BUS-ORIENTED INTERCONNECTION TOPOLOGIES FOR SINGLE-HOP COMMUNICATION AMONG MULTI-TRANSCEIVER STATIONS

Yitzhak Birk
IBM Almaden Research Center
San Jose CA 95120

Fouad A. Tobagi
Stanford University
Stanford CA 94305

Michel E. Marhic
Northwestern University
Evanston IL

ABSTRACT

We present the topological design space of single-hop interconnections among multi-transceiver stations. (No intermediate switches and no forwarding.) Next, we explore a class of such interconnections whose throughput with a uniform traffic pattern is proportional to the *square* of the number of transceivers per station. These interconnections consist of a collection of buses, each of which interconnects a proper subset of the stations, and are referred to as *Selective-Broadcast Interconnections*. They are compared with a multiple-broadcast-bus interconnection in terms of station hardware, throughput, delay and power budget.

1. INTRODUCTION

1.1 Single-Hop Interconnections

We define a single-hop interconnection ($\mathcal{SHI}$) to be one in which a message travels from the sender to the recipient without any intervention; i.e., no intermediate switches (as in multistage interconnections) and no need for forwarding by other stations (as in multi-hop networks). The interconnection network can thus be entirely passive. $\mathcal{SHI}$'s are often desired due to their inherent reliability, low latency and simplicity in operation and maintenance, and are sometimes an absolute requirement.

1.2 Why Multi-Transceiver Stations?

Consider the single broadcast bus, commonly used in local-area networks (LAN's).* It is the simplest $\mathcal{SHI}$ and the only possible one among single-transceiver stations. The $\mathcal{SBB}$ can be characterized as follows.
(i) The transmission rate B must at least equal the aggregate network throughput S. Consequently, very fast channels are often required, resulting in high cost as well as in an inability to use certain physical media.
(ii) Even the smallest user of the network must be capable of transmitting and receiving at a rate that is equal to the aggregate network throughput. This can cause the cost of the communication interface to dominate the cost of the stations.
(iii) Since the bus is shared among all N stations, the average utilization of a station's communication hardware cannot exceed $1/N$ (averaged over stations and time, assuming single-destination messages). The $\mathcal{SBB}$ appears to lend itself to bursty transmissions, since a single transmission can use the entire communication bandwidth; at first glance, this suggests that one should use an $\mathcal{SBB}$ and make every effort to increase the transmission rate. However, network nodes can seldom take advantage of this due to the limited speed of their internal data paths. This mismatch between "raw" communication bandwidth and the rate that a node can sustain is most pronounced in high speed fiber optic networks.
(iv) The use of fiber optics in the implementation of an $\mathcal{SHI}$ raises a power budget problem. This is due to the noncoherent detection and low impedence of detectors. The power required at an optical receiver is known to be proportional to the transmission rate over a broad range of rates [1], corresponding to a requirement of a minimal number of photons per bit. An important manifestation of the power budget is a limitation on the number of stations. The actual numbers vary with the physical topology used to implement the bus; for the physical topologies required in conjunction with access schemes that are relatively efficient at high transmission rates, the numbers are very small. [2]
In order to remedy the above limitations, the interconnection must permit some degree of concurrency. To do so in an $\mathcal{SHI}$, each station must be equipped with multiple, say C, transmitters and/or receivers,

This research was carried out at Stanford University; it was supported in part by NASA and in part by IBM through a graduate student fellowship to Y. Birk.

*A token ring is logically similar to $\mathcal{SBB}$, but differs greatly in power budget.

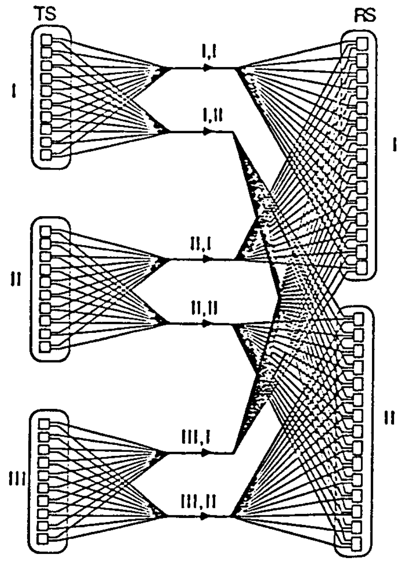

Fig. 1 Single-path, unidirectional, equal-degree $\mathcal{SBI}$ with disjoint subnetworks and grouping. $N = 30$; $C_T = 2$; $C_R = 3$.

and multiple buses must be used.* The straightforward way of interconnecting C-transceiver stations is using C broadcast buses. We refer to this as the *parallel broadcasts interconnection* ($\mathcal{PBI}$).

1.3 Notation and Representations

For purposes of clarity and generality, $\mathcal{SHI}$'s will be represented by unidirectional interconnections connecting N_T transmitting stations, each with C_T transmitters, to N_R receiving stations, each with C_R receivers. Fig. 1 depicts an $\mathcal{SHI}$. The general representation includes the interconnection of N bidirectional stations as a special case. We use B to denote the signaling rate on a channel, which is directly proportional to the data rate of an individual transmission, and S denotes the aggregate network throughput. We define *power split* to be the number of receivers reached by a transmitter and denote it by PS.

1.4 Performance Measures

In comparing $\mathcal{SHI}$'s, our primary performance measures are throughput and delay; single-destination traffic is assumed unless stated otherwise, and the throughput is always stated in conjunction with a traffic pattern. While $\mathcal{SHI}$'s invariably employ shared channels thus requiring access schemes to regulate the sharing [3,4], our focus is on topology. A perfect access scheme will generally be assumed, but the effect of imperfect schemes on the comparison will be stated. Also, we will use the term *concurrency* (defined to be the throughput with an ideal access scheme, usually expressed in packets per packet transmission time, thus representing the level of parallelism) instead of throughput.

