Mars 1974
SCH 527
INWG #60

A PROPOSAL FOR INTERCONNECTING PACKET SWITCHING NETWORKS

by

Louis POUZIN

Directeur du Réseau Cyclades
Institut de Recherche d'Informatique et d'Automatique (IRIA)
78150 Rocquencourt, France

Abstract.

Interconnecting packet switching networks may be a puzzle, when they are built as stand alone objects, and mix various other functions with the transport of packets. In working out international standards, agreement will likely be reached over a minimum set of basic characteristics, such as : - adressing plan, - packet size, - header format, - accounting. And packet transport will be the basic service. This allows each network to keep complete autonomy in internal design, technology, and operation. With a hierarchical address space, packets can be shipped across networks at a minimum overhead. Options in header format cater for shorter or longer subscriber addresses, depending on network size. The maximum length of information carried within a packet would be 255 octets. With a limited set of constraints, an aggregate of networks behave as a single logical network, called CATENET. Interfacing with existing networks usually requires a front end adaptor, or gateway, to wrap CATENET packets into local formats, unless the standard format is accepted internally. Other specific services, such as end-to-end control, are to be stopped at the gateways so that they do not interfere with CATENET workings.

To be presented at Eurocomp, Brunel University, 13-17 May 1974.

1. INTRODUCTION.

Packet switching networks are presently being built without apparent concern for interconnection. Thought out as stand alone objects, they are not even fit for connection with a twin-network, which is a tell-tale of egocentric design. Similarities between networks are misleading, since they hide differences in the details, and a sophisticated logic seems to be required in order to interface them. Actually no experiment has yet been performed to assess this very possibility.

Most of the apparent difficulties stem from peculiar functions which are not mandatory features of a packet switching service. They smack more of accidental design, rather than intrisic requirement. In this category fall virtual circuits, end-to-end acknowledgements, buffer allocation, and the like. As a close examination reveals, packet switching would not be the least impaired, should these functions be left out. On the other hand, any end-to-end function, which users might desire, can be implemented over one or several networks at the user interface.

Assuming that a network performs packet switching stricto sensu, interfacing with a kin turns out to be a trivial matter, as long as a minimal agreement has been reached for packet formats. This paper points out the major facets of a solution to interconnection. More details can be found in ref.[1].

2. TERMINOLOGY.

- Packet Switching Network (PSN). A set of communications computers (nodes) connected via physical circuits, and carrying bit strings using a store and forward technique.
- Packet. A bit string exchanged between nodes of a particular PSN.
- Message. A bit string exchanged between a PSN and an external device, (computer or other PSN).
- Gateway. A particular node capable of sending/receiving messages.
- Catenet. An abstract PSN resulting from the juxtaposition of several PSN's.

This terminology is used in the present paper. Readers should be aware that different authors may use different words, or mean different things with the same words.

3. TRANS-NETWORK SERVICES.

These are the services expected across several interconnected networks (Catenet). The minimum basic service is the capability to send a message to a destination, given a global address understandable by the traversed networks. Message contents should be delivered unaltered as an exact copy of the sender's message.

Some control services may be desirable, such as echoing, tracing, timing, accounting. Due to the independent design of each network, some variations are likely to occur, and it is more difficult to define standard services, although some proposals are made in ref.[1].

Other services such as end-to-end control are not expected from Catenet. They may be implemented as added services on a mutual agreement basis.

E.g. some existing or planned networks offer services such as virtual circuits, end-to-end flow control, short addressing, hot circuits, broadcasting, mailboxes, multiple addresses, etc... It is not our intent to discuss the appropriateness of any particular property. But since they are likely to be just a little different from one network to another, at best, they should be considered as local features to be ignored when defining a general trans-network service. Of course, the fact of not using them must not require any peculiar set up, from the external user standpoint. Practically speaking, it is more realistic to seek a multi-network agreement based on a minimum service, rather than a maximum.

