Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

P.M.B. Vitányi

Locality, communication and interconnect length in multicomputers

69 B70, 69 C20, 69 D50, 69 F 21, 69 F22, 69 G12

Locality, Communication and Interconnect Length in Multicomputers

Paul M.B. Vitányi

Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

The lower bound on the average interconnect (wire) length in d-dimensional embeddings of "symmetric" circuits is proved to be within a small constant multiplicative factor of the obvious worst-case upper bound. The proof is simple, geometrical and works for wires with zero volume, thus increasing the range of applicability of the results. For instance, to optical (fiber) or photonic (fiberless, laser) communication. While getting rid of the 'von Neumann' bottleneck in the shift from sequential to non-seqential computation, a new communication bottleneck arises. This results from the interplay between locality of computation, communication, and the number of dimensions of physical space.

1. Introduction

In many areas of the theory of parallel computation we meet graph structured computational models. These models suggest the design of parallel algorithms where the cost of communication is largely ignored. Yet it is well known that the cost of computation - in both time and space - vanishes with respect to the cost of communication in parallel or distributed computing. As multiprocessor systems with really large numbers of processors start to be constructed, this effect becomes more and more apparent. Thinking Machines Corporation of Cambridge, Mass., has just marketed the "Connection Machine," a massively multiprocessor parallel computer. The prototype contains microscopically fine grained processor/memory cells, 65,536 of them, each with 4,096 bits of memory and a simple arithmetical unit. The communication network connecting the processors is packet-switched and based on the binary 16-cube. (A binary n-cube network consists of 2^n nodes, each node identified by an n-bit name, and an edge between nodes which differ in a single bit.) This is implemented by packing a cluster of 16 processors and one router circuit on a single chip. The 4,096 routers (in casu chips) are connected by 24,576 bidirectional wires in the pattern of the binary 12-cube. The last chapter of [Hillis1985a], "New Computer Architectures and their Relationship to Physics or, Why Computer Science is No Good," expresses the dissatisfaction of the designers with traditional computer science, "which abstracts the wire away into a costless and

volumeless idealized connection. [The] old models do not impose a locality of connection, even though the real world does. ... In classical computation the wire is not even considered. In current engineering it may be the most important thing.'' Here we shall argue that, while getting rid of the so called 'von Neumann' bottleneck*, in the shift from serial to non-serial computing, we run into a new *communication* bottleneck due to the three dimensionality of physical space.

Let us first look at some recent theoretical models for parallel computation.

(1). 'Parallel random access machines (PRAMs).' Such machines usually can, at each point in their computation, spawn a couple of offspring PRAMs to perform some subcomputations. Broadly speaking, we can therefore imagine the computation as a binary tree of processors. The 'time' the computation takes is then related to the depth of the tree.

(2). This idea is sometimes translated in terms of 'very large integrated circuits.' A complete binary tree with processors in each node is then claimed to solve NP-complete problems like the 'traveling salesman problem' in linear time. This, on the grounds that the processor at the root can send a copy of the problem instance to each of the leaves, and each of the leaves can try one candidate solution. A simple scheme can guarantee that each leaf tries a different solution, each solution is tried by some leaf, and all answers are percolated upwards to the root. If positive answers win over negative ones in the fan in, the root can determine the solution, or the absence of a solution, after it receives answers from both descendants. Cf, [Mead1980a] Chapter 8.

(3). One of the currently flourishing parts of the theory of parallel computation is 'NC-computation.' A problem is in 'Nick's Class'† if it can be solved in polylogarithmic 'time' using a polynomial number of processors. Here, 'time' means the length of the longest chain of causally related steps.

