[Li86] *Kai Li*. “Shared Virtual Memory on Loosely Coupled Multiprocessors”. PhD thesis, Yale University, New Haven, CT (1986).

[NPK$^+$94] *A. Nowatzyk, M. Parkin, E. Kelly, M. Browne, G. Aybay, and D. Lee*. The S3.mp Scalable Shared Memory Multiprocessor. In “Proceedings of the HICSS”, Mountain View, CA (1994).

[PMM93] *Douglas M. Pase, Tom MacDonald, and Andrew Meltzer*. MPP FORTRAN Programming Model. Technical Report, Cray Research, Inc. (November 1993).

[SSC93] *Rafael H. Saavedra, Gaines R. Stockton, and J. Carlton, Michael*. Micro Benchmark Analysis of the KSR1. In “Proceedings of the Supercomputer Conference” (1993).

[SWG91] *J.P. Singh, W.D. Weber, and A. Gupta*. Automatic and Explicit Parallelization of an N-body Simulation. Technical Report, Stanford University (1991).

4.3.2 Interconnection Networks for Parallel and Distributed Systems

by Harald Richter

4.3.2.1 Introduction

Various types of interconnection may feature in local or global computer networks, telephone systems, and networks for parallel computers. Qualitative und quantitative classification of these is important for the design, testing, and operation of networks. A distinction has to be made between point-to-point, multicast, inverse multicast, broadcast and conference links and the exact number of interconnection combinations present has to be known to allow proper determination of the state and capacity of a network. The increasing availability of ATM switch elements has given rise to growing interest in network techniques that are not based on the time sharing principle of a common medium, e.g. Ethernet or FDDI, but have spatially separate interconnection paths between individual ports. This contribution describes the type and number of parallel interconnections in a general network, presents formulae for quantitatively determining them and explains these by means of examples. Emphasis is placed on interconnection networks for parallel computers.

4.3.2.2 Types of Interconnections and their Quantitative Analysis

Assume we have a general interconnection network I with n $(n = 1, 2, ...)$ connected users $U_1 - U_n$, which may be, for example, computers or computer nodes such as are applied in local networks or parallel computers. In I various types of interconnection can be distinguished for the n ports present: Firstly there is the point-to-point type of interconnection, where one port is coupled with any other port. Secondly, there is the broadcast type of interconnections, where any port can be coupled with all the others. Thirdly, there is the multicast type, which is the generalization of the first two types, and in which the interconnection is made between one port and several (m) other ports, where $0 \leq m \leq n$. The multicast type yields the point-to-point case for m = 1 and the broadcast type for m = n. There are two other cases which not only play a role in counting all combinations of interconnection types, but also have a certain practical importance, viz. the interconnection of a port with itself (e.g. for testing purposes) and the interconnection of a port with no other port (tristate or m = 0).

Besides these 5 types of interconnection, there are 4 types of operations that have to be distinguished: unidirectional transmission, unidirectional reception, bidirectional, and inactive mode. The inactive mode, where a user neither transmits or receives, can be merged with the tristate type of interconnections to yield a single case, leaving just 3 modes. The various types of interconnection and operation that can be present for each port or user can be arranged in 15 different combinations. An example of the combinability of interconnection types with operation modes is a transmitting port coupled with several receiving ports (multicast) or a receiving port with several transmitting ports (inverse multicast). The inverse multicast is also referred to as the gather function and is often applied in interconnection networks for parallel computers in order, to compute the sum of the elements of a distributed vector from a computer node for example. A second example of combinability is several bidirectional ports coupled with several other bidirectional ports, yielding a conference link where each user can receive the other's data. (In order to avoid interference, it is necessary that only one transmitter be active at any one time and that the others receive.) The conference link constitutes complete intermeshing of a group of users and can be composed of either individual multicast or inverse multicast types. For this reason it is sufficient, without restriction of generality, to consider just the two basic types, multicast and inverse multicast, since all other cases such as point-to-point, broadcast, and conference links can be traced back to these.