4. ON GATEWAYS.

When it comes to link networks, one approach, which is often advocated, consists in shipping messages through gateways in charge of interfacing between networks.

Gateways would look like a host for each network, i.e. receive and forward messages, exchange adequate signaling information, perform some error and flow control, accounting, etc. In other words, a gateway could be viewed as two hosts face to face, (fig. 1).

One might think of gateways as nodes of a super-network, in which transmission lines happen to be local networks. Then, one could implement within gateways whatever properties should be deemed fit for inter-network communications, using more or less of the local features to carry messages between gateways.

Actually, this last approach boils down to the construction of yet another network, and ultimate at that. The very feasibility of such an objective can be seriously in doubt. Should network inter-connection depend on a universal network, schisms would indefinitely raise the problem one level up.

This is not to say that a general inter-network agreement will never happen, but that it will happen gradually, starting with a few recommendations on fundamental issues, rather than attempting to design a universal network whereas hardly any has reached out an experimental stage.

A simpler approach is to consider gateways as mere gatekeepers, in charge of looking after individual messages, to make sure that they are acceptable. It bears some similarity with the policeman at the border. Each PSN is perfectly entitled to protect itself from undesirable traffic, and to check message formats as against local regulations. Of course, a set of internationally or widely agreed tolerancies should be offered by each PSN to entering traffic. Or else, it would belong to those peculiar networks, which messages had better keep away from.

What functions should be implemented in gateways in actually local PSN business. The only requirement is that opposite gateways interfacing two networks should agree on a common transmission procedure. But this is merely a border problem, not a world-wide issue.

5. CATENET STRUCTURE.

Catenet is an aggregate of independent PSN's. Each one retains complete autonomy in internal design, control mechanisms, and operation. The only constraint is to be able to carry messages through, using a standard message format. On the other hand, within their own boundaries, PSN's may carry local traffic using whatever local format they may desire. PSN's may be physically connected through direct circuits between nodes. They only need agree on a local transmission procedure for each circuit. Any pair of PSN's may be connected through several circuits, (fig. 2).

While Catenet should be ideally a mere piecing together of independent PSN's, it should nonetheless display a consistent behavior equivalent to a unique simple PSN. This criterion will be exercised to assess the validity of its functional structure.

On the other hand, no assumption is made about the particular techniques used in each PSN to carry messages through. They may be fragmented in several packets, and reassembled in an out-gateway. Or several may be bundled into a single packet, and unbundled later on. Virtual circuits may be set up, if this appears desirable, etc... As seen as Catenet components, PSN's are only supposed to carry independent messages, no matter how they do it.

6. NAME SPACE.

6.1. Network hierarchy :

When a message travels within a PSN toward another, there is no need to interpret the destination address any more than required to find an appropriate gateway in the correct direction. Putting gateway names in addresses is unacceptable, as it would tie up addressing, and network topology. Thus, only PSN names should be used as Catenet addresses. Delivering a message to a final destination is carried out only by the final PSN.

So as to allow each PSN to segregate easily Catenet and local address formats, a complete Catenet address is structured as follows :

<Format> <PSN name> <local name>

The first two components may be of fixed length, since the total number of PSN's is roughly predictable. The local name is to be of variable length, in order to accomodate local requirements.

6.2. Area hierarchy :

Depending on local PSN authorities, PSN addresses may well be structured with an area name and a local subscriber name. This stems directly from designing a PSN along the same lines as Catenet, i.e. as an aggregate of local area networks.

If it happens that local traffic makes up a significant fraction of the total traffic, then shorter addresses may reduce message overhead, and simplify the design of equipment intended only for local traffic.

To the extent that adjacent or overlapping PSN's comply with a homogeneous format, an area may well belong to different PSN's, or straddle over boundaries. This is most convenient for subscribers which can be known by the same name within different PSN's, or can be reached with short addresses from an adjacent PSN, as if they were in the same net. But all that is just borrowing from the telephone system, which is an interesting and sophisticated model.

