Effects of the Dynamic Processor Scheduling in a Function Distributed System

SATORU IKEHARA*

Due to advances in hardware technology, the cost of a small scale processor is rapidly decreasing; and it will be possible to build a computer complex, comparable to a large scale system in performance, with many small scale computers.

The computer complex can be classified into a multiprocessor (MP) and a distributed function system (DP). It is said that the response time of DP is longer than that of MP, because this system uses more queues than the multiprocessor and the waiting time in the queues increases.

In other words, every processor is fixed to some functions; then the processors of DP have more idle time than that of MP.

If we let some processor of DP flexibly be assigned to any function dynamically, performance improvement will be obtained.

This paper evaluates the performance improvement and the cost performance improvement which are able to be obtained by the dynamic processor scheduling in the DP system. The conditions under which the dynamic processor scheduling act effective are derived. In this evaluation, it is assumed that the distributed processor consists of general purpose processors, which are dynamically scheduled to busy stages, and specialized processors, which are bound to functions.

1. Introduction

Since great advances in hardware technology have brought cost decreases (especially in small scale computers), it seems possible to build a computer complex system, using these small scale computers, more economically than an ordinary large scale computer can be built [1], [2].

The computer complex can be classified into Multiprocessor (MP) and a Function Distributed Processing (DP) System. In MP, every processor plays the same role and processes the same process. However, in DP, the functions are divided among the processors of which DP is constructed. Thus, it seems possible to obtain better cost performance by DP, because every processor, of which DP consists, is able to be designed as a function oriented processor. Therefore, these processors are expected to become faster and cheaper than the general purpose computers of which MP is constructed.

On the contrary, compared with the MP (in which any job can be assigned to any processor which is not busy), the job in the DP is processed walking around the processors; so the job will wait as long as the required processors are busy. Thus, it seems that the DP response time is longer than that of MP.

The author has shown [3] that if the performance of individual processors of DP is twice as good as that of MP, the response time of DP becomes shorter than MP.

*Electrical Communication Laboratory of Nippon Telegraph and Telephone Public Corporation.

This paper studies the effect of dynamic processor scheduling methods. When the load on some processors is heavy, while other processors are empty, a good efficiency effect will be achieved if portion of the load, which is now being processed by the busiest processors, is assigned dynamically to processors which are empty. To evaluate these effects, DP* (modified DP) is constructured by both the function oriented processors and general purpose processors. The general purpose processors are allocated to any busy stages dynamically. Then, improvements in throughput and cost performance for the system are evaluated.

2. Evaluation Models

Let's divide the software, which operates the system, into m function groups and set up special purpose processors for each stage. These groups are called stages. The job enters the systems at any stage, walks through the stages being processed by each function and exits at a certain stage. Job transitions among stages are controlled by queues which are placed in front of stages.

Thus, these systems are represented by the queuing network shown by Fig. 1 where:

① Job transition sequences are chosen according to the transition probability matrix.

② Service rate of each stage distributes exponentially.

③ Exactly when a job exits the system, a new job is scheduled to keep the same number of jobs in the system.

④ Every processor is arranged to balance the load among various stages.

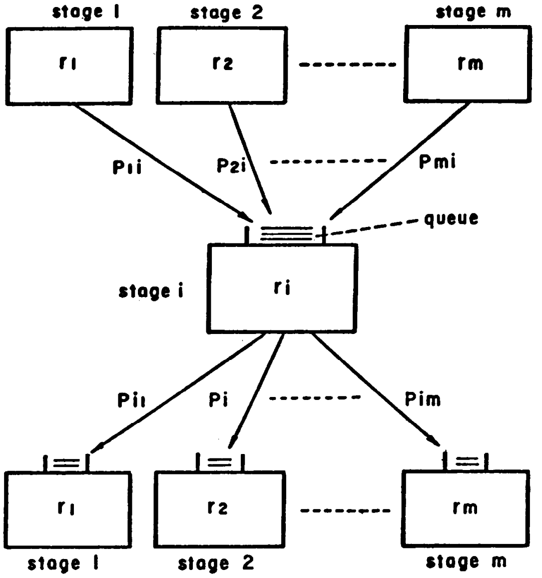

Fig. 1 *i*-th stage Distributed Processing System (DP) model.