Another aspect of an interconnection's performance, which is of particular importance to fiber optic implementations, is the *power budget*. This will be touched upon briefly; for more details, see [5].

1.6 Outline

In section 2, we describe the design space of $\mathcal{SHI}$'s and state some of their properties. Section 3 is devoted to a special class of $\mathcal{SHI}$'s, referred to as unidirectional, equal degree $\mathcal{SBI}$'s with grouping and disjoint subnetworks, and section 4 discusses a bidirectional derivative of this class. The focus is on performance in terms of concurrency. Section 5 is devoted to a delay comparision between one such $\mathcal{SBI}$and $\mathcal{PBI}$. Section 6 outlines several implementation and application issues, and section 7 summarizes the paper.

2. $\mathcal{SHI}$ DESIGN SPACE

2.1 Classification of $\mathcal{SHI}$'s

A given $\mathcal{SHI}$ can be characterized as possessing or not possessing each of the following properties:

- *Standard Stations.* All TS's have an equal number of transmitters, and all RS's have an equal number of receivers.
- *Equal Degree.* All transmitters reach an equal number of receivers, and all receivers can be reached by an equal number of transmitters. In the bipartite graph representing such an $\mathcal{SBI}$, all nodes on the left (transmitters) have an equal degree, as do all the nodes on the right (receivers).
- *Grouping in the Weak Sense of TS's (RS's).* Two TS's, say i and j, are said to belong to the same *group* if and only if they have equal numbers of transmitters and, for each transmitter of i, there is a transmitter of j such that the two transmitters reach identical sets of receivers. Similarly for RS's with receivers and transmitters exchanged.

*B is the signaling rate on the channel. We therefore exclude spread-spectrum techniques as means of permitting S to exceed B, in spite of the fact that concurrent transmissions are possible.

- *Disjoint Subnetworks.* The sets of receivers reached by any two transmitters are either identical or disjoint.
- *Grouping in the Strong Sense.* In addition to grouping in the weak sense, transmitters of TS's that belong to different groups reach disjoint sets of receivers (similarly for RS's).
- *Bidirectional $\mathcal{SHI}$.* Station i can reach station j over a given subnetwork if and only if j can reach i over the same subnetwork.

An $\mathcal{SHI}$ can also be characterized by

- The *number of disjoint paths* between each TS and RS. (A k-path $\mathcal{SHI}$.)

Selective-Broadcast Interconnections. We refer to the subclass of $\mathcal{SHI}$'s in which (i) a transmission is heard by a destination-dependent proper subset of the stations and (ii) the concurrency for a uniform traffic pattern grows quadratically with an increase in C as "selective-broadcast interconnections", or $\mathcal{SBI}$'s. In [6], we presented specific instances of this class, focusing on fiber optic implementation. In this paper, we cover a broader range of $\mathcal{SHI}$'s and include delay in the performance comparisons.

2.2 Properties of $\mathcal{SHI}$'s

Proposition 1. Grouping in the strong sense implies disjoint subnetworks; a bidirectional $\mathcal{SHI}$ always has disoint subnetworks; bidirectional media implies disjoint subnetworks; bidirectional $\mathcal{SHI}$'s cannot have grouping in the strong sense (except $\mathcal{SBB}$ and $\mathcal{PBI}$). The proofs are left out for brevity. See [5].

Proposition 2. For a uniform traffic pattern, the maximum concurrency of a k-path $\mathcal{SHI}$ with disjoint subnetworks is $C_T \cdot C_R/k$; The maximum utilization of transmitters is $N_T/(kC_R)$, and that of receivers is $N_R/(kC_T)$; the minimal power split is $k \cdot N_R/C_T$. (Utilization and power split for a given interconnection are taken to be the worst over all transmitters or receivers.) The bounds are tight.

Proof. From symmetry considerations and since the various measures are taken to be the worst case, it follows that all subnetworks are of the same size. In a k-path $\mathcal{SHI}$, a TS must reach at least $k \cdot N_R$ receivers using C_T transmitters; therefore, each subnetwork must include at least kN_R/C_T receivers. Similarly, it must include kN_T/C_R transmitters. This proves the utilization and power-split bounds. Dividing the total number of transmitters or receivers by the size of each subnetwork, we see that one can have at most $C_T \cdot C_R/k$ subnetworks, hence the maximum concurrency. Tightness will be shown by construction. □

The discussion in the remainder of this paper is limited to $\mathcal{SHI}$'s with standard stations and disjoint subnetworks (bus-oriented); the term grouping will be used synonymously with grouping in the strong sense.

3. UNIDIRECTIONAL, EQUAL-DEGREE $\mathcal{SBI}$'s WITH GROUPING

3.1 The Unidirectional, Single-Path $\mathcal{SBI}$ [6]

3.1.1 Construction

The TS's and RS's are arranged in C_R and C_T disjoint groups of equal sizes, respectively. Next, $C_T \cdot C_R$ subnetworks are constructed, such that subnetwork (i,j) connects the jth transmitter of every TS in group i to the ith receiver of every RS in group j. Fig. 1 depicts such an $\mathcal{SBI}$; observe that each TS has only one subnetwork in common with any given RS. When $N_T{=}N_R{=}N$ and $C_T{=}C_R{=}N-1$, this $\mathcal{SBI}$ reduces to a fully connected topology with a point-to-point link from each TS to each RS; when $C_T{=}C_R{=}1$, it reduces to $\mathcal{SBB}$.