The total of 15 combinations of interconnection types and operation modes define at each individual port 15 different states which the port can assume. The question that arises from this is: how many states or combinations of interconnections the complete network can have? The answer to this question is important because this number allows one to determine the interconnection capability and diversity of a network. Capability may be defined as the property whereby a network is strictly non blocking, non blocking by rearrangement of internal paths or blocking [Wu91]. Interconnection diversity denotes whether a network allows point-to-point links only or also multicast and inverse multicast, broadcast, and conference links [Hwa93].

STATES IN THE NETWORK

The question of the total number of states in the network is not trivial, since with n ports and 15 states per port there cannot be just 15^n different network states. This is because, on the one hand, the ports are not independent of one another, i.e. when a port is transmitting, for example, another has to receive; and because, on the other hand, one must take into account which port is coupled with which. It thus follows that the number of network states must be much higher than 15^n. If one is restricted to, for example, point-to-point interconnections, n simultaneously active users will result in n! permutations of potential interconnections [HB85]. This number can become very large. In the case of the parallel computer CM-5, for example, the complete configuration affords an interconnection network coupling 16 K processors with one another [Lei92]. This network has to be capable of switching 16K! = 10^{160000} different point-to-point interconnections. (The number of atoms in the universe, in contrast, is "only" approx. 10^{65}.) The point-to-point interconnections alone thus already impose high requirements on the network. In the following, we shall calculate the total number of all states possible in an general network, taking the individual states at the ports into account. For this purpose we proceed stepwise, first determining the number of combinations, i.e. states, in the network which can arise from all the possible multicast interconnections of all ports with one another. This is facilitated by imagining that all receivers of a certain multicast are merged into one receiver group. This allows us to assign all (active) receivers of the network to different multicast groups. The receiver groups are disjunctive, i.e. no receiver can belong simultaneously to more than one group. In the second step we shall then calculate the inverse multicast interconnections, whereupon it will be found that for symmetry reasons the results of the (normal) multicast can be taken over. Finally, in the third step, we calculate the total number of all states possible in a network.

NUMBER OF MULTICAST INTERCONNECTIONS

To determine the number of multicast interconnections, we proceed on the following basis: In the initial state of the network all ports are inactive and no port is coupled with any other. An arbitrary first transmitter can then "select" its first receiver from a batch of n receivers (including itself). With the 1st transmitter in multicast mode there are then still n - 1 candidates available for selection for its 2nd receiver, n - 2 for its 3rd receiver, etc., so that there are $n - i_1$ receivers left at the end, where i_1 denotes the number of receivers connected with the 1st transmitter. It holds that $0 \leq i_1 \leq n$. The selection procedure described continues with the 2nd transmitter, it selects its i_2 receivers from the remaining $n - i_1$ receivers, etc. In the end at most n transmitters are connected with their i_1 to in multicast receivers, and it holds that $i_1 + i_2 + ...i_n \leq n, 0 \leq i_j \leq n$. As the multicast case contains the inactive case for $i_j = 0, j = 1, ..., n$, the point-to-point case for $i_j = 1$, and the broadcast case for $i_j = n$, only the multicast type need be considered. For at most n transmitters in multicast there can be a maximum of n groups of receivers, where each group can have between 0 and n receivers. (Of course, the sum of all receivers from all groups can be at most equal to n.) For selection of the i_1 members of the first group there are (Eq. 1) possibilities [HJ81] since the sequence of selection within a group of multicast receivers plays no role. Accordingly, there are still (Eq. 2) selection possibilities for the i_2 members of the 2nd receiver group, so that for the last, i.e. n-th, transmitter there are (Eq. 3) choices of receivers left. As the total number of possibilities is given by the product of the individual possibilities, we obtain (Eq. 4) for the total number K of multicast states in a general interconnection network. This can also be written in the form (Eq. 5) [Ric95] which is a new result.