All of the above models may be valuable for investigations into the inherent parallelizability of algorithms to solve certain problems. This often takes the form of distributing copies of the entire problem instance, or pieces of the problem instance, among an exponential number of processors in a linear number of steps. Or, as in NC, among a polynomial number of processors in a polylogarithmic number of steps. The way a problem instance can be divided and partial answers put together may give genuine insight into its parallelizability. However, it can *not* give a reduction from an asymptotic exponential time best algorithm in the sequential case to an asymptotic polynomial time algorithm in *any* parallel case. At least, if by 'time' we mean time. This folklore fact can be seen easily as follows. If the parallel algorithm uses 2^n processing elements, regardless of whether the computational model assumes bounded fan-in and fan-out or not, it cannot run in time polynomial in n, because *physical space* has us in its tyranny. Viz., if we use 2^n processing elements of, say, unit size each, then the tightest they can be packed is in a 3-dimensional sphere of volume 2^n. No unit in the sphere can be closer to all other units than a distance of radius R,

---

* When the operations of a computation are executed serially in a single Central Processing Unit, each one entails a 'fetch data from memory to CPU; execute operation in CPU; store data in memory' cycle. The cost of this cycle, and therefore of the total computation, is dominated by the cost of the memory accesses which are essentially operation independent. This is called the 'von Neumann' bottleneck, after the brilliant Hungarian mathematician John von Neumann.
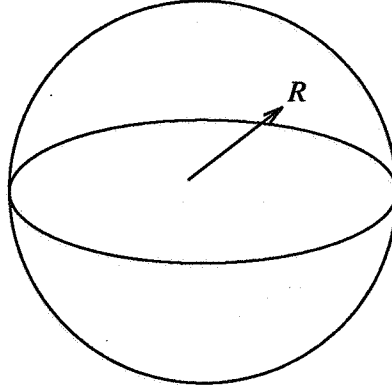† Named after Nicholas Pippenger.

Figure 1

$$R = \left[\frac{3 \cdot 2^n}{4\pi}\right]^{1/3}$$

Unless there is a major advance in physics, it is impossible to transport signals over $2^{\alpha n}$ ($\alpha > 0$) distance in polynomial $p(n)$ time. In fact, the assumption of the bounded speed of light says that the lower time bound on *any* computation using 2^n processing elements is $\Omega(2^{n/3})$ outright. Or, for the case of NC computations which use n^{α} processors, $\alpha > 0$, the lower bound on the computation time is $\Omega(n^{\alpha/3})$.*

1.1. Higher Dimensions

Of course, one may want to keep open the option of embedding physical circuits in hyper dimensions. Assume that a node (processor) has unit volume in any number of dimensions we care to consider. This, in order to obtain comparable reasoning to the physical relevant case of 3 dimensions. Our intuition about higher dimensional Euclidean geometry turns out to be quite unreliable. The Euclidean volume V_d of a d-dimensional sphere of radius R_d is

$$V_d = \frac{(R_d)^d \pi^{d/2}}{\Gamma(1+d/2)} \tag{1.1}$$

see e.g., [Hamming1980a]. With radius 1, this gives for dimensions $d=1,2,\cdots$, the volumes 2, 3.14, 4.18, 4.93, 5.26, 4.72 , 4.06, $\cdots$ The volume of the unit radius sphere comes to a maximum for $d=5$, and falls off rather rapidly toward zero as d approaches infinity. Setting $V_d=N$ and $d=2k$, we have

$$N = (\frac{\pi^k}{k!})(R_{2k})^{2k} = \frac{(\pi(R_{2k})^2)^k}{k!}$$

By Stirling's approximation,

$$R_{2k} \sim \sqrt{\frac{k}{e\pi}(N\sqrt{2\pi k})^{1/k}}$$

* It is sometimes argued that this effect is significant for large values of n only, and therefore can safely be ignored. This is a curious defense in an area were all results are of asymptotic nature, i.e., hold only for large values of n.