6.3. Private hierarchy :

Carrying this scheme one step further, a PSN subscriber may be a private network, using the remaining part of the local address to route a message toward a final sub-subscriber.

6.4. Address format :

Some specific sizes must be proposed for the fields of an address, in order to visualize what this scheme might turn out.

Since addresses may have to be composed from keyboards or panels, and are likely to be displayed on diagnostic equipment, it appears convenient to use digits, decimal or possibly hexa. Of course a conversion logic might well translate internal formats into human readable ones, and vice versa. But if it happens not to waste any significant bandwidth, it is probably simpler to keep internal formats as readable as possible.

Fields could be as following :

- Format (1 digit)

  = 0 local PSN address
  = 1 1-digit global name
  = 2 2-digit " "
  = 3 3-digit " "
  = 4-F reserved for future use

- Global name (1-3 digits)

  = 1-9 very large common carrier or PTT network, (continental)
  = A-F " " well known network
  = 10-99 common carrier or PTT national network
  = 9A-FF well known international network
  = 100-FFF reserved for future use

- Local name (1-10 digits)

  No particular standard is imposed, as it relates to local PSN characteristics. However, it would probably be convenient to allocate names along the preceding lines, see e.g. ref.[2].

  To sum up, the address size would be 3 to 14 digits, depending mainly on local name format. Except for the destination PSN, other networks would only be concerned with format and global name, i.e. 2 to 4 digits. This should be quite economical.

6.5. Remarks :

This question of numbering plan has been analyzed in the context of packet switching. Actually it applies verbatim to any type of data transmission network. In the case of a (virtual or physical) circuit-minded network, the setting up of a call implies the routing of one or several initial messages, based on some generally understood addressing plan. Only after this initial operation can messages be routed in reference to an established circuit, instead of a full destination address. Therefore, the issue of an international addressing plan could be studied quite independently of transmission techniques, and interface problems.

7. PACKET IDENTIFICATION.

In order to insure message transfer reliability some naming scheme is required to identify individually each message between source and destination. This is the only practical way to pinpoint lost or duplicate messages,[3] . No matter how reliable each network authority will claim to be, it should be expected that end users will rely primarily on their own error control machinery when using Catenet services. Consequently they will have to reserve a name field within each message. This is typically what is called a sequence number in conventional transmission procedures.

Furthermore, occasionally there will be some arguments between networks and users about weird traffic behavior. Trouble tracing requires some message identifier. It does not have to be unique network-wide as long as there is a sufficiently low probability to find the same in a different message. If they occur, ambiguities are resolved by looking at other fields. In an other domain this would be called an audit trail.

Putting the identifier in the text field would be undesirable, as PSN's should look at text as a sealed envelop, possibly encrypted. Furthermore, everyone's need being about the same, a fixed length field should fit. 12 to 16 bits are a reasonable figure.

8. NETWORK FACILITIES.

Although it is somewhat early to propose standards for network facilities, it is clear that every network needs various options accessible from possibly privileged subscribers, for maintenance, statistics, trouble-shooting, etc... Thus, a field should be reserved for each PSN convenience.

Accounting is a far from settled issue. Since the service offered may vary with priority, routing, traffic schedule, to name but a few factors, it is to be expected that tariffs will not be based only on a flat rate per packet. Therefore some field should be assigned to control various parameters yet to be determined.

More control information could be added for custom tailored services, e.g. end-to-end error and flow control, message sequencing, character translation, delivery confirmation, etc... It is far from clear that user requirements can be determined a priori. Consequently, a conservative approach is to leave that question open. A few spare bits, and the possibility to define additional formats in the future are likely enough at the present time.

More detailed discussions on network facilities can be found in ref.[1] .

9. MESSAGE FORMAT.

9.1. Text :

Since fragmentation and reassembly of data items are to be necessary functions of subscriber protocols, no further fragmentation appears to be desirable at PSN level. However each PSN is entirely free to do so, if it wishes, as long as it delivers unaltered messages to the next PSN. Consequently, a substantial saving in overhead would result from a standardized maximum text size, fitting within a single packet. A present trend is 255 octets,[4,5,6,7].

