

Constraints in the Construction of Computers with ever Larger Numbers of Processors

Joseph W. Goodman
Department of Electrical Engineering
Stanford University
Stanford, CA 94305 USA

and

Haldun M. Ozaktas
Department of Electrical Engineering
Bilkent University
06533 Bilkent, Ankara, Turkey

1. INTRODUCTION

As the parallelism of digital computers increases, the limitations associated with interconnecting a large number of processors becomes a greater and greater constraint on system performance. For example, as we pack more and more processors of a given size together in a single machine, naturally the physical dimensions of the machine must grow, and with that growth comes an increase in the maximum time delay experienced in communicating between the most distant processors in the array. To a degree that depends on the algorithm being executed and its communication requirements, that communication latency between different parts of the machine ultimately poses a limit to the speed with which the machine can solve problems. The growth of size of the machine and the associated latency limitations in fact take place even in the idealized case of processors that occupy no volume themselves. We rapidly find that these infinitesimal processors must be separated by finite distances for both of two reasons: 1) a volume of space is required to allow the many interconnections between such processors to be realized, and 2) a minimum separation between processors is required to allow heat generated by both the processors and the interconnections to be removed from the system.

Our purpose in this paper is to discuss some of the fundamental aspects of the problem described above. We consider in what follows two different cases: 1) all interconnections are optical, and 2) all interconnections are electrical. Of course with a hybrid set of interconnects (i.e. part optical and part electrical), better performance can be achieved. However, we do not consider hybrid strategies in this paper. This work is the result of a Ph.D. thesis at Stanford University¹ from which a number of publications resulted²⁻⁶.

2. MODELS AND ASSUMPTIONS

We assume an ideal model of a parallel computer consisting of N processors that can communicate with one another. The computer can be modeled as a graph, where the nodes of the graph represent computing elements, while the edges of the graph represent interconnections between computing elements (see Fig. 1). Naturally the graph structure contains all necessary information about the interconnectivity of the elements. The bisection H of a system is defined as the number of graph edges crossing an imaginary boundary dividing the system in two. A graph edge may be implemented with $\chi \geq 1$ parallel interconnections (e.g., physical wires). Thus χH physical interconnections cross the imaginary boundary.

For the purposes of this discussion, we assume that the processing elements in the array are arbitrarily fast, and therefore that all time constraints are imposed by the interconnections. Of course, finite processor speed can also be included in the theory, but we will not do so here.

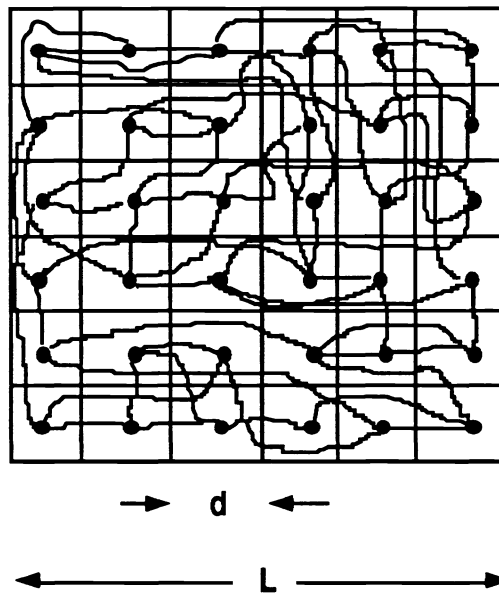


Figure 1. Connected grid model

Rent's rule may be used as a model for the "reach" of the interconnections from one processor across the array. If there are k interconnections leaving each processor, then Rent's rule states that the number of interconnections emanating from an array of N' nodes in an array of N total nodes is given by

$$P(N) = kN'^p, \quad (1)$$

where p is the Rent exponent ($0 \leq p \leq 1$) and we have assumed that $N' \leq N/4$.

For a system satisfying Rent's rule, and for $p > (e-1)/e$, where e is the dimension of the layout ($e = 2$ or 3), the bisection is given by

$$H = \kappa N^p, \quad (2)$$

where κ is a coefficient of the order of unity (c.f. Eq. (14)). Note that this relation implies that H grows faster than $N^{1/2}$ in two dimensions and faster than $N^{2/3}$ in three dimensions.

