)\right)=0$ we have $n_1=\frac{2}{\sqrt{3}}(1+\sqrt{3})$, $n_2=(1+\sqrt{3})$, $n_3=\frac{1}{\sqrt{3}}(1+\sqrt{3})$.

Example 3. Let the DAS consists of eight PEs. Let them use the **3**-bit binary words $\mathbf{A_0=000}$, $\mathbf{A_1=001,...,A_7=111}$ as their arbitration priorities. As every PE can participate in the arbitration operation, only nine different cases can occur for all possible values of τ_j, $0<\tau_j<\infty$. According to **(8)** they correspond to one of the conditions

$$0<x_j\le1,\ x_6+x_5\le1,\ x_6+x_3\le1,\ x_6+x_1\le1,\ 2x_5+x_3\le1,$$
$$2x_5+x_2\le1,\ x_4+x_3\le1,\ x_4+x_2\le1,\ x_4+x_1\le1,\ x_2+x_1\le1, \qquad (18)$$

where the parameters $\lambda_j=\frac{1}{n_j}$ $(j=0,...,7)$ are used. To speed up the arbitration operation according to the discussions above it is necessary to minimize the function $F(\lambda_0,...,\lambda_7)=\sum_{j=0,...,7}\frac{1}{\lambda_j}$ under conditions **(18)**. Below we prove that to reach the minimum it is sufficient to satisfy only five linear independent equations, i.e., the <u>essential conditions</u>, from the set

$$\lambda_6+\lambda_5=1 \quad \lambda_6+\lambda_3=1, \quad \lambda_6+\lambda_1=1, \quad 2\lambda_5+\lambda_3=1, \quad 2\lambda_5+\lambda_2=1,$$
$$\lambda_4+\lambda_3=1, \quad \lambda_4+\lambda_2=1, \quad \lambda_4+\lambda_1=1, \quad \lambda_2+\lambda_1=1.$$

That is the values of four other expressions must be less than one.

Let us, first, suppose that the number of essential conditions is greater than five. Then at least two conditions will be linearly dependent. Secondly, if the number of essential conditions is less than five then at least one variable will approach infinity. This contradicts the system **(18)**.

Obviously, the following five conditions are linearly independent

$$\lambda_6+\lambda_1=1, \quad 2\lambda_5+\lambda_3=1, \quad 2\lambda_5+\lambda_2=1, \quad \lambda_4+\lambda_3=1, \quad \lambda_4+\lambda_1=1.$$

Then the optimal solution of the formulated optimization task is $F(\lambda_0,...,\lambda_7)=14$, and the optimal values of parameters λ_j are

$$\lambda_7=\lambda_0=1,\ \lambda_6=0.536,\ \lambda_5=0.268,\ \lambda_4=0.536,\ \lambda_3=0.464,\ \lambda_2=0.464,\ \lambda_1=0.464.$$

Now using the relation $\lambda_j=\frac{1}{n_j}$ it is possible to find the unknown coefficients n_j. If we compare the value of the optimized function to the result of [9] we conclude that according to the suggested method the duration of the arbitration operation can be reduced by a factor of **1.5** without any changes of the hardware.

Summary. The suggested method of identification of the duration of distributed operations can be used to enhance the performance of a DAS. It allows

1) to compute the optimal individual waiting time of PEs where each of them has no information about speed characteristics of others,

2) to speed up the operation without any changes to the PEs and to the algorithm of the operation,

3) to determine the duration of the operation dynamically,

4) to find subprocesses which contribute most to the global waiting time, and

5) to minimize the duration of the operation by selection an appropriate hardware configuration of a DAS.

References

1. Borrill P.L.: Futurebus+: A Tutorial. VMEbus applications 4 (1990) 25-34.
2. Taub D.M.: Clockless Synchronization of Distributed Concurrent Processes, IEE Proceedings-E, Vol. 139, No. 1 (1992) 88-92.
3. Taub M.D. Corrected Settling Time of the Distributed Parallel Arbiter, IEE Proc. E, Vol. 139, No. 4, 1992, 348-354.
4. Makhaniok M. Cherniavsky V., Männer R., K.-H. Noffz, Extremal Codes for Speed-Up of Distributed Parallel Arbitration, Parallel Algorithms and Applications, Vol., 6, pp. 1-16, 1995.
5. Makhaniok M., Cherniavsky V., Männer R., Stucky O., Massively Parallel Realization of Logical Operations in Distributed Parallel Systems. in: H. Burkhart (Ed.): Proc. CONPAR ´90 - VAPP IV, Lect. Notes in Comp. Sci. 457, Springer, Heidelberg (1990), 796-805.
6. Voronoi G.F.: Sur Quelques Proprietes des Formes Guadratiques Positives Parfaites, J. Reine und Angew. Math., Vol. 133 (1908) 79-178.
7. Taub D.M.: Arbitration and Control Acquisition in the Proposed IEEE 896 Futurebus, IEEE Micro, Vol. 4 (1984) 28-41.
8. Taub D.M.: Improved Control Acquisition Time for the IEEE 896 Futurebus. IEEE Micro, Vol. 7, No. 3 (1987) 52-62.