ABSTRACT

The traffic characteristics of two burst-switched network topologies are evaluated. These ring and tree configurations have potential application in the local-access-area of the public switched telephone network. Performance aspects are addressed primarily through simulation. A brief overview of burst switching is given and network models are described. Methodology and results from a performance evaluation of the two integrated networks are presented and discussed. Sensitivity of voice freezeout, data delay, and data queue sizes to traffic loads are investigated for the two network topologies. Finally, a promising control scheme for the tree topology is evaluated.

1. INTRODUCTION

Efficient integration of voice and data in a digital network is a goal for network designers. A new network architecture based on the burst switching concept [1,2,3] shows promise as a basis for such an integrated network. Presented in this paper are performance evaluation analyses and results for the burst-switched local-access-area derived from mathematical models and computer simulations. Mathematical modeling is limited due to the advanced features of burst switching and the complexity of voice and data interactions. An overview of burst switching is given below followed by a presentation of network topological design considerations. Two topologies are investigated herein. Simulation models for a ring network and a tree network are discussed in Section 2. Voice and data traffic models used in the simulations are described in Section 3.

Ring and tree traffic performance results and a comparison of the two topologies will be given in Section 4. In all cases in this paper, the results assume zero community of interest (COI), i.e., all calls terminate outside the local-access-area. This limitation is considered appropriate since an assumed zero COI produces worst case performance. Also, various estimates of voice traffic COI for the group size we are considering range from only 1% to 10%, depending on population density and numerous other factors.

Another parameter included in the simulation models but not investigated in this paper is intentional voice delay. This parameter was investigated previously in the burst switched ray configuration [3] and it was shown that a small degree of improvement in voice freezeout performance could be obtained with intentional voice delay in a range around 10 ms. These voice-only results are similar for the ring and tree configurations.

Computer simulation is a powerful tool which makes possible the study of numerous network conditions and combinations of parameters. Simulation results and exact mathematical formulae should be compared where possible to test modeling assumptions. Mathematical results applicable to the burst-switched local-access-area are available for only idealized situations. We use these results where possible (Section 4) but have chosen not to limit use of the simulation to those cases. In fact, the simulations allow many possible areas of study beyond the scope of this paper and these are listed in Section 5. Also in Section 5 is a summary of the conclusions of this study.

1.1 Review of Burst Switching

Burst switching has been a significant area of research at GTE Laboratories for over four years. It employs a highly distributed network architecture which places link switches in the outside plant. In a burst switched network, information is transmitted in variable-length packets called bursts which can contain either voice or data. Transmission is channelized and channels are occupied only when speech energy or data are present. Speech/silence detectors are placed at each voice port in the switching node. Voice has priority over data but will not preempt a data burst. If there are more active sources than available channels, speech will be lost (freezeout) or data will be queued only for as long as congestion exists. Otherwise, bursts are sent through the switching nodes (link switches) with negligible delay and storage. That is, burst switching differs from conventional packet switching in that the information is not routinely stored in the switching nodes. Channels are dynamically allocated as bursts pass through intermediate link switches. End-to-end delay is small due to a burst header of only four bytes for voice and six bytes for data.

1.2 Network Topological Design Considerations

Selection of a network topological must take into account many factors, such as:

- **Node Complexity** — Includes both the processing capability required and the number of processors as determined by the degree of the node.
- **Path Length and Delay** — Usually a function of the number of hops in the path (however, an example is shown later of a short path which does not reduce delay).
- **Network fault tolerance** — How well a network is able to withstand component failures.
- **Outside Plant Engineering** — The physical connectivity of switches as opposed to the logical connectivity. Two objectives in the digitalization of the public switched telephone network are: (1) use existing outside plant, and (2) minimize customer loop length.
• Modularity — How easily new links and switches can be added to an existing network.
• Traffic — The throughput, delay, and availability characteristics of a network.

No topology has a clear advantage for all applications. Two basic topologies, the tree and ring, are selected for investigation in this paper. Relative to the ring topology, the tree generally offers delay and modularity advantages but disadvantages in node complexity and network fault tolerance. Outside plant considerations are heavily dependent on the application. The many factors which play a role in network design suggest that there is a need for topological alternatives examined in this paper.

2. LOCAL-ACCESS-AREA NETWORK MODELS

The first configuration considered consists of a variable number of link switches (LS) connected in a ring by T1-rate (1.544 Mb/s) transmission links (Figure 1). The bidirectional ring is connected to the outside world by two adjacent interface link switches (ILS) which do not support traffic sources. This entire configuration is referred to as a link group.