⑤ All of the jobs are the same type (Single job type).

Let r_i be the number of processors of which i-th stage consist and assume that both the multiprocessor and the distributed processor systems are constructed of the same number of processors M.

Then $m=1$, $r_1=M$ represents the multiprocessor system and $m \neq 1$, $\sum_{i=1}^{m} r_i=M$ represents the distributed processor systems.

Let's denote the type of the system by "$(m \times r)$", assuming the number of processors of every stage is the same. The relation $m \times r=M$ holds.

Now, let's consider three types of the models.

(1) DP (Distributed Processor) Model

Consider "$(M \times 1)$ type" Distributed Processing System, each processor of which is specialized to its own stage so that it has better cost performance than a general purpose processor.

(2) DP* (Modified Distributed Processor) Model

Let's consider another $(M \times 1)$ type of Distributed Processing System, which is constructed with specialized processors and general purpose processors. In this model, assume that general purpose processors are not fixed to stages but are dynamically assigned to any stages which are busy. Let's denote these systems simply by "DP*".

(3) MP (Multiprocessor) Model

It is assumed that $m=1$ and $r=M$. Every processor has the same function and plays the same role.

Assume that the population of jobs in these systems are the same as N and not smaller than the number of processors M. And, the average workloads of the jobs which are processed by these three models are the same, i.e., the average of the sum of the dynamic steps of jobs is the same in these models.

3. Performance Evaluation

3.1 Analysis of Each Model

(1) DP model analysis

It has been shown that in this model the processor utilization ρ_{DP} and the response time T_{DP} are,[3]

$$\rho_{DP}=\frac{N}{M+N-1} \tag{1}$$

$$T_{DP}=(M-1)\frac{1}{\lambda}\frac{\rho_{DP}}{1-\rho_{DP}}. \tag{2}$$

Where λ denotes the throughput of this model.

(2) DP* model analysis

Let x be the number of general purpose processors in DP*. The population of the jobs in the system N is assumed not to be smaller than the number of processors M; so it is guaranteed for all of the general purpose processors to have jobs to be processed at any time.

(i) Processor Utilization

In the DP* system, the total number of processors is M; then the number of stages, m, is equal to $M-x$.

The number of distinguishable states, $(n_1, n_2, \cdots, n_{M-x})$, where every stage has $n_1, n_2, \cdots, n_{M-x}$ jobs, respectively, is represented by

$$w(n_1, n_2, \cdots, n_{M-x})=\frac{N!}{n_1!n_2!\cdots n_{M-x}!}, \tag{3}$$

where $\sum_{i=1}^{M-x} n_i=N$.

Let $W_T(N)$ denote the sum of w for all the arrangements of $\tilde{n}$; $W_T(N)$ is equal to the number of partitions of N jobs among $M-x$ stages:

$$W_T(N)=\sum_{\tilde{n}\in(\Sigma n_i=N)} w(n_1, n_2, \cdots, n_{M-x}) = \binom{M+N-x-1}{N}. \tag{4}$$

Now, let us define the total processor utilization, ρ_{DP^*}, as follows:

$$\rho_{DP^*}=\left\{\sum_{i=1}^{M-x} P(n_i \geq 1)+x\right\}/M \tag{5}$$

where, $P(n_i \geq 1)$ is the probability that the i-th stage has jobs to be processed; i.e., the utilization of the processor at i-th stage. Thus, $P(n_i \geq 1)$ is represented as

$$\left.\begin{aligned} P(n_i \geq 1)&=\sum_{j=1}^{N} P(n_i=j) \\ P(n_i=j)&=\frac{1}{W_T(N)}\sum_{\substack{\tilde{n}\in(\Sigma n_i=N)\\ \& n_i=j}} w(\tilde{n}) \equiv \frac{w(n_i)}{W_T(N)} \end{aligned}\right\} \tag{6}$$

Where, $w(n_i)$ represents the number of states where the population of jobs at i-th stage is n_i. The second term in Eq. (5) implies that the utilization of all the processors, which are able to be assigned to any stage, will always be 1.0, because of the condition wherein $N \geq M$.