Differentiating, we find that R_{2k} reaches its minimum for

$$k \sim \log(N\sqrt{\pi/2}) \sim \log N \tag{1.2}$$

Therefore, $R_{2k} \geq k/\sqrt{e\pi}$, and by (1.2) we finally obtain

$$R_{2k} \geq \sqrt{\frac{\log N}{\pi}} \tag{1.3}$$

Thus, within Euclidean space, whatever the number of dimensions used, it is impossible to embed an N-node network such that nodes, which are pairwise farthest apart, have Euclidean distance less than indicated in (1.3). Therefore, the present argument allows networks with diameter $\log N$ to be embedded such that the Euclidean distance between nodes is linear in $\log N$. This, however, requires a number of physical dimensions which rises unbounded with N.

2. Physical Space and Communication: the Wiring Problem

The situation is worse than it appears on the face of it. Let us analyse the amount of wire involved. To prevent arguments that the results hold only asymptotically, or that processors are huge and wires thin, we calculate precisely without hidden constants and assume that wires have length but no volume and can pass through everything.

Note. Deriving the total necessary wire length for embeddings of networks in Euclidean space, I do not make *any* assumptions about the volume of a wire of unit length, or the way they are embedded in space. This in contrast with previous VLSI related arguments, see e.g. [Ullman1984a], which are the only ones on this issue known to me. It is consistent with the results that wires have *zero* volume, and that *infinitely* many wires pass through a unit area. Such assumptions invalidate the arguments used elsewhere. In contrast with other investigations, the goal here is to derive lower bounds on the total wire length irrespective of the ratio between the volume of a unit length wire and the volume of a processing element. The lower bound on the total wire length below is independent of this ratio, which changes with different technologies and granularity of computing components. For instance, the results also hold for optical communication networks, intraconnected by optical wave guides such as glass fiber or guideless by photonic transmission in free space by lasers.

Consider an architecture such as the binary n-cube. Recall, that this is the network with $N = 2^n$ nodes, each of which is identified by an n-bit name. There is a two-way communication link between two nodes if their identifiers differ by a single bit. The network is represented by an undirected graph $C = (V, E)$, with V the set of nodes and $E \subseteq V \times V$ the set of edges, each edge corresponding with a communication link. There are $n 2^{n-1}$ edges in C. Let C be embedded in 3-dimensional Euclidean space, and let each node have unit volume. Let x be any node of C. There are at most $2^n/8$ nodes within Euclidean distance $R/2$ of x, where R is as above. Then, there are $\geq 7 \cdot 2^n/8$ nodes at Euclidean distance $\geq R/2$ from x. Construct a spanning tree T_x in C of depth $\leq n$ with node x as the root. There are 2^n nodes in T_x and $2^n - 1$ paths from the root to a node in T_x. Let P be such a path from x to a node in T_x. Note that the number of edges $|P|$ in P is at most n. Let $l(P)$ denote the Euclidean length of the embedding of P. Since 7/8th of all nodes are at Euclidean distance at least $R/2$ of root x, the average Euclidean length of such a path P, in the 3-space embedding of C, is given by

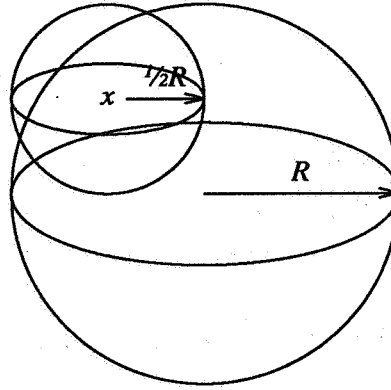$$\frac{1}{2^n - 1} \sum_{P \in T_x} l(P) \geq \frac{7R}{16} \tag{2.1}$$