3.1.2 Comparison with $\mathcal{PBI}$

Unlike the concurrency with $\mathcal{PBI}$, which is always C, the concurrency provided by this $\mathcal{SBI}$varies between 1 and C^2, depending on the traffic pattern. A comparison can therefore only be conducted for specific patterns.

Uniform traffic pattern and single-destination transmissions

In this case, the $\mathcal{SBI}$'s subnetworks are equally loaded. Consequently,

$$S^{\mathcal{SBI}} = K^{\mathcal{SBI}} \cdot C_T^{\mathcal{SBI}} \cdot C_R^{\mathcal{SBI}} \cdot B^{\mathcal{SBI}}. \qquad (1)$$

For $\mathcal{PBI}$,

$$S^{\mathcal{PBI}} = K^{\mathcal{PBI}} \cdot C^{\mathcal{PBI}} \cdot B^{\mathcal{PBI}}. \qquad (2)$$

K is a constant which depends on the channel access scheme. ($0 < K \leq 1$.) For the time being, ideal access schemes are assumed, so $K = 1$. To permit comparison, it is also assumed that $C_T{=}C_R$ and that N_T and N_R are the same in both systems. The above expressions can then be interpreted in several ways:

- With $C^{\mathcal{SBI}} = C^{\mathcal{PBI}} = C$ and $B^{\mathcal{SBI}} = B^{\mathcal{PBI}}$, the aggregate throughput of $\mathcal{SBI}$ is C times higher than that of $\mathcal{PBI}$, since it increases quadratically rather than linearly with C.
- With $S^{\mathcal{SBI}} = S^{\mathcal{PBI}}$ and $C^{\mathcal{SBI}} = C^{\mathcal{PBI}} = C$, the transmission rate required with $\mathcal{SBI}$ is C times

lower than that required with $\mathcal{PBI}$; i.e., slower (and cheaper) transmitters and receivers may be used for the same throughput.

- With $S^{\mathcal{SBI}} = S^{\mathcal{PBI}}$ and $B^{\mathcal{SBI}} = B^{\mathcal{PBI}}$, $C^{\mathcal{SBI}} = \sqrt{C^{\mathcal{PBI}}}$; i.e., $\mathcal{SBI}$ requires fewer transmitters and receivers.

Since each subnetwork of $\mathcal{SBI}$ serves only N/C transmitting stations and N/C receiving stations, as compared with N in $\mathcal{PBI}$, the average utilization of transmitters and receivers is C times higher than that of $\mathcal{PBI}$. Similarly, the power split is C times smaller. This $\mathcal{SBI}$ is optimal in the sense of proposition 2.

Multi-destination packets. Concurrency with $\mathcal{SBI}$ is up to C times higher than that with $\mathcal{PBI}$, depending on the number of transmitters that must be used in $\mathcal{SBI}$ to reach all destinations.

Nonuniform traffic patterns. When all the traffic is from one TS group to one RS group, the concurrency with this $\mathcal{SBI}$ drops to one, whereas that with $\mathcal{PBI}$ is always C. Also, for any given source-destination pair, the maximum instantaneous data rate with $\mathcal{SBI}$ is B, as compared with $C \cdot B$ with $\mathcal{PBI}$.

3.2 Muli-Path, Unidirectional $\mathcal{SBI}$

The fact that the concurrency with the single-path $\mathcal{SBI}$ can be as low as 1 and as high as C^2, depending on the traffic pattern, whereas that of $\mathcal{PBI}$ is always C, raises the question of whether one can do better than $\mathcal{PBI}$ without sacrificing flexibility. Lang, Valero and Fiol [7] addressed this question, with the assumption that a single station never does more than one thing at a time. Given C buses, they therefore consider performance not to be degraded as long as any C source-destination pairs, such that no source or destination appears more than once, can be accommodated concurrently. They have shown that the maximal combined reduction in the total number of transmitters and receivers is $C(C+1)$. The relative saving, $(C+1)/N$, becomes negligible as the number of stations increases.

From the above results, it follows that there is a trade-off between the maximum concurrency for a uniform traffic pattern, $\mathcal{C}_{\max}$, and the guaranteed concurrency, $\mathcal{C}_{\min}$(for any pattern); $\mathcal{PBI}$ and the single-path $\mathcal{SBI}$ are two extremes. We next present 3 parameterized compromises. In all cases, the guaranteed (minimum) concurrency is equal to the number of alternate paths between any two stations, and the maximum concurrency is equal to the number of disjoint subnetworks.*

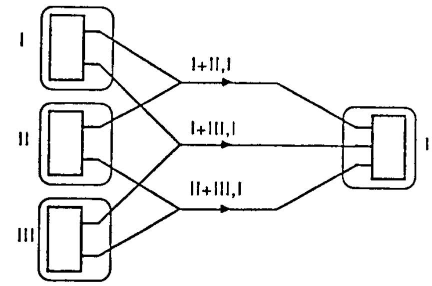

Fig. 2 Unidirectional 2-path equal-degree $\mathcal{SBI}$ with disjoint subnetworks and grouping. (*SMP*.) A representative station is shown for each group. G_T=3, k_T=2, G_R=k_R=1; C_T=2, C_R=3.