If we consider the meaning of Eq. (6), it is obvious that $P(n_i \geq 1)$ is given by

$$P(n_i \geq 1) = \frac{W_T(N-1)}{W_T(N)}. \qquad (7)$$

Then, in Eq. (5), the factor inside the $\sum$ term has become independent of i, so it can be simplified by making use of Eq. (4):

$$\rho_{DP*} = \frac{M(N+x) - x(x+1)}{M(N+M-x-1)}. \qquad (8)$$

Thus, we are led to calculated ρ_{DP*} by giving some values to M, N and x.

(ii) Response Time

It is assumed that both the number of jobs in DP and DP* are equally N, and that the loads of jobs are the same in these two systems. Let λ_{DP} and λ_{DP*} denote the throughputs of DP and DP*, respectively. Then, the throughputs which are proportional to the processor utilization are written as

$$\lambda_{DP} = k\rho_{DP}, \quad \lambda_{DP*} = k\rho_{DP*} \qquad (9)$$

where, k is a constant and ρ_{DP}, ρ_{DP*} are processor utilizations of DP and DP* given Eqs. (1) and (8), respectively.

Now, let's define the difference ratio of the response time, $\Delta T/T$, as

$$\frac{\Delta T}{T} = \frac{T_{DP} - T_{DP*}}{T_{DP}}, \qquad (10)$$

where, T and T^* are the response times of DP and DP*. Here, T_{DP} and T_{DP*} are represented by N/λ_{DP} and N/λ_{DP*} respectively. Then, Eq. (10) becomes

$$\frac{\Delta T}{T} = 1 - \lambda_{DP}/\lambda_{DP*}. \qquad (11)$$

Substituting Eq. (9) into Eq. (11) and making use of Eqs. (1) and (8), results in

$$\frac{\Delta T}{T} = \frac{x\{(M+N-1)(M-x-1) + MN\}}{\{M(N+x) - x(x+1)\}(M+N-1)}. \qquad (12)$$

Thus, it has become known that the DP* response time is shorter than that of DP according to Eq. (12)

(iii) Maximum Throughput Value

It is evident that the total utilization of DP and DP* become 1 if the number of jobs approaches infinity. Therefore, the maximum throughput values for these two systems agree.

(3) MP model analysis

Apparently the processor utilization $\rho_{DP} = 1$ and the response time $T_{MP} = (M/N)(1/\mu_{MP})$, because $N \geq M$. Where, μ_{MP} is the service rate of the individual processors.

3.2 Evaluation

(1) Value of x Which Makes ρ_{DP*} Maximum

Let's derive the number of general purpose processor, x, that makes the processor utilization, ρ_{DP*}, maximum, assuming that the number of jobs in the system is fixed.

Solving the equation that is obtained by differentiating Eq. (8) with respect to x and equating it to zero, we obtain

$$x_0 = (M+N-1) - \sqrt{N(N-1)} \qquad (13)$$

as x which makes ρ_{DP*} maximum. Here, the maximum value of ρ_{DP*}, ρ_{DP*max}, is given by

$$\rho_{DP*max} = \frac{1}{M}\{M + 2N - 1 - 2\sqrt{N(N-1)}\}. \qquad (14)$$

From Eq. (13), for N, M, the relation,

$$M - 1 \leq x_0 \leq M,$$

can be easily proved. Since DP* system, where x is equal to $M-1$, is just the same as MP system, it is possible to conclude that the DP* processor utilization is always smaller than that for MP.

(2) Special Cases

We shall evaluate ρ_{DP*}, ρ_{DP} and $\Delta T/T$ in special cases, where $N=M$ and $N=2M$. In these cases, ρ_{DP}, ρ_{DP*} are plotted in Fig. 2 from Eqs. (1) and (8); and $\Delta T/T$ are plotted in Fig. 3 from Eq. (12).

