Proposal for the Development of a National Communication Service for on-line Data Processing

Summary

Starting from the assumption that on-line data processing will increase in importance, and that the users of such services will be spread out over the country, it is easily seen that data transmission by a switched network such as the telephone network is not matched to the new communication needs that will be created.

The user of an on-line service wishes to be free to push keys sporadically, and at any rate he wishes, without occupying and wasting a communication channel. But he does not expect a reply from the computation service for less than a ‘message’ of several characters, typically between 10 and 100.

A message communication service in which short messages are temporarily stored in computers situated at the nodes of the network, and forwarded in turn, can give great economies in the use of transmission paths. Further economies are afforded by the use of digital transmission plant, with regenerators in place of linear amplifiers. The result of these two factors is that transmission cost can be extremely low by present day standards.

The computers which collect and assemble messages from many terminals and distribute messages to these terminals will have taken over some of the functions normally ascribed to multi-access computers. Correspondingly, their capital cost will be high. Only a rough estimate of cost can be made short of a proper design study, but the result is encouraging.

The starting point of this proposal was communication between people and machines, but a message system could be employed for very many other purposes in which industrial or data collection equipment communicates with a central computing system.

The paper concludes with proposals for investigations needed to increase our knowledge of the profitability and benefit to the community of a message communication service of the kind proposed.

1. The Terminal Equipment

A keyboard and cheap printer today form the simplest terminal for people to use. At present, the need for data storage with luminous displays based on C.R. tubes causes them to be rather too expensive for general use, but this may not always be so. For the present, let us assume a keyboard and printer. Although it is not ideal, the teleprinter will probably be used for some time as a terminal.

The choice of the best signalling interface between the terminal and the system is bound up with problems of multiplexing to reduce the cost of local lines, and cannot be decided without more study, but some points are made on this subject in section 4.

The amount of use a terminal gets, i.e. the traffic it generates, depends on how it serves the organisation that rents the line. The biggest traffic will come from professional users, such as clerks acting as an intermediary between the public and a public service (such as transport, power utilities, banking or insurance).

A typical request, such as an enquiry for the booking of airline seats, will occupy no more than a line of typing and the reply will be about the same size. Such messages would be exchanged only about once a minute for a given terminal.

Most other users would be more casual in their use of a terminal. Some of these might share a terminal, for example in a design office for reference to information services, design calculation services and so forth. The average use of a terminal would be much less than one message a minute.

In a large city, the density of terminals and their usage would be high, and an exchange for 10,000 lines would handle about 200 messages per second. Out of 100 such exchanges, if half the traffic were local, would come 10,000 messages per second (peak rate) and a given route would be called upon to handle 1,000 messages per second.

These very rough estimates are made in order to give numbers to assess the transmission and message handling problems. If terminals with displays become cheaper, and can be widely used, this might encourage the more lavish use of the communication system, and this, more than any other factor, could increase the traffic beyond the figures mentioned. For the case of printers only, our figures are probably overestimates.

2. Digital Transmission

The best evidence available comes from the existence and economy of Pulse Code Modulated speech channels. They are economical in spite of their high bandwidth compared with the speech they carry because regenerators for digital data can completely restore it after some degree of distortion. It is clear that the carrying of digital data on digital channels rather than on speech channels will reduce the transmission cost greatly. Hartley and Thomas in the IEE Colloquium of October 1965 estimated a factor of 25 to 1.

A corresponding result is that data transmission, done properly, will occupy only a small amount of transmission capacity compared with speech. The bit rate of 1.4×10^6 used in P.C.M. junctions will carry several times the amount of message data we estimate as originating in a very large town. All the data could be carried in the pauses in speech if several P.C.M. channels were being used.

If P.C.M. speech transmission becomes widespread, only that portion of the channels needed for the actual message data load need be diverted to that purpose, and the message channels will be elastic.

Because speech will always occupy much more channel capacity than message data, the economics of transmitting speech will decide the form of the country’s communication network. The use of P.C.M. seems economical for short and medium distances, and there are arguments which tend to show that its range of economical use will grow.

The cost of transmission will therefore be ignored.