Multiple single-path $\mathcal{SBI}$'s (*MSP*)
m single-path $\mathcal{SBI}$'s are constructed, each of which utilizes *1/m* of the transmitters and receivers. The concurrency provided by an *MSP* and the required values of C are as follows:

$$\mathcal{C}_{\min} = m; \; \mathcal{C}_{\max} = \frac{C_T C_R}{m}; \text{ Power split} = m\frac{N_R}{C_T}. \quad (3)$$

A single multiple-path $\mathcal{SBI}$ (*SMP*)
The sets of TS's and RS's are partitioned into G_T and G_R groups, respectively. A subnetwork is constructed to connect each possible combination of k_T groups of TS's with each possible combination of k_R groups of RS's. Fig. 2 depicts an *SMP*. The performance of the *SMP* is as follows:

$$\mathcal{C}_{\min} = \binom{G_T - 1}{k_T - 1} \cdot \binom{G_R - 1}{k_R - 1}; \; \mathcal{C}_{\max} = \binom{G_T}{k_T} \cdot \binom{G_R}{k_R};$$

$$C_T = \frac{k_T}{G_T} \cdot \binom{G_T}{k_T} \cdot \binom{G_R}{k_R}; \quad C_R = \frac{k_R}{G_R} \cdot \binom{G_T}{k_T} \cdot \binom{G_R}{k_R};$$

$$\text{Power split} = k_R \cdot \frac{N_R}{G_R}. \quad (4)$$

A hybrid $\mathcal{SBI}$ -$\mathcal{PBI}$ interconnection
C' transmitters and C' receivers of each station are are used for a $\mathcal{PBI}$, and the remaining ones are used for a single-path $\mathcal{SBI}$. An example of such an $\mathcal{SBI}$ is

*Our criterion for worst-case concurrency is stronger than that in [7], since we require k disjoint paths between any two stations. Our criterion is equivalent to fault tolerance.

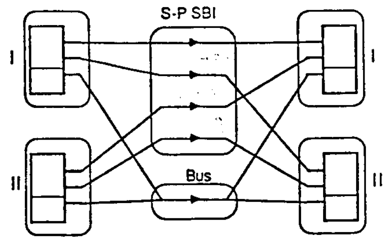

Fig. 3 Unidirectional hybrid $\mathcal{SBI}$-$\mathcal{PBI}$. C=3, C'=1.

depicted in Fig. 3. Note that the hybrid is not equal-degree, and grouping is only in the weak sense. The performance is as follows:

$$\mathcal{C}_{\min} = C'+1; \qquad \mathcal{C}_{\max} = C'+(C_T-C')\cdot(C_R-C');$$

$$\text{Power split (worst case)} = N_R. \tag{5}$$

Comparison

To simplify the comparison, let $C_T = C_R = C$, $G_T = G_R = G$, $k_T = k_R = k$, and $N_T = N_R = N$. The comparisons, conducted by equating C and $\mathcal{C}_{\min}$ for the different topologies and then comparing $\mathcal{C}_{\max}$ and the power split show *MSP* and *SMP* to be equal. Furthermore, they are optimal in the sense of proposition 2. While this suggests that the two are merely different ways of describing the same interconnections, this is not the case.

Using *MSP* terminology, the hybrid configuration has

$$\mathcal{C}_{\max} = (C-m+1)^2 + m - 1, \tag{6}$$

as compared with C^2/m for *SMP* and *MSP*. It can be shown that the performance is equal if $m = C$ or $m = 1$, and the hybrid performs better in all other cases. (This is proved by showing that, for any given C, the difference is zero only at two values of m, namely 1 and C, and that for $m = C/2$ the hybrid is always superior.) Furthermore, the performance advantage of the hybrid increases as C increases, for any fixed value of m. By letting $m = \alpha C$, $0 < \alpha < 1$ and substituting in the above equations, we see that

$$\mathcal{C}_{\min} = \Omega(C), \qquad \mathcal{C}_{\max} = \Omega(C^2). \tag{7}$$

The hybrid thus incorporates the performance advantages of the two constituent configurations, up to a constant factor. However, it is as bad as the $\mathcal{PBI}$ in terms of utilization and PS due to its $\mathcal{PBI}$ part.

The hybrid has another advantage, namely the flexibility in the allocation of hardware to the two components. This is illustrated in Fig. 4, which shows, for each of the three configurations, all feasible combinations of $\mathcal{C}_{\min}$ and $\mathcal{C}_{\max}$ subject to the constraint that $1<\mathcal{C}_{\min}<\mathcal{C}_{\max}$ and $C \leq 20$. One can also see a significant advantage of *MSP* over *SMP* in this respect. Triangles, boxes and plus signs correspond to *SMP*, *MSP* and the hybrid, respectively.

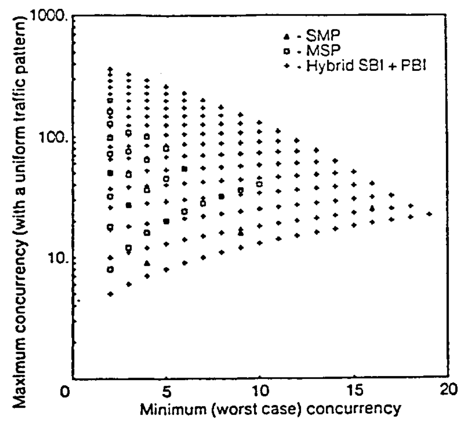

Fig. 4 Feasible ($\mathcal{C}_{\min}$,$\mathcal{C}_{\max}$) combinations with unidirectional *SMP*, *MSP*, and a hybrid.

4. BIDIRECTIONAL (k, C)-PATH $\mathcal{SBI}$

These $\mathcal{SBI}$'s provide k paths between any two stations in different groups and C paths between stations in the same group.

4.1 The Bidirectional $(1, C)$-Path $\mathcal{SBI}$

This $\mathcal{SBI}$ is derived from the unidirectional single-path $\mathcal{SBI}$ with $C+1$ transmitters and receivers per station by (i) combining subnetwork (i, j) with (j, i) for $1 \leq i, j \leq C+1$, and (ii) doing away with subnetworks (i, i) along with one transmitter and one receiver per station. Stated differently, the stations are divided into $(C+1)$ groups, and a subnetwork is constructed between each pair of groups. There are $\frac{C\cdot(C+1)}{2}$ subnetworks, which also provide C paths between any two stations within a group, hence the characterization as a $(1, C)$ $\mathcal{SBI}$. The $(1, 3)$ bidirectional $\mathcal{SBI}$ is depicted in Fig. 5.