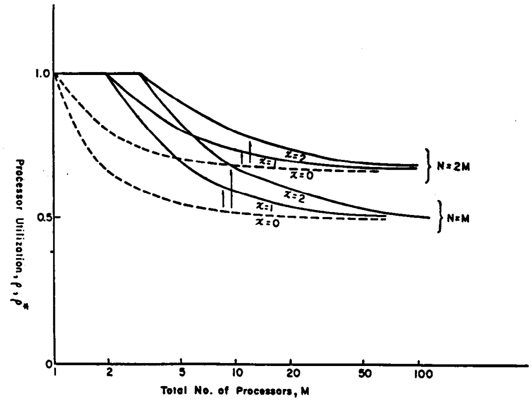

Fig. 2 Processor utilization improvement.

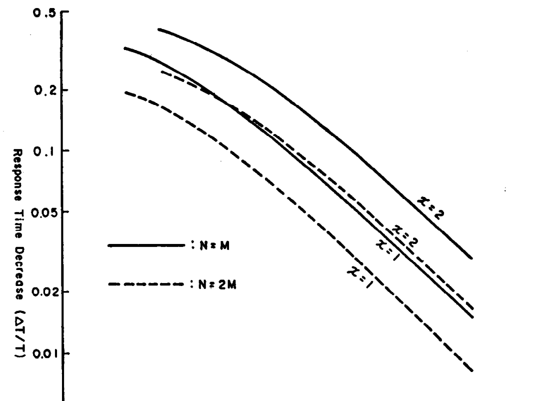

Fig. 3 Response time decrease.

It is obvious that processor utilization fairly increases by letting some processor free to be assigned to any stage. For example, it is seen that it yields better utilization to let a single processor free if $N \le 5$ or to let two processors free if $N \le 9$, rather than letting the number of jobs in the system increase twice.

From Fig. 3, it is also seen that the response time reduces up to 30∼40 percent for a relatively small number of M by the dynamic processor scheduling.

4. Cost Performance Evaluation

According to the results derived above, it is possible to improve processor utilization, i.e. throughput, by dynamic processor scheduling.

Since those processors that are dynamically scheduled must fit any of the stages, it is necessary to use general purpose processors which are more expensive than specialized processors of which DP consists. Therefore, the DP* system cost is more expensive than that of DP. In this section, the DP cost performance is evaluated, compared with that of DP and MP.

Now, assume that the cost of two kinds of processors are:

specialized processor . . . 1
general purpose processor . . . $\xi(\ge 1)$.

and that these processors have the same process speed (*e.g.* average execution time/instruction). Let cost performance η be the ratio of processor utilization (i.e. throughput) which is to be obtained when there are N jobs in the system to the system cost C;

$$\eta = \rho / C \tag{15}$$

Let us evaluate η for DP*, DP and MP.

4.1 Evaluation of η

The cost MP system, C_{MP}, is written as

$$C_{MP} = M\xi, \tag{16}$$

where M is the number of general purpose processors.

Since the DP* system consists of x general purpose processors and $M-x$ specialized processors, the cost of DP*, C^*_{DP}, is written as

$$C^*_{DP} = \xi x + (M-x). \tag{17}$$

The processor utilization of MP, ρ_{MP}, is exactly equal to 1 if $N \ge M$. Therefore, cost performance η_{MP} can be written:

$$\eta_{MP} = \frac{1}{\xi M}. \tag{18}$$

Since the processor utilization of DP*, ρ^*_{DP}, has been given by Eq. (8), cost performance, η^*_{DP}, is as follows:

$$\eta^*_{DP} = \frac{M(N+x) - x(x+1)}{M\{(\xi-1)x + M\}(M+N-x-1)}. \tag{19}$$

In the special case of this system, where $x=0$, we obtain η_{DP} for DP such as,

$$\eta_{DP} = \frac{N}{M+N-1} \cdot \frac{1}{M}. \tag{20}$$

Let us define the cost performance ratio of η^*_{DP} to η_{MP} and to η_{DP} as follows:

$$\psi = \eta^*_{DP} / \eta_{MP} \tag{21a}$$

$$\varphi = \eta^*_{DP} / \eta_{DP}. \tag{21b}$$

By making use of Eqs. (18)∼(20),

$$\psi = \frac{\xi\{M(N+x) - x(x+1)\}}{\{(\xi-1)x + M\}(M+N-x-1)} \tag{22a}$$

$$\varphi = \frac{M+N-1}{N} \cdot \frac{M(N+x) - x(x+1)}{\{(\xi-1)x + M\}(M+N-x-1)}. \tag{22b}$$

Figs. 4 and 5 represent the relations of ψ and φ versus x/M.