Figure 2: At most ⅛th of all nodes in the large sphere are also contained in the small sphere centered on x.
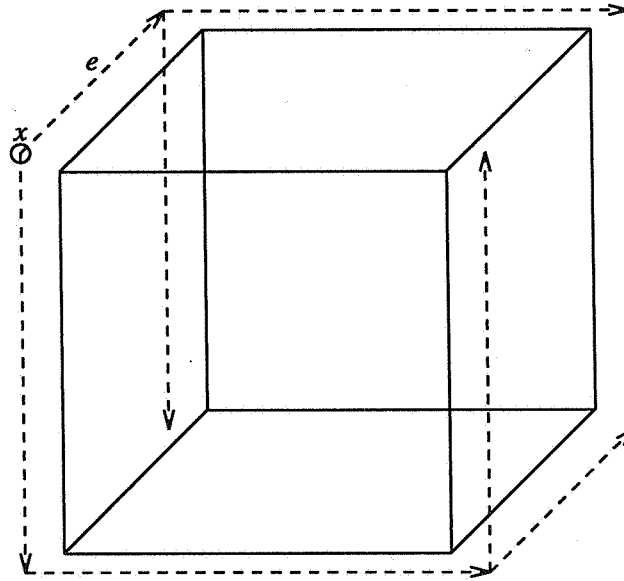
Figure 3: Binary 3-cube with spanning tree T_x.

Let $l(e))$ denote the Euclidean length of the embedding of edge e. By (2.1), the average Euclidean length of an embedded edge *in a path P* is bounded below as follows:

$$\frac{1}{2^n-1} \sum_{P \in T_x} \left[\frac{1}{|P|} \sum_{e \in P} l(e) \right] \geq \frac{7R}{16n} \tag{2.2}$$

This does not give a lower bound on the average Euclidean length of an edge, the average taken *over all edges* in T_x. To see this, note that if the edges incident with x have Euclidean length $7R/16$, then the average edge length in a path from the root x to a node in T_x is $\geq 7R/16n$, even if all edges not incident with x have length 0. However, using the symmetry of the binary n-cube we can establish the following.

Lemma 1. *The average Euclidean length of the edges in the 3-space embedding of C is greater or equal to $7R/16n$.*

Proof. Denote a node a in C by an n-bit string $a_1 a_2 \cdots a_n$, and an edge (a,b) between nodes a and b differing in the kth bit by:

$$(a_1...a_{k-1}a_k a_{k+1}...a_n, \quad a_1...a_{k-1}(a_k \oplus 1)a_{k+1}...a_n)$$

where $\oplus$ denotes modulo 2 addition. Since C is an undirected graph, an edge $e=(a,b)$ has two representations, namely (a,b) and (b,a). Consider the set A of automorphisms $\alpha_{c,j}$ of C consisting of

(1) modulo 2 addition of a binary n-vector to the node representation, followed by

(2) a cyclic rotation.

Formally, let $c=c_1 c_2 \cdots c_n$, with $c_i=0,1$ ($1 \leq i \leq n$), and let j be an integer $1 \leq j \leq n$. Then $\alpha_{c,j}: V \rightarrow V$ is defined by

$$\alpha_{c,j}(a) = b_{j+1} \cdots b_n b_1 \cdots b_j$$

with $b_i = a_i \oplus c_i$ for all i, $1 \leq i \leq n$.

Consider the set S of spanning trees of C, each tree isomorphic with T_x above,
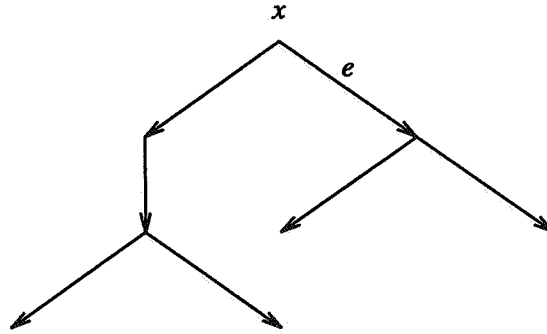
$$S = \{\alpha(T_x): \alpha \in A\}$$

Figure 4: The chosen "generic" spanning tree T_x

For each edge $e=(a,b)$ in T_x and each edge $e'=(c,d)$ of C there are exactly two distinct isomorphisms α_1 and α_2 in A such that $\alpha_1(e)=\alpha_2(e)=e'$. (Namely, $\alpha_1(a)=c$, $\alpha_1(b)=d$ and $\alpha_2(a)=d$, $\alpha_2(b)=c$.)