4.2 Multi-Path Bidirectional $\mathcal{SBI}$'s

The performance tradeoff as well as the proposed compromises are similar to those in the unidirectional case, and subnetworks now interconnect groups of bidirecional stations.

Multiple $(1, C/m)$, Equal-Degree Bidirectional $\mathcal{SBI}$'s (*MSP*)

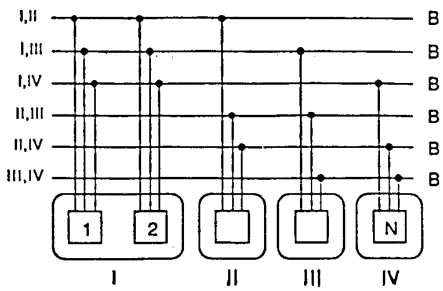

Fig. 5 A bidirectional representation of a bidirectional, $(1, C)$-path $\mathcal{SBI}$ with equal-degree and grouping. $C = 3$.

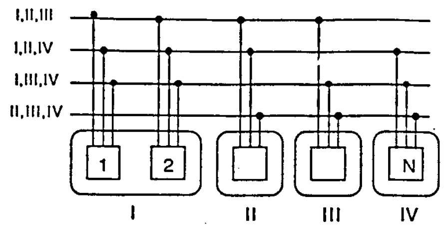

Fig. 6 Bidirectional *SMP*. $G = 4$, $k = 3$, $C = 3$.

These are simply m identical $(1, C/m)$ bidirectional $\mathcal{SBI}$'s, each using $\frac{C}{m}$ transceivers per station. The performance is (assuming m divides C)

$$\mathcal{C}_{\text{min}} = m; \quad \mathcal{C}_{\text{max}} = \frac{C(C+m)}{2m}; \quad \text{PS} = \frac{2mN}{C}. \tag{8}$$

A Single Multiple-Path, Equal-Degree Bidirectional $\mathcal{SBI}$ (*SMP*)

As depicted in Fig. 6, each subnetwork interconnects k of the G groups of stations.

$$\mathcal{C}_{\text{min}} = \binom{G-2}{k-2}; \quad \mathcal{C}_{\text{max}} = \binom{G}{k}; \quad C = \frac{k \cdot \binom{G}{k}}{G}.$$

$$\text{Power split } = k \cdot \frac{N}{G}. \tag{9}$$

Hybrid $\mathcal{SBI}$ -$\mathcal{PBI}$ interconnection

The performance with the hybrid configuration is

$$\mathcal{C}_{\text{min}} = C' + 1; \quad \mathcal{C}_{\text{max}} = C' + (C - C') \cdot (C - C' + 1)/2;$$

$$\text{Power split (worst case) } = N. \tag{10}$$

Note that this hybrid $\mathcal{SBI}$ is not equal-degree.

Comparison

Again, we equate $\mathcal{C}_{\text{min}}$ and solve for $\mathcal{C}_{\text{max}}$ and for PS. The result is

$$\left(0.5 + \frac{k}{2G}\right) \cdot \mathcal{C}_{max}^{SMP} \leq \mathcal{C}_{max}^{MSP} \leq \mathcal{C}_{max}^{SMP}, \tag{11}$$

with equality only in the case that both reduce to a single bus (k=G) or to a single $(1, C)$-path $\mathcal{SBI}$ ($k = 2$). As for power split:

$$\text{PS}^{MSP} = \frac{2(k-1)}{G-1} \cdot N \geq \frac{k}{GN} = \text{PS}^{SMP}. \tag{12}$$

The bidirectional *SMP* thus also provides a better power split than the *MSP*.

Hybrid

$$C = \frac{k}{G} \cdot \binom{G}{k}; \qquad C' = \mathcal{C}_{\text{min}} - 1 = \binom{G-2}{k-2} - 1;$$

$$\mathcal{C}_{\text{max}} = \binom{G}{k} \frac{k}{2G(G-1)} \left[\frac{(G-k)^2}{k-1} \binom{G\text{-}2}{k\text{-}2} + 3G\text{-}k\text{-}2 \right]. \tag{13}$$

It can be shown that the hybrid outperforms the *SMP*, with equality only when $k = 2$, $G - 1$, or G [5]. The hybrid is also more flexible in terms of feasible combinations. The actual choice between *SMP* and *MSP* would again depend primarily on feasibility.

5. DELAY COMPARISON OF $\mathcal{SBB}$, $\mathcal{PBI}$ AND THE SINGLE-PATH $\mathcal{SBI}$

5.1 Outline

Here, we compare the performance of the single-path, equal degree, unidirectional $\mathcal{SBI}$ with grouping with those of $\mathcal{PBI}$and $\mathcal{SBB}$. (Delay is the time interval from the creation of a message until its successful delivery to its destination). Focusing on the effect of topology, we assume an ideal access scheme. A uniform traffic pattern is assumed, but a nonuniform pattern can be easily accommodated by looking at the busiest subnetwork.