![Figure 1. Burst switching local access area ring configuration](image-url)

The second configuration considered consists of two connected binary trees each with link switches again connected by T1 links. Only binary trees are considered in order to minimize the degree of the nodes. The tree link group is shown in Figure 2. Two ILSs are also present in the tree configuration. The path to an ILS in the tree generally has a small number of hops. This fact and the fact that the two trees are connected makes possible alternate routing of congested speech bursts (as opposed to conventional circuit-switched telephone networks which perform alternate routing of calls).

Traffic on both the ring and tree may be classified as either internal or external. Internal traffic both originates and terminates on the link group and is routed to avoid the ILSs. External traffic is routed via the shortest path, i.e., through the nearest ILS. External voice traffic sources are assumed to be in pairwise conversation, so there is exactly one source outside the link group associated with each external voice source on the link group. Only external traffic is considered in this paper.

Both network models are represented by discrete event simulations developed in the GPSSH language. The simulations can be used to examine behavior of the network under various conditions, e.g., number of link switches, transmission channels, and number of voice and data sources.

![Figure 2. Tree configuration with six link switches](image-url)

In the ring simulation, eight link switches are connected by links with 24 T1 channels. The number of switches simulated is not varied because the ring results are identical for any number of switches greater than six when there is zero COI and other parameters are held constant. The reason is that blockage occurs in no more than three switches in tandem, even in the congested cases which will be considered. The addition of more than six switches (three on each half of the ring) will not affect the distribution of congestion on the links. Another constant condition in the ring simulation is the location of the host computers. They are assumed to be connected to link switches 4 and 5, so that the longest possible path is taken by bursts going off the link group from the hosts and also by external bursts coming on the link group to the hosts.

The tree simulation models six switches. Only two levels of the binary trees are modeled because congestion does not occur in the lower levels of the tree for the cases considered in this paper.

In both models each half acts independently of the other when neither COI nor rerouting for congestion is present. In both the ring and tree simulations error-free channels are assumed. Also, traffic sources are uniformly distributed among the switches in the ring. The tree and ring have the same number of sources at the nodes adjacent to the Interface Link Switches in order to get a direct comparison of the two configurations. The remaining sources in the tree are uniformly distributed.

3. VOICE AND INTERACTIVE DATA MODELS

The voice traffic model is based on empirical results published in [4]. This model represents measured English conversational speech through use of talkspurt and pause length probability distributions. Voice burst mean and standard deviation are 284 ms and 241 ms, respectively. Silent interval mean and standard deviation are 480 ms and 1379 ms, respectively. The resultant speech activity level is 37.1%. It is important to use a realistic voice model. O'Reilly [5] shows that data performance in an integrated voice-data environment is greatly affected by the assumed voice talkspurt and pause distributions.
The interactive data traffic model is derived from measurements of actual computer-terminal traffic in a scientific environment found in [6]. In order to incorporate this data model in a communication network simulation, it is decomposed into: (i) a model of traffic generated by the user at a terminal, and (ii) a model of traffic generated by the computer in response to users.

Interarrival times are the sum of computer reaction time, user time, and transmission periods in both directions. The transmission periods are 0.2% of the total interval and are ignored. Therefore, the single user interarrival time distribution is formed by convolving the distributions for the two remaining times.

The single user model is used to form an aggregate model for multiple users. The random arrival of bursts from a single terminal can be modeled as a stationary renewal point process. Therefore the aggregate model is formed by a composition of renewal processes as described in [7]. Users are assumed to be independent. The aggregate model represents traffic from a number of interactive terminals on a common controller. We model 32 terminals on a common controller which has a single data port at the link switch. Mean values for the model parameters, assuming 64 kbps channels, are as follows:

- mean burst length (terminal to host) = 7 ms (448 bits)
- mean burst length (host to terminal) = 36 ms (2304 bits)
- mean interarrival time (single terminal, both directions) = 27323 ms
- mean interarrival time (32 terminals, both directions) = 854 ms

Note that the total channel utilization is only 0.008 and 0.042 in the two directions for the 32 interactive sessions. The simulation allows any mix of terminal and host traffic to enter and leave the link group; however, all simulation results in this paper have the same number of host and terminal sources both entering and leaving the link group.