This message size is slightly longer than existing packet sizes, so that no further fragmentation would be necessary. Furthermore it is reasonably efficient in terms of bandwidth and transit time. Anyhow, nothing prevents subscribers to dispatch shorter messages if they really want shorter transit delays.

On the other hand, it is contended that restricting all messages to be shorter than 256 octets would not allow using larger sizes such as Arpanet [8] (about 1000 octets). This is actually not a PSN interconnection issue, since the Arpanet packet is only 1000 bit long ; it is rather a host interface issue. As far as packet switching is concerned, PSN's using longer messages should only break them into Catenet messages, when handling outbound traffic. No problem would arise with inbound traffic.

As it should be obvious now, the problem does not lie in carrying messages across PSN's, but in host-to-host protocols, since there is no longer a one-to-one correspondence between messages sent and received. But this would be the subject of another paper.

9.2. Header :

Since hosts will have different word length, it is not superfluous to reduce formatting overhead in trying to match somehow header and word length. In that respect, some interesting header lengths are :

. 72 = 9	octets,	multiple of	6	8	12		18	24		36	
. 96 = 12	"	"	6	8	12	16		24	32		
. 120 = 15	"	"	6	8	12			24			60
. 128 = 16	"	"		8		16			32		
. 144 = 18	"	"	6	8	12	16	18	24		36	

Therefore, recommended sizes could be 9, 12, 15, 18 octets. But this is just for host convenience, and has no effect on Catenet workings.

A suggested header format is on fig. 3, it is based on the following lines :

- . fixed size fields are at the beginning, to reduce parsing overhead.
- . fields fit in 8- or 16-bit word boundaries, since switching nodes are likely to be that kind of computers.

- format information is located at the beginning in order to direct parsing more easily in case of multiple formats.
- control information is gathered in one field, in order to facilitate a redefinition of its components, if at all necessary.
- global addresses are put together before local addresses, because the former are Catenet standards, while the latter concern only end networks. On this way global addresses may be identified easily without knowing about local naming schemes. Furthermore, by mutual agreement between correspondents, local addresses may extend into the text field without requiring any prior Catenet arrangement. This is particularly useful for network hosts, or simply multiple subscribers hosts, which need to append some intrinsic addresses to locate the final recipient of a message.
- destination precedes source, as it will likely be processed more often in the switching machinery.

	Bits
Reserved for local PSN	2
Header format	2
Header length (-8 octets)	4
Text length (octets)	8
Message identifier	16
Local PSN control	3
Echo	1
Trace	1
Other functions	11
Global destination format	4
" " address	4 to 12
" source format	4
" " address	4 to 12
Local destination address	4 to 40
" source address	4 to 40

<u>Fig. 3</u> - Header format

The header length varies from a minimum 64 bits up to a maximum 160 bits. It actually contains 4 major fields : -formats, -control, -Catenet addresses, -PSN addresses.

The format proposed here might seem bulky. But it is not so much after all. With typical 12-bit PSN names (2 digits + format) the first 9 octets are all that is necessary to route a message across Catenet. This also happens to be a recommended size. The rest of the header is actually devoted to the user (host, subscriber). It can contain port or process names which must be transmitted anyway. With the present scheme users can take advantage of the header format flexibility to make text length almost fixed, rather than chop it down for internal addressing needs. This would help compatibility between network applications and services.

10. NETWORK CONTROL.

10.1. Signaling :

It may happen that some specific information will have to be exchanged between PSN's, related or not to a particular subscriber message. Some of it may require Catenet standards, or just mutual agreement between directly connected PSN's. Not enough experience is available to bring up any meaningful proposal.

Nevertheless, steps should be taken so that signaling be introduced gradually, if necessary, with experience being gained. A way to do just that is to consider signaling messages as regular traffic sent to special destinations within PSN's, instead of outside.