5.2 Queueing Models

$\mathcal{SBB}$ will be modeled as a $G/G/1$ queue. The single-path $\mathcal{SBI}$ will be modeled as a collection of independent $G/G/1$ queues. The relative arrival rate to each of those queues is a function of the traffic pattern; nevertheless, they can be analyzed separately. $\mathcal{PBI}$ will be assumed to permit only bit-serial transmissions; however, concurrent transmissions between a pair of stations are possible. Therefore, it will be modeled as an $G/G/m$ queue (a single job can be served by only one server, but this can be any of the servers). For simplicity in obtaining rough results, we will assume Poisson arrivals and exponentially distributed packet lengths. Let λ and $\frac{1}{\mu}$ denote the mean rate of packet generation and the mean packet transmission time, respectively. For an $M/M/1$ queue, the mean packet delay is then given by

$$T = \frac{\frac{1}{\mu}}{1-\frac{\lambda}{\mu}}. \tag{14}$$

For an $M/M/m$ queue, let p_0 denote the probability that the system is empty and $\rho \triangleq \frac{\lambda}{m \cdot \mu}$. It has been shown that [8]

$$p_0 = \left[\sum_{k=0}^{m-1} \frac{(m\cdot\rho)^k}{k!} + \left(\frac{(m\rho)^m}{m!}\right)\left(\frac{1}{1-\rho}\right)\right]^{-1}. \tag{15}$$

It can also be shown, by finding the mean number of messages in the system and using Little's theorem, that the mean delay is given by

$$T = \frac{p_0}{\lambda}\left\{\sum_{k=1}^{m} \frac{k\,(m\rho)^k}{k!} + \frac{m^m}{m!}\cdot\frac{\rho^{m+1}}{(1-\rho)^2}\,(m\,(1\text{-}\rho)+1)\right\}. \tag{16}$$

5.3 Comparison

The comparison will be conducted for the same three cases that were used in the concurrency comparison. For clarity, λ_0 and μ_0 will denote the values of λ and μ for $\mathcal{SBB}$, and the values for $\mathcal{SBI}$ and $\mathcal{PBI}$ will be expressed in terms of these. Also, m will denote the number of buses in $\mathcal{PBI}$; the number of subnetworks in the $\mathcal{SBI}$ will be a case-dependent function of m. As before, S represents the aggregate throughput (concurrency) B is the transmission rate, and C is the number of transmitters and receivers per station.
In all three cases, average delay for $\mathcal{SBB}$ is

$$T_{\mathcal{SBB}} = \frac{\frac{1}{\mu_0}}{1-\frac{\lambda_0}{\mu_0}}. \tag{17}$$

1) Equal S for all three schemes; equal B for $\mathcal{SBI}$ and $\mathcal{PBI}$.
$\mathcal{PBI}$ with m buses: the following substitutions should be made in the $M/M/m$ results.

$$\lambda = \lambda_0;\ \mu = \frac{\mu_0}{m};\ \rho = \frac{\lambda_0}{\mu_0}. \tag{18}$$

$\mathcal{SBI}$ with m subnetworks: substituting $\lambda = \frac{\lambda_0}{m}$ and $\mu = \frac{\mu_0}{m}$ in the $M/M/1$ results yields

$$T_{\mathcal{SBI}} = m \cdot T_{\mathcal{SBB}}. \tag{19}$$

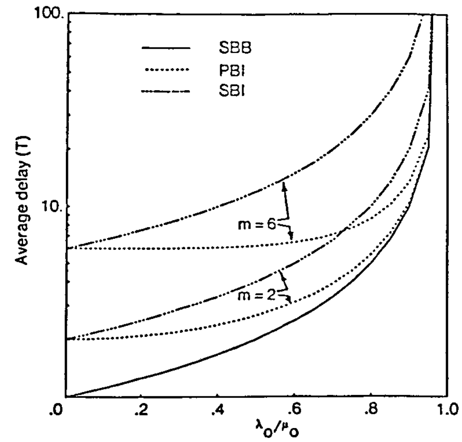

Fig. 7 Delay comparison of $\mathcal{SBI}$, $\mathcal{PBI}$ and $\mathcal{SBB}$. Equal S for all; equal B for $\mathcal{SBI}$ and $\mathcal{PBI}$.

Fig. 7 presents the mean packet delay as a function of λ_0/μ_0 for $\mathcal{SBB}$, $\mathcal{PBI}$ and $\mathcal{SBI}$; results are presented for $m = 2, 6$. At low values of λ_0/μ_0, $\mathcal{PBI}$ and $\mathcal{SBI}$ perform equally well, but both are m times worse than $\mathcal{SBB}$. This is so because in this range delay consists primarily of the transmission time, which is m times longer for $\mathcal{PBI}$ and $\mathcal{SBI}$ than for $\mathcal{SBB}$ due to the m-fold fragmentation of the aggregate bandwidth. As λ_0/μ_0 increases, the queueing delay becomes a major factor. Here, the fact that $\mathcal{PBI}$ permits any message to be transmitted over any channel enhances its performance with respect to $\mathcal{SBI}$ and brings it closer to that of $\mathcal{SBB}$. As the load on the interconnection grows even further, the delay with $\mathcal{PBI}$ becomes essentially equal to that with $\mathcal{SBB}$, whereas that with $\mathcal{SBI}$ remains m times higher.

2) Equal S for all schemes; equal C for $\mathcal{SBI}$ and $\mathcal{PBI}$.
$\mathcal{PBI}$: same as 1). (m buses.)
$\mathcal{SBI}$: there are now m^2 subnetworks, each with a transmission rate that is $1/m^2$ of the aggregate bandwidth. Therefore,

$$T_{\mathcal{SBI}} = m^2 \cdot T_{\mathcal{SBB}}. \tag{20}$$

Fig. 8 presents the mean packet delay as a function of λ_0/μ_0 for $\mathcal{SBB}$, $\mathcal{PBI}$ and $\mathcal{SBI}$; results are presented for $m = 2, 6$.