K is thus the number of multicast combinations between transmitters and receivers that is possible with at most n transmitters and their n groups of at most n receivers per group. K is a function of the variables i_1 to i_n, which determine the number of receivers in each group where n is a free parameter. The equation for K allows one to cover all combinations of interconnections where not more than one transmitter is coupled to a receiver group. The inverse multicast, i.e. gather function, is not covered. The equation for K can be explained as follows: K is the volume of an n-dimensional cube formed from n mutually perpendicular axes of the length (Eq. 6.) The lengths of the coordinate axes correspond to the number of possibilities of combining the at most

$$\text{Eq. 1}: \quad \binom{n}{i_1}$$

$$\text{Eq. 2}: \quad \binom{n-i_1}{i_2}$$

$$\text{Eq. 3}: \quad \binom{n-i_1-i_2-...-i_{n-1}}{i_n}$$

$$\text{Eq. 4}: \quad \binom{n-i_1}{i_2}...\binom{n}{i_1}\binom{n-i_1-i_2-...-i_{n-1}}{i_n}$$

$$\text{Eq. 5}: \quad K=\prod_{\text{j=1}}^{\text{n}}\binom{F_j}{i_j} \text{ with } F_j=n-\sum_{l=1}^{j-1} i_l \text{ and } 0 \le i_j \le F_j$$

Figure 4.11: Eq. 1 to Eq. 5.

n receiver groups. Three simple examples may illustrate (Eq. 5): For $i_1 = i_2 = ... = i_n = 1$ we obtain (Eq. 7) which is the known result for point-to-point permutations [HB85]. And for $i_1 = n, i_2 = i_3 = ... = i_n = 0$ we have (Eq. 8) which is the one and only possibility for a broadcast from the first transmitter to all receivers. The third special case of (Eq. 5) is present if every receiver is assigned a receiver group, i.e. if the receiver groups are complete. We then have (Eq. 9). With this additional constraint, (Eq. 5) can be significantly simplified to (Eq. 10) which is also a well known result in combinatorics [BS83]. In practice it may often happen that no receiver is inactive, then (Eq. 10) can be applied.

NUMBER OF INVERSE MULTICASTS

In this section we determine the number of combinations of interconnections for the inverse multicast case in which more than one transmitter is connected with any receiver. For this purpose we utilize the fact that inverse multicasts are mirror images of (normal) multicasts, so that all results obtained for these are transferable provided that the roles of transmitters and receivers are interchanged. The decomposition $i_1, i_2, .., i_n$ is thus now used to denote the assignment of at most n transmitters to at most n disjunctive groups, where all transmitters of one group are connected with the same receiver, and no transmitter can be connected with more than one receiver. As in the normal multicast case, $0 \le i_j \le n(j = 1, ..., n)$ is allowed here as well, i.e. the number of transmitters in a group can vary between 0 and n. Of course, it must hold that $i_1 + i_2 + .. + i_n \le n$. With the application of (Eq. 5) to transmitter groups it is possible to determine, differentiated with respect to transmitter groups, the number of inverse multicasts, also including point-to-point connections for n = 1. For the special case where $i_1 + i_2 + .. + i_n = n$, (Eq. 5) can likewise be simplified to (Eq. 10).

Eq. 6 :
$$\binom{Fj}{ij}$$

Eq. 7 :
$$K = \binom{n}{1}\cdot\binom{n-1}{1}\cdot\binom{n-2}{1}\cdots\binom{n-(n-1)}{1} = n!$$

Eq. 8 :
$$K = \binom{n}{n}\cdot\binom{n-n}{0}\cdot\binom{n-n}{0}\cdots\binom{n-n}{0} = \binom{n}{n} = 1$$

Eq. 9 :
$$\sum_{j=1}^{n} i_j = n, \text{ with } 0 \le i_j \le n$$

Eq. 10 :
$$K = \frac{n!}{i1!i2!\cdots in!}$$

Figure 4.12: Eq. 6 to Eq. 10.

NUMBER OF CONFERENCE LINKS

The number of conference links is calculated on the same lines: The number n of users is assigned to at most n disjunctive groups of at most n users each. (The exact number of users per group is defined by the variables $i_1, i_2, .., i_n$) In each group all users of a certain conference link are present. Accordingly, (Eq. 5) and (Eq. 10) are valid.