3. Tandem Message Exchanges

A node in the network which is not connected to terminals presents a much simpler computer problem than an exchange serving terminals, therefore it will be discussed first. Such an exchange can also be connected to several computer services. These services communicate at high speed in complete messages so they appear at the exchange much like the transmission paths of the network. They do, however, need a small extra amount of processing to add to each message a note of its source, and possibly to replace the stated destination by a route indication.

The task of such a tandem exchange is basically simple. Very simple ADT devices could place incoming messages into small buffers in a core store. A central computer could examine these buffers in turn and sort the messages, according to routeing information, into output queues. The output queues would fluctuate in length somewhat, and waiting here would be the main source of delay. Simple ADT devices could transmit the messages at the rate allowed by the channels, whenever the queue was non-zero.

An approximate model (valid for the case of many channels) is that arrivals in the output queue are Poisson distributed and output is at a regular rate determined by the time *t* taken to transmit one message. We expect *t* to be about one millisecond. The mean time of waiting in the queue would then be

$$W = \frac{t}{2} \frac{\rho}{1-\rho}$$

where ρ is the mean arrival rate divided by the steady (potential) departure rate, i.e. the degree of saturation. Since saturation would be disastrous to the service and traffic cannot be accurately assessed it is not likely that ρ would be greater than 0.8 and therefore the mean waiting time only 2*t*.

This approximate model illustrates that appreciable queuing delays indicate near-disaster. If the arrival is less random than Poisson this is even more true.

Assuming that to carry a message no more than 5 tandem exchanges are ever needed, and therefore that such messages are held about 7 times in short buffers and 5 times in output queues and are transmitted six times, the total delay time would average about 23*t*. This could be kept down to 100 milliseconds if all the communication channels had a capacity of at least 250 messages per second. With digital transmission this sort of capacity would easily be provided, and correspond to a few telephone (PCM) channels.

The rate of access to the store might however be a limitation, or additional core stacks a cost burden. But if a message fits into 10 store words, and each message is handled three times in transit through a tandem exchange, a rate of 10,000 messages for the whole exchange would need only a single 3 microsecond store. Each queue would need two pointers which might also be in the core stores, but references to these in the main store might run a single store into access limitations for a 10,000 per second exchange.

For a rough estimate, the computer needed would cost about £100,000 for the largest of such exchanges, allowing for considerable redundancy.

4. Message Exchanges with Terminals

Since local lines will add considerably to the capital cost if they exceed a mile, and only in commercial centres will there be enough density of terminals to connect directly to a large exchange, arrangements for multiplexing lines in small numbers are essential.

Time division multiplexing of slow serial signals, like teleprinter signals, involves many samples per digit. Considering the signals from the keyboards to the exchange, this creates quite a problem of unscrambling the digits and assembling them into characters.

Assume that each terminal must be sampled 500 times per second for reliability and to avoid jitter. A group of 500 such lines can be multiplexed in a number of stages to make one synchronous system, giving a bit rate of 250,000 to be dealt with. The interpretation of the samples from each channel can be expressed as fixed logic in terms of the incoming bits and a certain status word. A core store can hold the status words and partially assembled characters, one per terminal.

The recognition logic could be built-in, since it would have to operate on a new channel and status word every 4 microseconds.

The distribution of characters to printers at the terminals could share the same store but have separate logic.

The communication of complete characters to and from the main store, with their channel numbers, would not occupy many store cycles.

It is difficult to see how such a character assembler and distributor could cost less than £100 per terminal, and it is probably the most expensive equipment, next to the terminal equipment itself.

A rough calculation of capital cost per line can be made as follows:

	£	£
Terminal equipment	400	(5)
Local line	25	(50)
Part of Multiplexer	25	(–)
Character Assembler/distributor	100	(–)
Main Computer or Exchange equipment	25	(35)
Tandem Exchanges	25	(10)
Lines carrying dense traffic	5	(50)
Buildings	35	(35)
	640	(185)

The figures in brackets give for comparison some figures for the telephone system, probably not accurate, but indicating the main differences. The biggest part of the capital cost for message communication is in the terminal equipment and character assembler/distributor.