3) Equal $B \cdot C$ for all schemes; equal B and C for $\mathcal{PBI}$ and $\mathcal{SBI}$.
$\mathcal{PBI}$: same as 1). (m buses.)
$\mathcal{SBI}$: There are m^2 subnetworks, each with transmis-

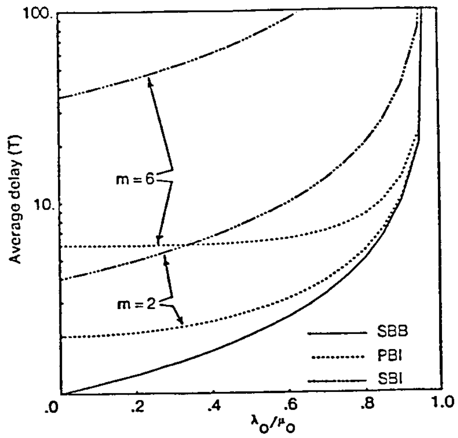

Fig. 8 Delay comparison of $\mathcal{SBI}$, $\mathcal{PBI}$ and $\mathcal{SBB}$. Equal S for all; equal C for $\mathcal{SBI}$ and $\mathcal{PBI}$.

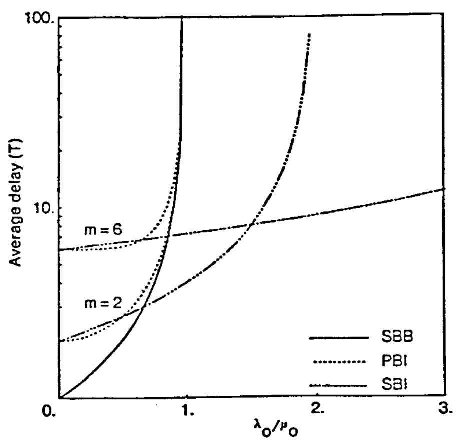

Fig. 9 Delay comparison of $\mathcal{SBI}$, $\mathcal{PBI}$ and $\mathcal{SBB}$. Equal S for $\mathcal{SBB}$ and $\mathcal{PBI}$; equal B and C for $\mathcal{SBI}$ and $\mathcal{PBI}$.

sion rate μ/m. Therefore, $\lambda = \lambda_0/m^2$, $\mu = \frac{\mu_0}{m}$, and

$$T_{\mathcal{SBI}} = \frac{\frac{m}{\mu_0}}{1 - \frac{\lambda_0}{m \cdot \mu_0}}. \tag{21}$$

Fig. 9 presents the mean packet delay as a function of λ_0/μ_0 for $\mathcal{SBB}$, $\mathcal{PBI}$ and $\mathcal{SBI}$; results are presented for $m = 2, 6$. The performance of $\mathcal{SBB}$ as well as that of $\mathcal{SBI}$ are again unchanged. As for $\mathcal{SBI}$, there are m^2 subnetworks, (as in case 2,) but each of them has the same bandwidth as a $\mathcal{PBI}$ bus. (Recall that the constraint here is equal transmission rate per station, not equal bandwidth of the medium.) As a result, the delay at low load is equal to that with $\mathcal{PBI}$ and m times higher than with $\mathcal{SBB}$. As the load increases, $\mathcal{PBI}$ becomes superior to $\mathcal{SBI}$ due to the pooling of the servers; however, the difference is very slight. As the load increases further, the fact that the capacity with $\mathcal{SBI}$ is m times higher begins to play a major role, and $\mathcal{SBI}$ begins to outperform $\mathcal{PBI}$. Finally, as the load approaches $\lambda_0/\mu_0 = 1$, the delay with $\mathcal{SBB}$ and $\mathcal{PBI}$ grows very rapidly, whereas that with $\mathcal{SBI}$ exhibits only a moderate growth, since the load on an $\mathcal{SBI}$ subnetwork is only $\frac{\lambda_0}{m^2}/\frac{\mu_0}{m} = \frac{1}{m} \cdot \frac{\lambda_0}{\mu_0}$. In fact, $\mathcal{SBI}$ outperforms $\mathcal{SBB}$ for

$$\frac{\lambda_0}{\mu_0} > \frac{m-1}{m - \frac{1}{m}}. \tag{22}$$

This is of course in addition to the fact that $\mathcal{SBI}$ can carry m times more traffic with finite delay.

Summary. The above comparison leads one to conclude that, under a constraint of fixed media bandwidth, (i) the $\mathcal{SBI}$ is always inferior in terms of delay and (ii) it is desirable (from a standpoint of delay) that the hardware savings of $\mathcal{SBI}$ (compared with $\mathcal{PBI}$) be in the form of a reduced number of transceivers rather than a reduced transmission rate. However, this is guaranteed to be true only when the coefficient of variation of packet lengths is less than 1. [9] Under a constraint of fixed bandwidth per station, the comparison between $\mathcal{SBB}$ and $\mathcal{PBI}$ is not affected. However, the performance of $\mathcal{SBI}$ improves due to the fact that the aggregate transmission bandwidth increases as the fragmentation increases. This counteracts the negative effects of fragmentation, and $\mathcal{SBI}$ can actually outperform $\mathcal{SBB}$ and $\mathcal{PBI}$. (The crossover point depends on the packet length distribution.)

6. IMPLEMENTATION AND APPLICATIONS

In this section, we briefly present several issues pertaining to the implementation and applications of $\mathcal{SBI}$. For more detail, see [5].