In other words, part of the address space in each PSN would be reserved for internal components involved in signaling processing. This technique systematically used in Cigale [5] is highly flexible, since it does not require anticipating all possible needs. Furthermore, the address space is large enough so that a sufficient chunk be reserved for internal use without putting any restriction on subscriber services. Finally, no additional overhead is incurred in switching signaling information.

10.2. Diagnostic messages :

Customarily PSN's send some error messages back to the source whenever they recognize conditions precluding the correct forwarding of a regular message.

Even in the case of a single PSN, it is not clear whether this practice is actually useful. This is even more debatable in Catenet, as it could raise a tariff question about whom should be charged for diagnostics traffic. Furthermore, some preoautions are necessary so that successive failures do not wind up in priming a message explosion. Finally, some standardization would be mandatory if they are to be any useful for end recipients.

Consequently, it is suggested that diagnostics messages be confined within each PSN. Their propagation through Catenet should not be allowed until recognized necessary by experience.

10.3. Routing :

Various routing schemes[10] have been and will be experimented. Routing information exchanged between PSN nodes is both sensitive and tailored to a specific network. There is presently no obvious necessity for inter-connected PSN's to cooperate on routing, although it may appear theoretically attractive. As long as each PSN can find a good path through itself, one may conjecture that the total path is an acceptable solution. This approach lacks a rigourous basis, it resorts more to heuristics.

10.4. Congestion :

As for routing, it is not clear how networks could cooperate, while preserving their autonomy of management and internal structure. The most realistic approach is to leave each PSN responsible for fighting its own congestion, as it should do in any case. It has been shown [9] that congestion can be under control without requiring any specific USER-PSN interface. Owing to transmission procedures, no message can enter a PSN without being explicitly allowed. Thus, it is a matter of local PSN resource management to allow in traffic only to the extent that it does not deteriorate throughput. Although this is still a research problem, simple mechanisms can be used to prevent congestion. They could certainly be more efficient. With more experience PSN's will accept higher traffic load without getting congested, and this will be also reflected at Catenet level.

11. EXISTING NETWORKS.

Every standardization process is caught in a dilemma. Either no implementation exists, and there is no experience to build on. Or there are already instances of real objects, and their owners do not want to lose their investment. When it is not too early, then it is too late. Thus, a major concern is how to adapt existing PSN's to meet proposed standards at minimum cost.

Here appears the special role which can be devoted to gateways. Instead of being just gatekeepers, they can act as front end adaptors, in order to make a particular PSN look as if it would go by Catenet standards. E.g. wrapping Catenet messages into an internal packet frame, or mapping Catenet headers into internal headers, or generating and stamping out internal PSN signaling such as RFNM's and similar species, or managing specific internal resources like slots, links, packet numbers, buffers, etc... Local fragmentation and reassembly handled across a particular PSN are perfectly acceptable, as well as any other feature, insofar as they do not leak out into Catenet.

Physically, gateways need not be separate pieces of hardware. They can be implemented as additional modules within regular PSN nodes, in charge of "foreign affairs" with the Catenet world. Experiments under way between CYCLADES [11] and the NPL network should provide an excellent opportunity to assess the cost-effectiveness of the Catenet approach.

12. CONCLUSIONS.

This paper attempts to put forward a realistic scheme allowing point to point message transfer across several independent packet switching networks. As can be expected, this requires a common agreement in formatting messages, but constraints are minimized while preserving implementation freedom and efficiency.

The proposed message format should make unnecessary an internal reformatting or wrapping around, without resorting to exotic features, so that internal and trans-network traffic be handled by the same simple switching machinery.

13. REFERENCES.

1. POUZIN L. - Interconnection of packet switching networks. IFIP/TC 6.1/42, (Oct.73), 19 p. NIC 20792.

2. POUZIN L. - Address format in Cigale, IFIP/TC 6.1/20, (Jan.73), 5 p. NIC 14497.

3. POUZIN L. - Network protocols. NATO Advanced Study Institute on computer communications networks, Brighton, (Sep.73),25 p. IFIP/TC 6.1/50. NIC 22028.