3. WIREABILITY LIMITATIONS

Optical Interconnections

In two dimensions we assume that there is a minimum width λ associated with a single interconnection, where λ represents the optical wavelength. Practical constraints usually

prevent one from achieving the ultimate limit, so that the actual width is typically some (often large) multiple $f\lambda$, where $f \geq 1$. If the bisection of graph is H , then clearly the minimum linear dimension L of a planar processor array is $L = f\lambda H$, assuming only one layer of interconnections. Hence we see that we must be dealing with a planar array that is at least $L \times L$ in size, and the propagation delay between the most widely separated elements is of the order of L/v where v is the velocity of propagation on the interconnections. Equivalently the propagation delay τ is of the order of

$$\tau = \frac{f\lambda H}{v} = \frac{f\lambda k\kappa N^p}{v}. \quad (3)$$

In three dimensions, the cross-sectional area of a single interconnection is $(\lambda f)^2$, and therefore the cross-sectional area of the system must be at least $(f\lambda)^2 H$ for H channels to pass. The square root of this cross-section gives the minimum linear dimension of the array, $L = f\lambda H^{\frac{1}{2}}$. The propagation-limited delay becomes

$$\tau = \frac{f\lambda H^{\frac{1}{2}}}{v} = \frac{f\lambda (k\kappa N^p)^{\frac{1}{2}}}{v}. \quad (4)$$

We note a slower rate of growth of τ with N in three dimensions than in two dimensions, as might have been expected.

Let B represent the maximum bit repetition rate. The maximum value of this repetition rate is limited by the transducers, since both attenuation and dispersion are not significant over the short distances of interest here. If we wish to increase B beyond the limit B_{max} allowed by the transducers, we can do so only by adding parallel connections and transducers. As we do so, system size and delay will increase. Thus there is a relationship between B and propagation delay τ . For bit repetition rates above the transducer limit, as we increase B , we also increase τ .

The bisection-bandwidth product, HB can be arbitrarily increased at the cost of increasing τ . The bisection-inverse-delay product H/τ is found to rise as $N^{p(\epsilon-2)/(\epsilon-1)}$ with N , where ϵ again represents the dimension of the layout. Thus the bisection-inverse-delay product can increase with N if the dimensionality of the layout is greater than 2.

Electrical Interconnections

The signal delay of an unterminated electrical line of the RC type is given by

$$\tau = \frac{\alpha \ell^2}{W^2}, \quad (5)$$

where W is the line-to-line spacing (some multiple of the width of a line), ℓ is the length of the line, and the constant α is proportional to $\rho\epsilon$, where ρ is the resistivity of the

conducting material and ϵ is the dielectric constant of the insulator. In two dimensions the linear extent of the system must be at least HW . Thus the length of the longest line is approximately HW , and the delay becomes

$$\tau = \frac{\alpha(HW)^2}{W^2} = \alpha H^2 = \alpha(k\kappa N^p)^2. \quad (6)$$

For an unterminated line, B_{max} is equal to $1/\tau$. Thus B_{max} falls as N^{-2p} with increasing N , while τ rises in proportion to N^{2p} . Any attempt to increase B_{max} by increasing W (i.e. using wider wires with less resistance) is thwarted by the resulting increase in line lengths. On the other hand, any attempt to increase B_{max} by reducing W (so as to get shorter lines) is thwarted by the inverse dependence of τ on W .

The bisection-bandwidth product and the bisection-inverse-delay product are equal in this case, and have the dependence

$$HB = \frac{H}{\tau} \propto N^{p(\epsilon-3)/(\epsilon-1)} \quad (7)$$

With both two- and three-dimensional layouts, these quantities can not be increased by increasing N .

If we allow the electrical interconnections to be superconducting, then it can be shown that the square law dependence on H is removed, and electrical interconnections behave very similarly to optical interconnections.

4. HEAT REMOVAL LIMITATIONS

Optical Interconnections

Our ability to remove heat from a two- or three-dimensional system is characterized by a quantity Q , representing power removable per unit of cross sectional area. We assume that there is an upper limit to the value of Q that can be achieved. If P represents the power dissipated in the system, then we must satisfy the constraint

$$QL^2 \geq P, \quad (8)$$

where L is again the linear extent of the system.

Our system has N elements with k connections each operated at a bit rate of B bits/sec. The energy per transmitted bit required for optical interconnections is represented by E_o joules. Thus the total power dissipation associated with the interconnections is

$$P = kNE_oB. \quad (9)$$

This relation, combined with Eq. (9), leads us to the conclusion that the linear dimension of such a system must satisfy

$$L \geq (kNE_oB / Q)^{1/2}. \quad (10)$$

Electrical Interconnections

For simplicity, we consider here only unterminated lines. The terminated case is more complex but has been analyzed (see comments in the conclusions section). The energy required per transmitted bit is written

$$E_n = \gamma \ell, \quad (11)$$

where the constant γ is proportional to ϵV^2 , ϵ being the dielectric constant and V the voltage to be delivered to the end of the interconnection. Note in particular the linear dependence of energy on length, in contrast with the optical case, in which the required energy is, to a good approximation, independent of length.