4.2 Condition Where Maximum Cost Performance Value Exists

The value of x which makes ψ or φ maximum is able to be obtained by solving:

$$\frac{d\psi}{dx} = 0 \quad \text{or} \quad \frac{d\psi}{dx} = 0 \tag{23}$$

These equations lead to

$$ax^2 + bx + c = 0, \tag{24}$$

where, a, b, c are

$$a = M + N - \xi N, \tag{25a}$$

$$b = 2M + 2\xi MN - 2M^2 - 4MN, \tag{25b}$$

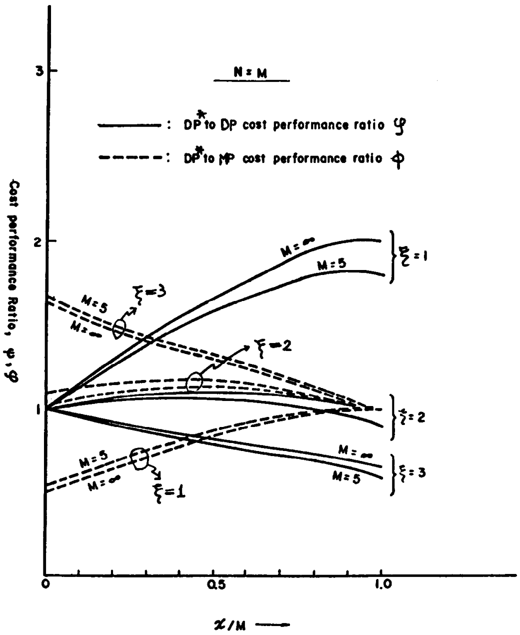

Fig. 4 System cost performance characteristics.

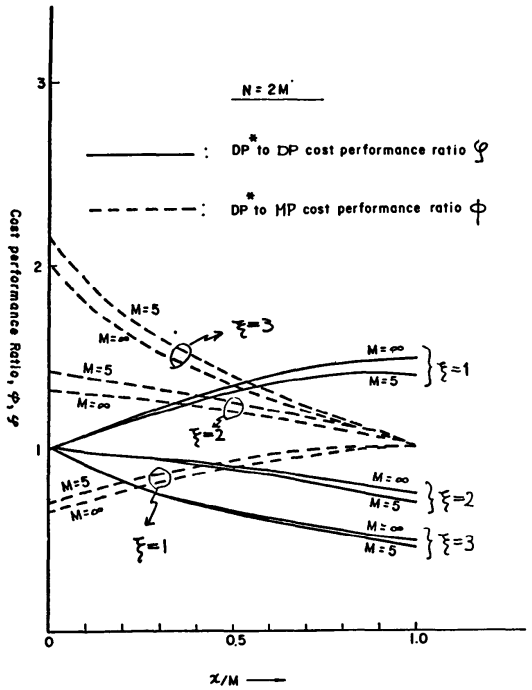

Fig. 5 System cost performance characteristics.

$$c=M-2M^2+MN^2-2MN+3NM^2 +\xi MN-\xi MN^2-\xi NM^2+M^3 \quad (26c)$$

respectively. Then, the maximum value of ψ and φ exists when the relation $b^2-4ac\geq 0$ holds. This relation can be reduced as

$$-N\xi^2+(M+2N-1)\xi-(N-1)\geq 0. \quad (27)$$

Solving this equation for ξ:

$$\alpha\leq\xi\leq\beta \quad (28a)$$

is obtained. Where, α, β are

$$\alpha=\frac{1}{2N}\{M+2N-1-\sqrt{(M-1)^2+4MN}\}, \quad (28b)$$

$$\beta=\frac{1}{2N}\{M+2N-1+\sqrt{(M-1)^2+4MN}\}, \quad (28c)$$

When $N=M$ and M approaches infinity:

$$\lim_{M\to\infty}\alpha=\frac{3-\sqrt{5}}{2}\simeq 0.382, \quad (29a)$$

$$\lim_{M\to\infty}\beta=\frac{3+\sqrt{5}}{2}\simeq 2.618. \quad (29b)$$

And when $N=2M$ and M approaches infinity:

$$\lim_{M\to\infty}\alpha=\frac{5-\sqrt{12}}{4}\simeq 0.384, \quad (30a)$$

$$\lim_{M\to\infty}\beta=\frac{5+\sqrt{12}}{2}\simeq 2.116. \quad (30b)$$

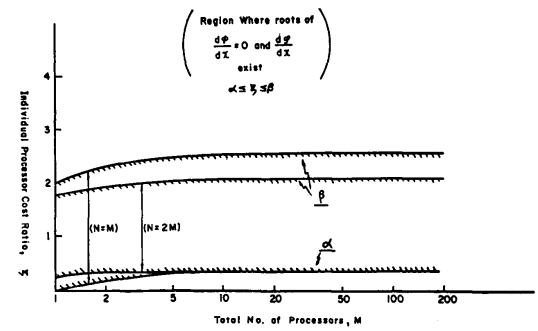

Fig. 6 Region where dynamic processor scheduling acts are effective (1).

The range of ξ, which satisfies Eq. (28a), is represented in Fig. 6.

From the above discussion, it has been found that:

① When $N=M$, (i) if the value about zero to 2.6 is taken as that of ξ, there exists a value of x for $0<x<M$, which makes ψ and φ maximum. However, (ii) if ξ takes a greater value, the maximum values of ψ or φ are obtained when $x=0$, i.e. in DP. (iii) on the contrary, if ξ takes a smaller value, ψ and φ become maximum in $x=M$, i.e. in MP.

② When $N=2M$, the range of ξ, where the maximum value of ψ and φ exists, becomes smaller than ①.

4.3 Condition Where $\psi\geq 1$ or $\varphi\geq 1$

The condition where the dynamic processor scheduling affords better cost performance than MP is obtained by letting $\psi\geq 1$ for ψ of Eq. (22a):

$$\xi\geq\frac{M+N-x-1}{N}. \quad (31)$$

The same relation is derived for DP by letting $\varphi\geq 1$ for φ of Eq. (22b):

$$\xi\leq\frac{M+N-1}{N}+\frac{M-x}{M+N-x-1}. \quad (32)$$

These values of ξ, in the two cases where $N=M$ and $N=2M$, are plotted for M in Fig. 7 and Fig. 8. If ξ takes the value indicated by shading in these figures, the DP* cost performance is better than that of MP and DP; so dynamic processor scheduling acts effectively.

4.4 Special Cases

(1) When $x=0$ or M

If $x=0$, ψ and φ reduces to

$$\psi=\frac{\xi N}{M+N-1}, \quad (33a)$$

$$\varphi=1. \quad (33b)$$

Here, letting $N=kM$, solve $\psi\geq 1$, then

$$\xi\geq\frac{(k+1)M-1}{kM}. \quad (34)$$

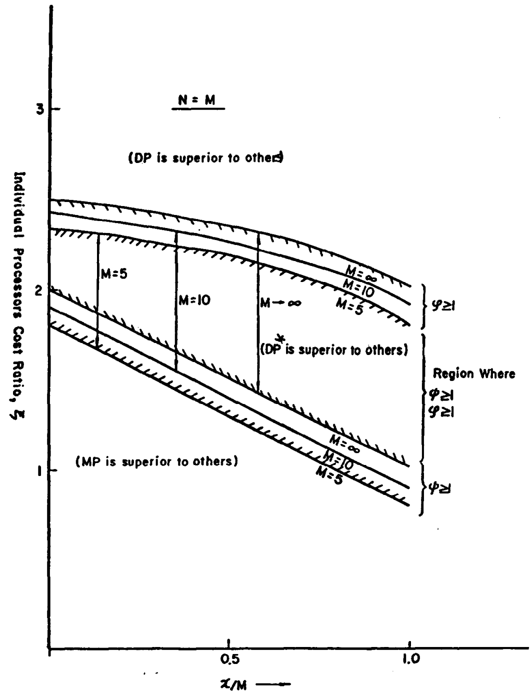

Fig. 7 Region where dynamic processor scheduling acts are effective (2).