These rough figures show that the capital cost is acceptable if the system makes a reasonable increase in the productivity of one operator. They are only rough figures, and some more detailed design would be necessary to refine them.

The results indicate the importance of better design of the signalling interface with the terminal. Suppose, for example, that the terminal device was able to store a character and serialise it rapidly (10–20 microseconds) when a pulse was received from the exchange. This arrangement would deliver complete characters to the exchange, and might well cost less than the £100 per terminal estimated for the same function at the exchange.

5. External Characteristics of the Proposed Message Communication System

Many external characteristics have already been described and further details can be added now that a rough idea of implementation has been stated.

Firstly it deals in characters, presumably of 8 bits in I.S.O. code. This is essential if it is to communicate reasonably cheaply with terminals designed for people to use. The so called ‘transparent’ interface allowing binary data without restriction would mean losing one information bit to avoid hitting a control character.

Secondly it does not transmit until it has collected a ‘message’. The size of message is an economic problem, decided by reference to the red tape characters that the system must add to each message, probably about 20 in number, to indicate origin, route and a few other things.

It is suggested that a message should be sent when the number of characters reaches a certain maximum, something like 50, or when carriage-return-line-feed (newline) is sent. Thus, typical conversational interactions would be segmented naturally by the users, whereas long messages would be arbitrarily segmented by the system. However, subscriber’s equipment for bulk transmission might recognise the known message maximum and segment its error detection blocks to suit it.

The reason for a maximum is to ease the housekeeping problems of the exchanges, and to share the time of channels equitably so as to keep delays to a minimum.

After studying the housekeeping problems it might be decided that a longer message maximum with a smaller unit of transmission and storage could be provided.

Thirdly, the delay in transmission due to intermediate storage could be kept down to about 100 milliseconds. Since the system is designed for a person at one end of the link, this is probably better than is strictly needed, but a short time of transmission will increase the range of uses to which the system can be put. It will, for example, help to speed up error checking in messages sent at high rates.

Fourthly, there will be high-rate connections to the system, suitable for the attachment of computers. These will handle complete messages. The communication system will therefore act as line concentrator, character assembler and message assembler for all the computers attached to it. This is obviously better than present arrangements in which the telephone network does the line concentration and transmission and each computer does the rest for itself.

6. Commands to the Communication System

It must be possible for the users, both computers and people, to tell the communication system what service they require by means of coding which is distinct from the data to be transmitted. The means of conveying this information and the services the system should offer will only be settled when the first public message systems go into operation. Fortunately, by using the versatility of the stored-program computer these details can be modified during the experimental period. Only when several exchanges are connected in a system with many dependent computer services will the need for fixed operating standards arise.

By analogy with some multi-access computers the instructions to the system will be called ‘commands’ and they will be represented by a name associated with arguments. Thus a simple requirement to send a message *Y* to terminal number *X* might be expressed:

to (*X*, *Y*)

In this notation, underlines are used to distinguish commands, and they are double to enable the data *Y* to contain underlines, which might arise in programming languages, or in the commands to a computer at the far end. If a computer has to be instructed to use double underlines in a reply message, some convention can be used so they are not sent like this in a message. It would be possible to make all these distinctions by the correct use of brackets, but this is unsafe, and not obvious to the user. The use of (, and) as punctuation in the above command is only one convention, derived from mathematical notation, and is not necessarily the best or the most natural to the user. An alternative which is more natural to the user but less so to the computer is:

(*Y* to *X*)

Some of the facilities that a message system might offer will now be described. A lot of thought will be needed to decide exactly what is needed in practice and can be economically provided.

Services attached to the message system that are needed by most users will have nation-wide names by which to address them. This will apply, for example, to computer services of a universal character.

A business user may wish to fix some terminals to a remote service and make it impossible for the operators to detach them. Other users may wish to fix their terminals voluntarily, and not have to give a command with each message. Terminals may want to exclude other incoming messages while they are connected to a certain computer service, or divert these messages to another terminal. A lock which prevents a keyboard sending the double underline would prevent tampering with a 'connection'. These 'connections' of course would be no more than entries in the appropriate tables in the main computer at the local exchange.