Let $\bar{\ell}$ denote the average interconnection length. If the elements in the array are spaced by distance d , then $\bar{\ell}$ can also be represented as $\bar{r}d$, where \bar{r} is a dimensionless quantity representing the average length of an interconnection measured in grid spacings.

The heat removal condition of Eq. (8) can now be stated

$$QL^2 \geq kN\gamma\bar{\ell}B. \quad (12)$$

Using $L = N^{\frac{1}{e}}d$, ($e=2$ or 3), we obtain

$$L \geq (k\gamma B/Q)\bar{r}N^{\frac{e-1}{e}}. \quad (13)$$

Using an argument that follows from Rent's rule, we may replace \bar{r} , through

$$\bar{r} = \kappa N^{p - \frac{e-1}{e}}, \quad (14)$$

(valid for $p > (e-1)/e$) so that

$$L \geq (k\kappa\gamma B/Q)N^p \quad (15)$$

independent of the value of e . In these expressions κ is the same constant appearing in the expression for H . We emphasize that the linear extent of a heat-removal-limited system does not depend on the dimension of the layout! Use of three dimensions rather than two merely helps with wireability, improving it so that it is no longer the dominating consideration, leaving heat removal as the dominant problem in three dimensions.

Since we assumed that $p > (e-1)/e$, we can be sure that $p > 1/2$. Comparing eqs. (10) and (15) we see that the linear dimension of the system grows with N faster in the electrical case than in the optical case. This is a consequence of the fact that the electrical

communication energy grows with length while the optical communication energy does not.

If $p < (e - 1)/e$, then \bar{r} is approximately constant, and in two dimensions the growth rate is the same for both optical and electrical interconnections. In this case, the average interconnection length does not increase with N . This corresponds to the locally interconnected case, which is less interesting than the more globally interconnected case.

In three dimensions, with $p < (e - 1)/e$, the electrical case is again worse. The reason again lies in the fact that the optical communication energy is independent of distance while the electrical communication energy is not. In both the electrical and the optical cases, the inter-element spacing grows with N due to heat removal. This can be seen from the fact that L increases in proportion to $N^{1/2}$ while in three dimensions there are only $N^{1/3}$ elements along an edge. Clearly, then, the inter-element spacing grows with N . In the optical case this causes no extra power dissipation in a given connection, but in the electrical case it does.

5. CONCLUSIONS

In two dimensions, wireability limitations are more important than heat removal, whereas the opposite is true in three dimensions.

In both the wireability-limited and heat-removal-limited cases, optical interconnections are better than electrical interconnections. The reasons lie in two facts: 1) the optical energy required per bit does not increase as strongly with length as does the electrical energy per bit, and 2) the optical delay is proportional to length, not to the square of length as in the electrical case.

Could terminating the electrical lines and propagating short pulses help the electrical case, since delay would then be proportional to length, rather than its square? For a given W , resistance increases with the length of the line. After a certain length, the lines become too lossy, and pulse transmission is not possible. Can the resistance be kept under control by making the lines wider as they become longer? No, for the lines then take up more space, requiring greater inter-element spacing, larger system size, and even longer lines.

A similar argument holds for energetic considerations. Can we save energy by sending short electrical pulses of constant length along a terminated line (instead of charging the line up), so that energy does not increase with length? No, for again the effects of resistance require that the energy per pulse increase with line length.

6. REFERENCES

1. Haldun M. Ozaktas, *A physical approach to communications limits in computing*, Ph.D. Thesis, Department of Electrical Engineering, Stanford University, Stanford, California, USA (1991).
2. H.M. Ozaktas, J.W. Goodman, "Lower bound for the communication volume required for an optically interconnected array of points", *J. Opt. Soc. Am. - A* 7, 2100-2106 (1990).

3. H.M. Ozaktas, J.W. Goodman, "The limitations of interconnections in providing communications between an array of points", in *Frontiers of Computing Systems Research, Vol. 2*, S.K. Tewksbury, Ed., Plenum Press, New York, NY, pp. 61-124 (1991).
4. H.M. Ozaktas, Y. Amitai, J.W. Goodman, "A three dimensional optical interconnection architecture with minimal growth rate of system size", *Optics Commun.* 85, 1-4 (1991).
5. H.M. Ozaktas, Y. Amitai, J.W. Goodman, "Comparison of system size for some optical interconnection architectures and the folded multi-facet architecture", *Optics Commun.* 82, 225-228 (1991).
6. H.M. Ozaktas, "Paradigms of connectivity for computer circuits and networks", to appear in *Optical Engin.*