The procedure for forming an aggregate model can be applied to any number of terminals. However, as the number of terminals increases the interarrival distribution approaches the exponential distribution. The general distribution from the actual measurements is used in this paper.

4. RING AND TREE PERFORMANCE RESULTS

Link switches support both voice and data traffic sources. Performance of integrated voice and data is evaluated using an interactive data model, following a brief discussion below of voice-only analyses. Voice performance is measured by cutout fraction which is the fraction of speech lost due to freezeout. Freezeout occurs as front-end clipping of the speech burst. Data performance is expressed primarily in terms of end-to-end queueing delay in the link group and also maximum queue length at a link switch.

In general, the ring and tree models yield similar voice-only freezeout results. Results are also very similar to those from the ray configuration discussed in [3]. Weinstein [8] shows that voice freezeout in the single node case is dependent only on the number of active sources, the mean activity level, and the number of channels. When one considers the multi-node case, the total mean freezeout is independent of the number and topology of link switches (see [9]) as long as the traffic is entirely external; the number of channels does not vary between nodes, and the particular local access area (LAA) considered has one link to the rest of the world. To obtain voice-only freezeout of sources at each link switch separately, the method of [9] can be used.

4.1 DATA DELAY AND QUEUE LENGTH RESULTS

Link group delays are investigated in order to determine regions of operation where the burst-switched local-access area will not contribute significantly to response times for interactive data users. Also, we will look at the relationship between delay and freezeout. Queue lengths provide indications of memory requirements in link switches for various levels of supported traffic.

Delay sensitivity to a variation in data traffic volume will be investigated first. Delay is defined to be total queueing delay through all link switches in the burst path within the local-access area. Processing delays are assumed negligible due to the cut-through nature of burst switching and low burst overhead. The 95th percentile of the delay distribution is used as a figure-of-merit.

The number of off-hook voice sources is held constant at 98 while the number of data sources is varied. The 95% data queueing delays resulting from simulation are shown in Figure 3. Delay increases steadily up to about 4800 data terminals as shown in Figure 3. With heavier data traffic, delay increases more rapidly. The number of data terminals was varied for both the ring and tree in the same increments to get a direct comparison between the two configurations. As seen in Figure 3, delays for the two configurations are very similar over the entire range of data terminals. This result occurs even though the bursts in the tree are traversing fewer links than in the ring. Speech freezeout is increasing, of course, while data sources are being added. This result is shown in Figure 4 for the same cases as in Figure 3. Freezeout level in Figure 4 is nearly a linear function of the number of data sources increasing at a rate of only 0.0011 for each additional 100 data terminals. Freezeout for both the ring and tree reaches 0.0196 ± 0.0010 with 98 voice sources and 6400 terminals. The system is stressed beyond the CCITT objective of 0.005 [10] because this objective is considered to be conservative.*

\*The 90% confidence interval half-lengths for the six cases shown in Figure 3 range from 7 ms to 30 ms. The same random number sequences were used to generate the voice traffic in all cases.

\*As argued in [3], the 0.005 objective may be conservative because current speech detectors [4] produce bursts of shorter average duration than assumed when the CCITT freezeout objective was established. Therefore more bursts are clipped but most are clipped a negligible amount (less than 20 ms) at the same 0.005 level. Since the resultant impairment in speech quality is less, the 0.005 objective is considered conservative.
Figure 4. Variation of freezeout with number of data Terminals

Also of interest are the lengths of the queues which build up in the link switches. These results are useful for buffer sizing. There are no restrictions on queue size in the simulation in order to determine the maximum queue buildup. The lengths of the largest queues in the same simulation cases as in Figure 3 are shown in Figure 5. Maximum buffers in both the ring and tree increase steadily with increasing data load in Figure 5. The trend toward a nonlinear increase seen with delay at heavy loads in Figure 3 is not seen in Figure 5. Note that maximum buffer is a single sample in a run and therefore is not a meaningful statistic but rather an observation. Its confidence interval can not be computed. However, each of the samples in Figure 5 comes from simulations of about 200,000 voice and data bursts; the maximum buffer was observed to be in steady state in these cases.

While 95% data queueing delay is a convenient figure of merit, the total cumulative distribution of delays provides a more detailed performance picture. The cumulative distributions of queueing delays on the link group for the ring and tree baseline cases of 98 voice sources and 960 data terminals are shown in Figure 6. These two cases were chosen as baseline because they just meet the 0.005 freezeout objective. Note that over 90% of the data bursts in the baseline simulation runs experience zero queueing delay. Mean data delays for the ring and tree are 5.1 ms and 4.7 ms, respectively.