Access schemes. Bidirectional $\mathcal{SBI}$'s can be operated using any of the existing access schemes, with the modifications necessary for choosing the appropriate transmitter. In the case of multi-path $\mathcal{SBI}$'s, multi-bus protocols can be adapted. [4]

The use of unidirectional $\mathcal{SBI}$'s is more problematic, since a station cannot sense all the channels over which it transmits. Possible remedies are the addition of sensing hardware, the use of schemes that do not require sensing (ALOHA, TDMA), and the use of the intragroup subnetworks (i, i) by the TS's in each group to coordinate their transmissions.

The use of a nonideal access scheme improves the performance of $\mathcal{SBI}$ relative to that of $\mathcal{PBI}$. The reason is the reduced efficiency of access schemes at high transmission rates [10, 11]. ($\mathcal{SBI}$ requires a lower rate than does $\mathcal{PBI}$ for achieving any given throughput.)

Subnetwork separation. The separation among the subnetworks need not be spatial; it can be by frequency*, time, or any combination thereof. Some of these separation methods result in hardware savings, passive as well as active. The separation in time is particularly intriguing, since such an $\mathcal{SBI}$ requires only a single transmitter and a single receiver per station; unfortunately, this has no significant throughput advantage over $\mathcal{SBB}$.

Hardware Utilization. The increased utilization of transmitters and receivers is particularly important in board-level interconnections among VLSI chips, since the pin count is often the performance bottleneck and an underutilized I/O pin is a true waste.

Media utilization. Certain physical media offer high bandwidth but a very limited data rate.** This is due to dispersion. Fragmenting the aggregate data rate into several channels (frequency separation), as is done by $\mathcal{SBI}$, can thus greatly increase the aggregate data rate. In other words, there are cases in which going from $\mathcal{SBB}$ to $\mathcal{SBI}$ will actually require no additional media.

Fault tolerance. While this was not addressed explicitly, our definition of worst-case concurrency is such that is is equal to the number of faults required for breaking connectivity.

7. SUMMARY

The topological design space of single-hop interconnections, which are usually thought of as either a single or multiple broadcast buses, is actually quite rich. For a uniform traffic pattern,† we showed how to construct interconnections whose throughput is proportional to the square of the number of transmitters and receivers per station. In other words, equipping stations with multiple transceivers and using multiple buses is not merely a necessary evil; it may actually be cost effective.

Clearly, $\mathcal{SBI}$'s cannot compete with multi-stage interconnections or with multi-hop ones in terms of performance [12]; however, in applications limited by throughput or transmission-rate, there may be an interesting performance range which is beyond the capabilities of an $\mathcal{SBB}$ but short of warranting a complex interconnection network, for which $\mathcal{SBI}$'s are appropriate; this, of course, is in addition to those applications in which one is limited to $\mathcal{SHI}$'s.

The unidirectional $\mathcal{SHI}$'s can also be interpreted as bidirectional interconnections between two groups of stations, such that intragroup communication is not required. This directly applies to shared-memory multiprocessor machines. In fact, the concept of $\mathcal{SHI}$'s is more general and can be applied in fields other than communication [5].

Lastly, it should be noted that the discussion here was restricted to $\mathcal{SHI}$'s with disjoint subnetworks. With unidirectional media, such as fiber optics with directional star couplers, more general $\mathcal{SHI}$'s can be constructed [5]

*We consider only non agile modems, as is the case in fiber optic and certain radio applications. With agile modems, one could simply allocate a frequency to each RS; this would permit concurrency of N with a single transmitter and receiver per station.

**An example is a coaxial cable; the same cable that can support hundreds of 5MHz TV channels can only support a transmission rate of several Mbps when used as an $\mathcal{SBB}$.

†Only a coarse uniformity is required, one which suffices to cause equal loading of the subnetworks.

REFERENCES

[1] T.V. Muoi, "Receiver design for high-speed optical-fiber systems," *IEEE J. Lightwave Tech nology*, vol. 2, pp. 243-267, June 1984.

[2] M.M. Nassehi, F.A. Tobagi and M.E. Marhic, "Fiber optic configurations for local area networks," *IEEE JSAC*, vol. SAC-3, no. 6, pp. 941-949, Nov. 1985.

[3] F.A. Tobagi, "Multiaccess protocols in packet communication systems," *IEEE Trans. Commun.*, vol. COM-28, Apr. 1980.

[4] M.A. Marsan and D. Roffinella, "Multichannel local area network protocols", it IEEE JSAC, vol. SAC-1, no. 5, pp. 885-897, Nov. 1983.

[5] Y. Birk, Ph.D. Dissertation, Stanford University, 1987

[6] Y. Birk, F.A. Tobagi and M.E. Marhic, "Selective-broadcast interconnections (*SBI*) for wideband fiber-optic local area networks", *Proc. SPIE Conference on Fiber Optic Broadband Networks*, Cannes, France, Nov. 1985.

[7] T. Lang, M. Valero and M.A. Fiol, "Reduction of Connections for Multibus Organization," *IEEE Trans. Comp.*, vol. C-32, no. 8, pp. 707-715, Aug. 1983.

[8] L. Kleinrock, *Queueing Systems, vol. I: Theory.* Wiley-Interscience, New York, NY, 1975.

[9] H. Kanakia and F.A. Tobagi, "On Distributed Computations with Limited Resources", *IEEE Trans. Comp.*, vol. C-36 no. 5, pp. 517-528, May 1987.

[10] F.A. Tobagi and V.B. Hunt, "Performance analysis of carrier sense multiple access with collision detection," in *Proc. Local Area Commun. Network Symp.*, Boston, MA, May 1979.

[11] F. Tobagi, F. Borgonovo and L. Fratta, "Expressnet: a high-performance integrated services local area network," *IEEE JSAC*, vol. SAC-1, Nov. 1983.

[12] H.J. Siegel, *Interconnection Networks for Large-Scale Parallel Processing*, Lexington books, Lexington, MA, 1985.