It would be possible to arrange that, by a suitable command, a user could assemble a long message (such as business letter) and not have it delivered until he has checked and edited it.

Messages could be recorded and timed on request, so that the time of sending was legally established and the fact of sending also. Facilities can be invented for ensuring that a message has arrived and arrived correctly, and for protecting it from arriving intelligibly at the wrong place even if there was a malfunction.

A user could deposit via his own terminal a certain 'key' such as a string of digits or a phrase, and use this secret message as authorisation to pick up messages from another terminal where he happens to be, or use it to get access to a service and charge the cost to his account.

The system facilities described are designed for the convenience of people using keyboard/printer terminals. Computer services could carry out such functions largely for themselves—such things as sorting incoming messages according to their origin, for example. Computer services would offer a great variety of convenient facilities to their users, while the communication system would only offer services with a very general application.

7. Some uses for a Message Communication Network

The original intention for its use, the connection of terminals to computer services, remains of primary importance. A selection of such services is listed:

- Numerical computation at various levels of generality
- Editing and typesetting of text
- Design services and problem oriented languages
- Availability of goods for sale
- Ordering of goods
- Invoicing, delivery notes, etc.
- Booking of transport
- Banking, establishing credit
- Remote access to national records, e.g. MPNI, tax, police, medical, on a secure basis
- Betting

The use of the system for people-to-people communication represents an elaboration of the Telex system. The greater distribution of terminals because of all the other services offered will make this method of communication more useful.

The use of the system for machine-to-machine communication depends on its reliability, i.e. freedom from congestion and visible faults. It would be possible for vital data to be sent to the two nearest exchanges by separate conduits, for protection against pick and shovel. Examples of machine uses, many of them having very low data rates per station, are:

- Road traffic control
- Monitoring and controlling of widespread plant, like pipelines, utility services or automatic Met. stations
- Burglar alarms and other security devices connected to monitoring computers and giving regular reports when asked
- The control of the telephone switching system

It is possible that the need to use very simple reporting devices for some of these applications will require certain services from the message communication system. For example to use a remote sensing device or motor attached to a pipeline one would have to be sure that it received a message only from its controlling computer. Thc local exchange could do this more cheaply than special hardware.

The most familiar form of data transmission today is bulk transmission at constant rate using a solid transmission channel. Bearing in mind that few organisations can use a rented line efficiently, there may often be a cost advantage in sending bulk data in message form in spite of the extra handling involved. Any slight variability in data rate at the received end would not matter if there was a small buffer. This variability should not be present at night, when a lower cost per message might be offered to encourage bulk traffic. Alternatively, a guaranteed channel capacity with fixed delay through the system could be offered by the use of special programs in the exchange concerned, and which method to adopt is an economic question.

The need for bulk data transmission should decrease as more on-line services come into use and data processing systems become rationalised.

The Development of a Message Communication System

Estimation of the demand for message communication of the kind proposed will be very difficult. The best that could be done is to approach computer users who have known real-time needs, such as airlines, pipeline operators or stockbrokers and describe the possibilities in some detail. They could then be asked whether and how much they would use such a system at various prices.

The improvement of our knowledge about the design of the system and the probable cost could be done in stages, and be started now. Companies with a stake in on-line computers and communications could do the work under contract.

A pilot service could be set up in central London, where it would attract the most custom, in order to learn about the problems and determine what operational features the customers would pay for. The first service would lose money for some time, and would need to be replaced when national and international standards were developed.

Such an experiment at an early stage is needed to develop the knowledge of these systems in the GPO and the British computer and communications industry. At present, one US company, Collins Radio, has specialised in providing message systems for private companies. It is very important not to find ourselves forced to buy computers and software for these systems from USA. We could, by starting early enough, develop export markets.

Research on the terminals and their signalling interface with the system will be essential if the cheapest overall solution is to be obtained. This should be carried on in parallel with the other developments and not left too late to influence the eventual standards.

D. W. DAVIES
15th December, 1965
AUTO 68/013