The comparison above was done in order to see the effect of spreading the data queues in the ring as opposed to the tree. Because traffic leaving the tree link group merges at LS 5 and LS 6 (Figure 2), it is not surprising that congestion is concentrated at these two switches. In the ring (Figure 1) one would expect congestion concentrated in switches 1 and S but also perhaps spread out along switches leading into these two switches. Some blocking of bursts was observed in the ring at switches 2 and S-1. However, for the cases considered in this paper, this spreading out of the blocking in the ring was not enough to cause any statistically significant differences in performance between the tree and ring.

4.2 Maximum Number of Data and Voice Sources on a Link Group

Consider now what happens when we incrementally add data traffic to the T1 links in a manner that maintains acceptable speech quality and also stays within acceptable data delay limits. We use here the 0.005 CCITT freezeout objective and an arbitrarily chosen 95% link group data queueing delay limit of 100 ms. Results of varying the number of voice and data sources are shown in Figure 7 and Table 1 and are discussed below. The relation between the number of active data terminals and off-hook voice sources is virtually the same for both the ring and tree. For this reason, one may refer to Figure 7 for both topologies. This curve extends to 100 voice sources with no data sources present.

Table 1 Ring simulation voice and data performance

<table>
<thead>
<tr>
<th>No. of Off-Hook Voice</th>
<th>No. of Data Ports</th>
<th>No. of Data Terminals</th>
<th>Channel Utilization</th>
<th>Mean Data Delay (ms)</th>
<th>95% Data Delay (ms)</th>
<th>Speech Freezeout (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0.77</td>
<td>0</td>
<td>0</td>
<td>0.45</td>
</tr>
<tr>
<td>96</td>
<td>39</td>
<td>960</td>
<td>0.77</td>
<td>5</td>
<td>3.2</td>
<td>0.48</td>
</tr>
<tr>
<td>96</td>
<td>42</td>
<td>1344</td>
<td>0.78</td>
<td>8</td>
<td>2.3</td>
<td>0.41</td>
</tr>
<tr>
<td>96</td>
<td>94</td>
<td>3008</td>
<td>0.79</td>
<td>13</td>
<td>1.7</td>
<td>0.48</td>
</tr>
<tr>
<td>90</td>
<td>220</td>
<td>7040</td>
<td>0.81</td>
<td>15</td>
<td>2.2</td>
<td>0.48</td>
</tr>
<tr>
<td>80</td>
<td>400</td>
<td>12800</td>
<td>0.82</td>
<td>18</td>
<td>3.2</td>
<td>0.48</td>
</tr>
<tr>
<td>60</td>
<td>768</td>
<td>24576</td>
<td>0.85</td>
<td>13</td>
<td>3.7</td>
<td>0.49</td>
</tr>
<tr>
<td>24</td>
<td>1410</td>
<td>45120</td>
<td>0.90</td>
<td>15</td>
<td>3.7</td>
<td>0.48</td>
</tr>
<tr>
<td>0</td>
<td>1890</td>
<td>60160</td>
<td>0.95</td>
<td>21</td>
<td>3.6</td>
<td>0.48</td>
</tr>
</tbody>
</table>
From the voice and data model parameters it is known that 15 data ports (480 terminals) produce approximately the same traffic volume (number of bits) as a single voice source. So, as expected, speech freezeout remains fairly constant when the number of voice sources is decreased by 2 from 100 to 98 and 30 data ports are added to the ring. However, when the number of voice sources is decreased by 10, 220 data ports can be supported at the same level of voice performance; the ratio of data ports to replaced voice sources has increased from 15:1 to 22:1. Note in Table 1 and Figure 7 that utilization of the channels into and out of the interface link switches (ILSs) has increased from 0.77 to 0.81 and also that data delay has increased but is still within our stated limit. This result satisfies our intuition that data bursts fill in the gaps between voice bursts and increase utilization. Voice performance is not greatly impacted because the data bursts are an order of magnitude shorter than the voice bursts. Therefore voice bursts, which have nonpreemptive priority, have a relatively short wait for any channels previously seized by data bursts.