4. BARBER D.L.A. - Hierarchical control of a communications channel. Nat. Phys. Lab. (May 72), 10 p.

5. POUZIN L. - CIGALE, the packet switching machine of the Cyclades computer network. IFIP/TC 6.1/51. NIC 22029, (Nov.73), 19 p. To be presented at IFIP Congress 1974.

6. Experimental Packet Switched Service. British Post-Office Telecommunications, (Nov.72), 50 p.

7. Spain's packet switching service. Computing, (29 Nov.73), 12-13.

8. CERF V./KAHN R. - Towards protocols for internetwork communication. IFIP/TC 6.1/39. NIC 18764, (Sep.73), 33 p.

9. PRICE W.L. - Simulation of packet switching networks controlled on isarithmic principles. 3rd Data Comm. Symp., Tampa, (Nov.73), 44-49.

10. FULTZ G.L. - Adaptive routing techniques for message switching computer-communication networks. UCLA-ENG-7252, (Jul.72), 435 p.

11. POUZIN L. - Presentation and major design aspects of the CYCLADES computer network, 3rd Data Comm. Symp., Tampa, (Nov.73), 80-87.

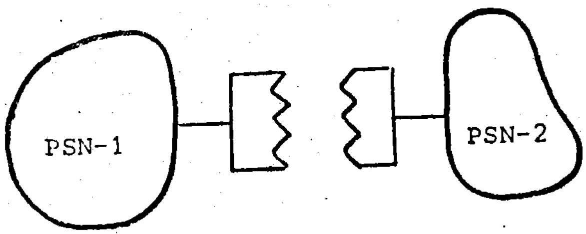

Fig. 1 - Gateways

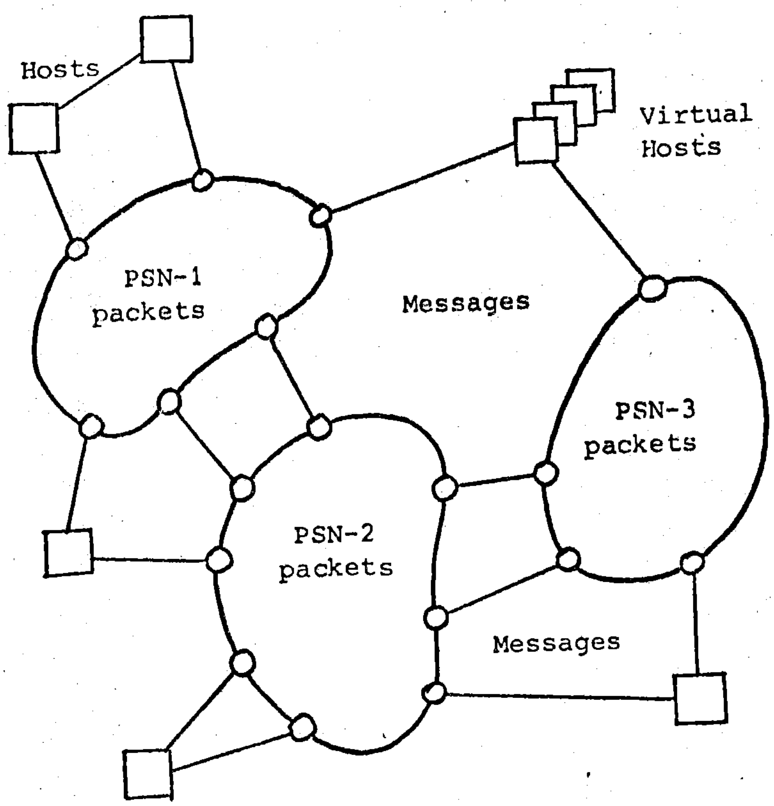

Fig. 2 - CATENET structure