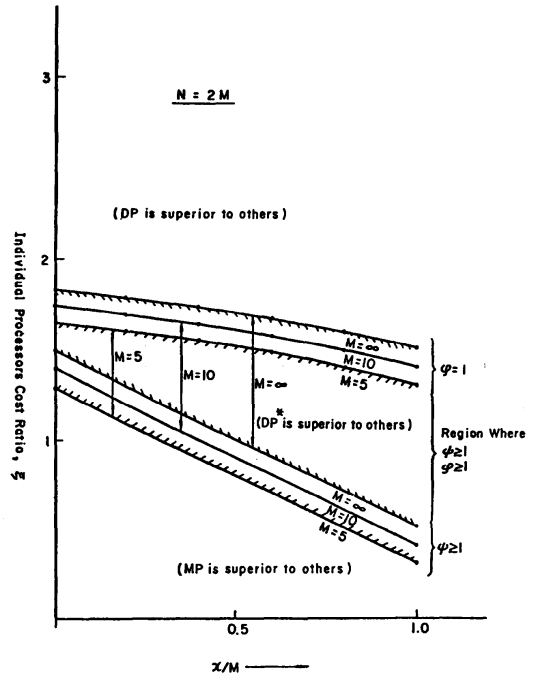

Fig. 8 Region where dynamic processor scheduling acts are effective (3).

If M approaches infinity, Eq. (34) becomes $\xi \geq (k+1)/k$. It is clear that, if $k=1$, $\rho=1$ in MP. In this case, the above condition results $\xi \geq 2$. This condition coincides with the performance condition derived in [3].

Next, when $x=M$, ψ and φ reduces to

$$\psi = 1, \tag{35a}$$

$$\varphi = \frac{M+N-1}{N} \cdot \frac{1}{\xi}. \tag{35b}$$

This accounts for MP. Similar to Eq. (34),

$$\xi \geq \frac{M+N-1}{N} \tag{36}$$

can be derived by $\varphi \geq 1$. Letting $N=M$, the same condition as the cost performance condition shown in [3] is derived,

$$\xi > 2 - 1/M. \tag{37}$$

(2) When $N=M$, $\xi=2$

In this case, (25) can be solved directly to yield that when

$$x = (M-1)/2, \tag{38}$$

we get maximum value of ψ and φ such as

$$\psi(\max) = \frac{10M^2 - 4M + 2}{(3M-1)^2}, \tag{39a}$$

$$\varphi(\max) = \frac{(2M-1)(5M^2-2M+1)}{M(3M-1)^2} \tag{39b}$$

Here, in the limit of $M \to \infty$, both ψ and φ approach 10/9. When about half of the processors are dynamically scheduled, the cost performance becomes the greatest and is better than MP and DP by about 10%.

5. Summary

The effects of the dynamic processor scheduling in the function distributed system are evaluated. It is found that letting the number of jobs in the system be fixed, the total processor utilization (i.e. throughput) can be increased by dynamic processor scheduling, where some processors are made available to any of the stages; and if some stages are busy, these processors help dynamically. These effects are calculated as 0~20% or 0~40% for the number of processors dynamically scheduled as one or two, respectively.

Since the processors in a stage can be specialized, it is expected that these processors can be obtained cheaper. However, processors which are scheduled dynamically seem expensive because these processors must be made for general purposes, just the same as in MP. Then, letting the cost ratio of these two kinds of processors be ξ, the condition that made the dynamic processor scheduling effective is derived as

$$0.4 \sim 1.8 \leq \xi \leq 1.8 \sim 2.6.$$

If ξ is greater than these values, DP is advantageous, because every processor is very cheap. On the contrary, if ξ is smaller than these values, MP is advantageous.

Acknowledgement

The author wishes to express his thanks to Dr. Y. Muraoka for the usual discussions.

References

1. Joseph, E. C. Future Computer Architecture-POLYSYSTEMS, *COMPCON* '72, 149–153.
2. Turn, R. Computers In The 1980s, Columbia U.P., 1974.
3. Ikehara, S. The Performance Evaluations of Function Distributed System by the Queuing Network Models, *Information Processing*, **18**, 11, 1102–1109.