As the number of data ports is increased even further while maintaining freezeout < 0.005, channel utilization continues to increase but at the expense of increasing data delay. Therefore, as the number of voice sources decreases to 80 and below, data delay resilient speech freezeout as the limiting constraint. Channel utilization continues to increase until it reaches 95% in an all data environment. At this point the link group supports 1880 data ports, equivalent to 60160 data terminals, and the 95% data queueing delay is 94 ms. We recognize that such a scenario is somewhat unrealistic. That is, it is unlikely one would find such a large number of terminals in an area which would be reasonably served by a single link group.

4.3 Rerouting for Congestion

A capability of the tree simulation is Rerouting for Congestion (RC). RC is applied to external voice traffic on a burst by burst basis as opposed to a call by call basis. Each talkspurt issued by a talker is free to take an independent path to its destination regardless of the paths of previous bursts. All bursts will attempt to follow their primary route first, but if that route is blocked the burst will follow its alternate route through the other ILS. Only one change of direction is allowed. Internal data traffic is not rerouted to avoid passing through the hub switch and also to avoid the (usually) most congested set of channels leading into the hub switch. Data traffic of any type is not rerouted because potential sequencing problems could result if a burst is queued.

RC was not implemented in the ring simulation. Most congestion is found at the last link switch and a ring could have a maximum of 16 LSs or 32 LSs. Therefore, most rerouted bursts in the ring would end up with a lengthened path taking them through nearly all the link switches.

Results from an all voice, zero COI case in Figure 8 show a sizable decrease in freezeout when RC is used in the tree simulation. With 100 off-hook talkers the reduction is about 83%, from 0.0046 to 0.0008. Most of this blocking occurs at the four channel sets leading to or from the ILSs. This is due to the fact that all sources on the LAA are competing for those last sets of channels leading to the rest of the world. The marked reduction in freezeout as a result of the RC option makes possible an increase in the number of off-hook subscribers that an LAA could support at the same freezeout level. It is seen in Figure 8 that this number could be increased from 100 to 107 and still meet the CCITT objective of 0.005 freezeout.

To implement RC in an actual burst-switched system would be feasible, albeit nontrivial, since at least one additional bit for routing control would need to be assigned and processed. A potential problem in using RC is that of voice bursts arriving out of sequence at the destination. As long as the minimum delay is increased but the bursts is greater than the additional processing delay caused by passing through a larger number of LS's, the bursts will always be in sequence entering the hub switch. A minimum pause can be ensured by the speech detector in order to avoid a sequencing problem on the link group. It is assumed that additional processing delays will be small so the integrity of the silence intervals will be preserved.

5. CONCLUSIONS

In summary, the performance of two burst switching network topologies are evaluated in terms of their traffic carrying capabilities. This paper gives insight into the effective design of an efficient burst switched network. An integrated voice and data environment was investigated via simulation. Mixing voice and interactive data in the ring and tree simulation models produces similar speech freezeout, data delay, and maximum data queue results over a wide range of traffic loads.

It was shown that channel utilization increases significantly with an increasing percentage of interactive data traffic relative to voice traffic. Keeping speech freezeout and data queueing delay below stated thresholds, channel utilization on the ring increased from 0.77 in an all voice environment to 0.95 in an all data environment. In the region of interest around the baseline case (98 off-hook voice sources and 960 interactive data terminals), the 95th percentile of the data queueing delay increases only about 40 ms with the addition of 2000 interactive data terminals. The maximum link switch buffer size used to queue data in the baseline case simulation runs was less than 5 kbytes.

Rerouting of blocked voice bursts was implemented in the tree simulation. Speech freezeout was greatly reduced as a result (Figure 8). In an all-voice situation while maintaining a 0.005 freezeout level, the number of off-hook voice sources in the simulated tree network could be increased from 100 to 107 with rerouting for congestion.

Figure 8. Rerouting for congestion improvement

The scope of the studies described in this paper has been performance aspects of voice and data in the local access area. Performance modeling activities need to be extended into the exchange area, which includes a single hub switch, and eventually into network-wide multi-hub studies. Possible extensions to past performance evaluation studies are to consider:

1. Multirate voice and data,
2. Traffic which both originates and terminates on the link group,
3. Other forms of congestion control,
4. Bulk data traffic on dedicated channels using a variable bandwidth concept, and
5. Variable host computer locations.
6. ACKNOWLEDGMENTS

The author wishes to thank S. Kopec for simulation development efforts and C. Jack for generating simulation performance results.

7. REFERENCES