By the argument used to obtain (2.1) and (2.2), the average of the Euclidean length $l(\alpha(e))$ of an edge $\alpha(e)$ in a path $\alpha(P)$ from root $\alpha(x)$ to a node in a tree $\alpha(T_x) \in S$ is $\geq 7R/16n$, for *each* α in A separately. Averaging additionally over all α in A, the same lower bound applies:

$$\frac{1}{n2^n} \sum_{\alpha \in A} \left[\frac{1}{2^n-1} \sum_{P \in T_x} \left[\frac{1}{|P|} \sum_{e \in P} l(\alpha(e)) \right] \right] \geq \frac{7R}{16n} \tag{2.3}$$
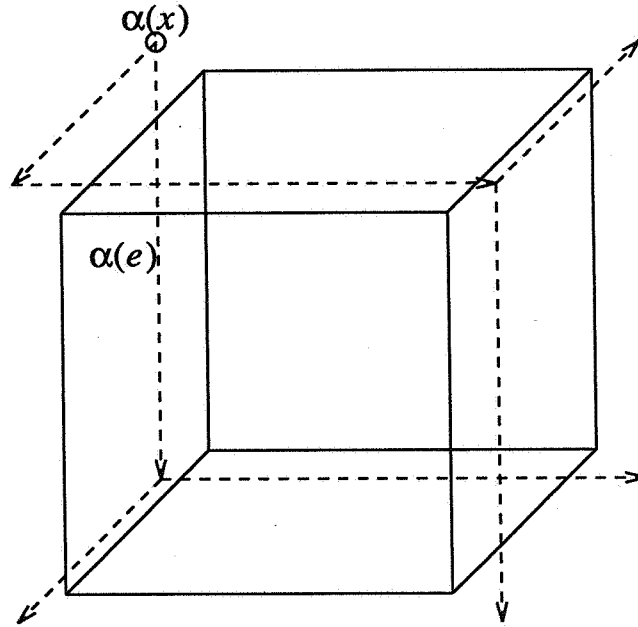
Figure 5: Binary 3-cube and rotated spanning tree $\alpha(T_x)$.

Now fix a particular edge e in T_x. We average $l(\alpha(e))$ over all α in A, and show that this average is independent of e. Together with (2.3) this will yield the desired result. For each edge e' in C there are $\alpha_1, \alpha_2 \in A$, $\alpha_1 \neq \alpha_2$, such that $\alpha_1(e) = \alpha_2(e) = e'$, and for all $\alpha \in A - \{\alpha_1, \alpha_2\}$, $\alpha(e) \neq e'$. Therefore, for each edge e in T_x, the summed Euclidean lengths of $\alpha(e)$, the sum taken over all α in A, equals twice the sum total of the Euclidean lengths of the embedded edges of C. Formally,

$$\sum_{\alpha \in A} l(\alpha(e)) = 2 \sum_{e \in E} l(e) \qquad (2.4)$$

Finally, let P be any path from root x to a node in T_x. Recall that P contains $|P|$ edges. By (2.4),

$$\sum_{e \in P} \sum_{\alpha \in A} l(\alpha(e)) = 2|P| \sum_{e \in E} l(e) \qquad (2.5)$$

Therefore, for each edge e in P, $l(\alpha(e))$ averaged over all α in A, equals the *average Euclidean edge length* $l(e)_{average}$ in the given 3-space embedding of C, and is therefore independent of P and e. By rearranging the summation order and by substitution, we obtain from (2.3) and (2.5) that $l(e)_{average}$ satisfies

$$l(e)_{average} = \frac{1}{n2^{n-1}} \sum_{e \in E} l(e)$$

$$\geq \frac{7R}{16n}$$

Since there are $n2^n/2$ edges in the binary n-cube, this sums up to an amazing *total* wire length L in any Euclidean 3-dimensional embedding of C:

$$L = \sum_{e \in E} l(e) \geq \frac{2^n 7R}{32}$$

$$\geq \left[\frac{3}{4\pi}\right]^{1/3} \cdot 7 \cdot 2^{(4n/3)-5}$$

Many network topologies are afflicted with this problem: n-dimensional cube networks, fast Fourier networks, butterfly networks, shuffle-exchange networks, cube-connected cycles networks, and so on. In fact, the arguments hold for networks with a small diameter which satisfy certain symmetry requirements. An example of a network with small diameter which is not symmetric in this sense is the tree. The fact that 7/8th of all paths from the root in a complete tree would have Euclidean length $\geq R/2$ in a 3-space embedding does not imply that the average Euclidean length of an embedded edge of the tree is larger than a constant. This is borne out by the familiar H-tree layout [Mead1980a] where the average edge length is less than 3 or 4.

Iterating this reasoning, but now adding the volume of the wires to the volume of the nodes, the greatest lower bound on the volume necessary to embed the binary n-cube converges to a particular solution in between a total volume of $\Omega(2^{4n/3})$ and a total volume of, say, $O(2^{2n})$ if we charge a constant fraction of the unit volume for a unit wire length. The lower bound $\Omega(2^{4n/3})$ ignores the fact that the added volume of the wires pushes the nodes further apart, thus necessitating longer wires again. The $O(2^{2n})$ upper bound holds under the assumption that wires of all lengths have the same volume per unit length (not more than a constant fraction of the unit volume of a node). In [Mead1982a, Vitányi1985a] it is shown that the latter assumption cannot always be made. (For instance, if we want to drive the signals to very high speed on chip.)

It is not difficult to see that no particular properties of the binary n-cube are needed to make the proof work, apart from its symmetry. Thus, the Lemma can be generalized as follows.

Let Γ be a group of permutations of a set S. Then, for each $s \in S$ the set $\Gamma(s) = \{\gamma(s): \gamma \in \Gamma\} \subseteq S$ is called the *orbit* (or transitivity set) of s. A *vertex [edge] automorphism* α of a graph $G = (V,E)$ is a permutation of V $[E]$ which preserves vertex [edge] adjacency.

A graph $G = (V,E)$ has a *vertex transitive automorphism group* [*edge transitive automorphism group*] if the group of vertex [edge] automorphisms of G has but a single orbit for each $v \in V$ $[e \in E]$, namely V $[E]$. A graph G is *symmetric* if it has both a vertex transitive automorphism group and an edge transitive automorphism group [Harary1969a]. The following property is fairly obvious.

Lemma 2. *For a symmetric graph $G = (V,E)$ and the smallest edge transitive automorphism group Γ of G there exists a constant c, such that for each pair of edges e, e' the set $\{\gamma \in \Gamma: \gamma(e) = e'\}$ has exactly c elements.*

Following the proof idea above we obtain:

Theorem. *Let R_d be the radius of a d-dimensional sphere of volume N. Let G be a symmetric N-node network. Let D be the diameter of G. Let a d-dimensional embedding of G in Euclidean d-dimensional space be such that each node has volume 1. Assume that a node is a sphere and not a "funny" form like a wire. Allow that wires have no volume and can cross through nodes in arbitrary ways. The average Euclidean length of an embedded edge in such an embedding of G is $\geq (2^d - 1)R_d / (2^{d+1} D)$.*

Proof. The obvious generalization of the proof of the Lemma 1, with the smallest edge transitive automorphism group of G in place of I, N instead of 2^n, d instead of 3, and D instead of n. ●

2.1. Higher Dimensions

For the 3-dimensional embedding of a complete graph K_N this results in an average wire length of $\geq 7R_3/16$, with $R_3 = (3N/4\pi)^{1/3}$.