TOTAL NUMBER OF MAPPINGS

The total number of mappings is given by the combination of multicast and inverse multicast connections, with each transmitter and receiver now being allowed to belong to several groups simultaneously. Any distinction according to disjunctive groups is therefore no longer appropriate. In order to determine the total number of all possible mappings, we abandon differentation regarding individual groups and consider the interconnection of at most n transmitters and receivers with one another as the set of all mappings of a transmitter vector T of n elements onto a receiver vector R. Here each individual mapping is performed by means of a composition matrix C, i.e. $R = C * T$. The composition matrix C has the dimension nxn and contains only variables of the quality "connected" or "not connected". That is, every element of C is of the Boolean type. The set of all different mappings C is given by the variation of all elements of C. With n^2 Boolean elements, one obtains 2^{n^2} mappings. Therefore the number of all combinations of connections in a general interconnection network is $K = 2^{n^2}$. This equation may be less useful than (Eq. 5) because it cannot differentiate with respect to transmitter and receiver groups. Finally, it is useful to order the number of combinations . This is accomplished in (Eq. 11).

4.3.2.3 Results

The various cases of interconnection have been described and summarized into two basic categories: multicast and inverse multicast out of which point-to-point, broadcast and conference links can be constructed. The combinations of interconnection types with operation modes has lead to 15 different states each port in the net can assume. The question how many states the net can have was answered by computing the number of multicast combinations, because for symmetry reasons the equations hold for inverse multicast and conference links as well. The computation was performed by assigning each active receiver to a receiver group, calculating the number of all sets of receiver groups and permuting them with all active transmitters. The resulting (Eq. 5) can be simplified if all n receivers are simultaneously active. The point to-point type of communication is always included as a special case of each of these types. The full set of interconnection possibilities was derived by using a mapping of Boolean elements from transmitters to receivers by means of a matrix-vector multiplication. By disregarding the decomposition of receivers or transmitters into groups the known results were obtained. The advantage of (Eq. 5) is that it allows more flexibility by explicitly defining the number of participants of each group. This can be used to decide about the interconnection capability and diversity of a given network. The results are summarized in Figure 4.13.

Eq. 11:

$$2^{(n^2)} >> n^n >> n! \geq \frac{n!}{i1!\,i2!\cdots in!} \geq \prod_{\mathrm{j=1}}^{\mathrm{n}} \binom{Fj}{ij} \quad \text{for n > 1}$$

Total number of combinations for connections in a network with n bidirectional ports:

$$2^{n^2}$$

Number of combinations of connections, if network is restricted to point - to - point, multicast, or inverse multicast connections. Traffic is decomposed into groups of i_j receivers or transmitters:

$$K = \prod_{\mathrm{j=1}}^{\mathrm{n}} \binom{F_j}{i_j} \quad \text{with} \quad F_j = n - \sum_{l=1}^{j-1} i_l \quad \text{and } 0 \leq i_j \leq Fj \quad \text{and} \sum_{l=1}^{j-1} i_l = 0 \text{ for } j = 1$$

Simplification of above equation if all network ports are simultaneously active:

$$K = \frac{n!}{i_1!\,i_2!\cdots i_n!}$$

Figure 4.13: Eq. 11 and Results.

References

[BS83] *I. Bronstein and K. Semendjajew*. " Taschenbuch der Mathematik ". Harri Deutsch, Frankfurt (1983).

[HB85] *K. Hwang and F. A. Briggs*. "Computer Architecture and Parallel Processing, pp. 332-354, 481-508". McGraw-Hill (1985).

[HJ81] *Hockney and Jesshope*. "Parallel Computers". Adam Hilger, Bristol (1981).

[Hwa93] *K. Hwang*. “Advanced Computer Architecture, pp. 75-96, 331-347”. McGraw-Hill (1993).

[Lei92] *C. Leiserson et al.* The Network Architecture of the Connection Machine CM5. In “4th Annual ACM Symp. on Parallel Algorithms and Architectures” (1992).

[Ric95] *H. Richter*. Quantitative Classification of Interconnections . In *B. Hertzberger and G. Serazzi*, editors, “Proceedings of the HPCN’95.”, page 942, Heidelberg (May 1995). Springer Verlag.

[Wu91] *C. L. Wu*. Tutorial on System Integration with Interconnection Networks. In “Int.Conf. on Par. Proc.”, St. Charles,ILL. (August 1991).