Let $N = n^\delta$, n a positive integer. Define a δ-dimensional mesh with wrap-around as a set of nodes $(i_1, \ldots, i_\delta)$, $i_j = 0, \ldots, N^{1/\delta} - 1$ $(1 \leq j \leq \delta)$. Node $(i_1, \ldots, i_\delta)$ is connected by an edge with node $(j_1, \ldots, j_\delta)$, if they are equal in all coordinates except one where they differ by $1 \bmod N^{1/\delta}$.

For a 3-dimensional embedding of a N-node d-dimensional mesh with wrap-around (e.g., a ring for $d=1$, and a torus for $d=2$), this results in an average wire length of $\geq 7R_3/(8N^{1/d})$.

Again assume that a node (processor) has unit volume in any number of dimensions we care to consider. For d-dimensional embeddings of N-node, δ-dimensional meshes with wrap-around we have an avarage interconnect length $\geq (2^d - 1)R_d/(2^{d+1}N^{1/\delta})$. This lower bound is a small positive constant for $d \geq \delta$, and d is small (this is necessary because of the curious behavior of the ratio between volume and radius in higher dimensions). Since the lower bound can be matched by an upper bound, such meshes are feasible architectures for large N.

One may think that it is the unfortunate accident of having a physical space of only 3 dimensions which makes it hard to embed symmetric graphs with small diameter. However, this is not the case. Using the formula for the volume of an d-dimensional Euclidean sphere (1.1), setting $d = 2k$ and $V_{2k} = N$, we obtain by Stirling's approximation

$$
\begin{aligned}
R_{2k} &= \left[\frac{N \cdot \Gamma(1+k)}{\pi^k} \right]^{1/2k} \\
&= \left[\frac{N k^k \sqrt{2\pi k}}{\pi^k e^k} \right]^{1/2k} \\
&\sim (2\pi k)^{1/4k} N^{1/2k} \sqrt{\frac{k}{\pi e}} \\
&\geq N^{1/2k} \sqrt{\frac{k}{\pi e}}
\end{aligned}
$$

Hence it follows from the Theorem that:

Corollary. The average Euclidean length of an embedded edge in an d-dimensional embedding of a symmetric graph G, as in the Theorem, is

$$
\geq \frac{N^{1/d}}{2D} \sqrt{\frac{d}{2\pi e}}
$$

This means that for δ-dimensional meshes with wrap-around the average interconnect length exceeds

$$
\frac{N^{(\delta-d)/d\delta}}{2} \sqrt{\frac{d}{2\pi e}}
$$

2.2. Robustness

The result is apart from general also *robust*. That is, the lower bound holds within a small multiplicative factor (say 2) for networks G which can be made into a network satisfying the conditions in the theorem, by adding, deleting or collapsing (say half the number of) nodes and edges.

2.3. Scalability

These surprising facts are a theoretical prelude to many *wiring problems* currently starting to plague computer designers and chip designers alike. Formerly, a wire had magical properties of transmitting data 'instantly' from one place to another (or better, to many other places). A wire did not take room, did not dissipate heat, and did not cost anything - at least, not enough to worry about. This was the situation when the number of wires was low, somewhere in the hundreds. Current designs use many millions of wires (on chip), or possibly billions of wires (on wafers). In a computation of parallel nature, most of the time seems to be spent on communication - transporting signals over wires. Thus, thinking that the von Neumann bottleneck has been conquered by non-sequential computation, we are unaware that a Non-von Neumann communication bottleneck is waiting for us. The following innominate quote covers this matter admirably:

"Without me they fly they think;
But when they fly I am the wings."

It is clear that these communication mishaps have influence on the algorithms to be designed for the massive multiprocessors of the future, and *vice versa*, existing algorithms influence the creation of novel architectures (e.g., the k-ary n-cube Mosaic of Caltech, the Fast Fourier Transform Butterfly of Bolt, Beranek and Newman, the shuffle-exchange Ultra computer of New York university) to run them on.

Another effect which becomes increasingly important is that most space in the device executing the computation is taken up by the wires. Under very conservative estimates that the unit length of a wire has a volume which is a constant fraction of that of a component it connects, we can see above that in 3-dimensional layouts for binary n-cubes, or for the other fast permutation networks, the volume of the 2^n components performing the actual computation operations is an asymptotic fastly vanishing fraction of the volume of the wires needed for communication:

$$\frac{\text{volume computing components}}{\text{volume communication wires}} \in o(2^{-n/3})$$

The impact of these arguments points strongly in the direction of mesh connected architectures as the ultimate solution for interconnecting the extremely large (in numbers) computer complexes of the future. Mesh architectures also have the very desirable properties of *scalability, modular extensibility* and *uniformity*, when embedded in Euclidean space. It is immediate, that all circuits with a lower bound $f(n)$, $f(n) \to \infty$ for $n \to \infty$, on the average interconnect length do not scale well. Namely, composing a larger such circuit from smaller ones, the average wire length needs to increase. Thus, embeddings of such circuits are not uniformly modular extensible. This positive dependency of the interconnect length on the number of nodes to be connected we call non-scalability. Thus we have:

Corollary. No symmetric graph on N nodes with a diameter $o(N^{1/d})$ is scalable (i.e., uniformly modular extensible) when embedded in d-dimensional Euclidean space.

Today it seems that the interconnect problem for mesh architectures is solved by optical communication, either wireless by means of lasers/infrared light or by using virtually unlimited bandwidth optical fiber or integrated waveguides [Tewksbury1986a]. For instance, we can obtain three dimensional interconnect structures by stacking wafer circuit boards and providing optical interconnections vertically between wafers over the entire wafer in addition to planar connections. This may use hybrid mounting of optical components, combined with integrated optical waveguides and lenses on a large area silicon wafer-scale integrated (WSI) electronic circuit combining electronic and photonic functions [Hornak1986a]. However, it is unlikely that any solution will free us of communication problems forever. Even while Nature is not malicious, she is subtle.

Acknowledgement
    Remarks by Evangelos Kranakis, F. Tom Leighton and Yoram Moses were helpful.

References
Hamming1980a.Hamming, R., (1980). *Coding and Information Theory*: Prentice-Hall.
Harary1969a.Harary, F., (1969). *Graph Theory*: Addison-Wesley, Reading, Massachusetts.
Hillis1985a.Hillis, W.D., (1985). *The Connection Machine*: MIT Press, Cambridge, Mass..
Hornak1986a.Hornak, L.A., Tewksbury, S.K., Hatamian, M., Ligtenberg, A., Sugla, B., and Franzon, P. (September, 1986). Through-wafer optical interconnects for multi-wafer wafer-scale integrated architectures, *Manuscript*, AT&T Bell Laboratories, Holmdel, N.J..
Mead1980a.Mead, C. and Conway, L., (1980). *Introduction to VLSI Systems*: Addisson-Wesley, Reading, Mass..
Mead1982a.Mead, C. and Rem, M., (1982). Minimum propagation delays in VLSI, *IEEE J. on Solid State Circuits*, SC-17, 773 - 775. Correction: *Ibid*, SC-19 (1984) 162.
Tewksbury1986a.Tewksbury, S.K., Hornak, L.A., and Ligtenberg, A. (June, 1986). The impact of component interconnects on future large scale systems, *Memorandum*, AT&T Bell Laboratories, Holmdel, N.J..
Ullman1984a.Ullman, J.D., (1984). *Computational Aspects of VLSI*: Computer Science Press, Rockville, Maryland.
Vitanyi1986a.Vitanyi, P.M.B., (1986). Non-sequential computation and Laws of Nature, pp. 108-120 in *VLSI Algorithms and Architectures*, Springer Verlag, Berlin.
Vitányi1985a.Vitányi, P.M.B., (1985). Area penalty for sublinear signal propagation delay on chip, pp. 197-207 in *Proceedings 26th Annual IEEE Symposium on Foundations of Computer Science*.