ROUTING ALGORITHM AND ROUTE OPTIMIZATION ON SITA DATA TRANSPORT NETWORK

J. M. Kaliszewski

Reprinted from IEEE Global Telecommunications Conference,
GLOBECOM '83, November 28-December 1, 1983, San Diego, CA

ROUTING ALGORITHM AND ROUTE OPTIMIZATION ON SITA DATA TRANSPORT NETWORK

J.M. KALISZEWSKI

ABSTRACT

After a short presentation of the SITA Advanced Network Programme, the routing algorithm for the packet-switching Data Transport Network is presented.

This routing is of the predictive type, i.e. alternate routes are off-line computed and stored in tables. This approach leads to some constraints which are analyzed. A solution to the routing table generation is presented and has been implemented in a set of pegrams operationally used for the SITA Network.

It appears that the solution to the routing problem may be seen as a way of computing and storing trees in a graph, this solution may easily be extended to generate all the trees.

SITA COMPUTER COMMUNICATIONS NETWORK has been operating since 1969 and been presented several times in the past. (e.g. Mischa SCHWARTZ - Computer Communication Network Design and Analysis. Prentice Hall 1977)

However, several years ago, SITA defined the New Telecommunications Network Architecture, that can serve the future needs of SITA Users in the most cost effective way.

Two different traffic types are transmitted over the SITA Network :

- Type A Traffic

  This is a conversational query-response traffic, between CRTs and Airline Computer Systems (ACS). For this, traffic transmission delay is the most important objective.

- Type B Traffic

  Telegraphic one way traffic, exchanged between TTYs (and eventually ACSs).

  For Type B traffic, the most important objective is a very low loss probability.

Both traffic types are transmitted over the network with the packet switching technique, but, type B traffic undergoes additional processing for message switching and traffic protection.

1. SITA ADVANCED NETWORK

The term Advanced Network refers to the new network architecture the implementation of which was started at the end of 1981.

The Advanced Network comprises four specific systems listed herebelow :

. The Data Transport Network (DTN)

. The Network Control System (NCS)

. The Message Storage and Handling System (MSS)

. The User Interface System (UIS)

Figure I shows a functional layout of this network, it also exhibits an HLS, which is a system used as a switching node in the former SITA High Level Network.

In the framework of the Advanced Network, the the HLS is relieved from a Type A and transit traffic treatment and is devoted to type B traffic (input/output and message switching).

1.1. The SPS is the currently used access system performing traffic concentration and protocol adaptation between user protocols and SITA internal protocols.

SPS can handle both type A and type B traffic only.

1.2. Data Transport Network (DTN)

The Data Transport Network is a packet switching network offering a datagram like service for data transmission.

The DTN nodes are called DIS and they perform all processing for type A and transit traffic.

From an ISO Model point of view, the DTN will offer all transport functions (i.e. level 4) necessary for existing and new

83CH1956-2/83/0000-0892 $1.00 © 1983 IEEE

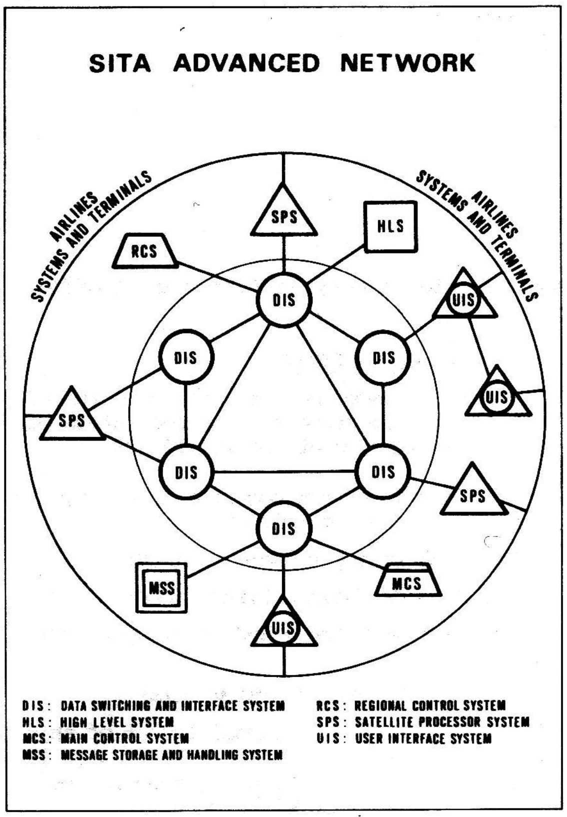

FIGURE 1

traffic types. However, the introduction of the DTN also results in immediate and significant improvements in the following areas :

- Throughput Capacity

The DTN provides additional throughput capacity for type A and transit type B traffic. DIS-DIS links are either medium speed, with a capacity of 9.6 Kbps or 14.4 Kbps, or wide-band with the capacity of 48 Kbps.

- Transparency and Data Compression

The DTN is transparent to any code and is able to transport any combination of 8 bit/ characters (e.g. EBCDIC).

Data Compression is achieved for messages coded in CCITT N° 2 and BCD, which results in 20% reduction in the volumes of data to be transported over the DTN.

Transparency and data compression result from the implementation of the new DIS-DIS LAP-B compatible protocol.

- Availability of Service

Type A and B traffic benefit from an improved node availability.

Improved type A service response time is made possible by the short DIS switching time and the ability to operate wide-band circuits.

1.3. Network Control System (NCS)

Presently, more than 9000 CRT terminals exchanging traffic with over 45 airlines computer systems and about 11000 low-speed terminals, are connected to the SITA Network. The Network itself comprises more than 500 medium-speed circuits.

The NCS is aimed at providing a man/network interface environment for network operations and network management.

The NCS is a hierarchical system composed of Regional Control Centers (RCC) and a Main Control Center (MCC).

Today the RCC of Paris is operational.

1.4. Message Storage & Handling System (MSS)

This new family of processors is designed to store and process messages, data requiring high protection (e.g. type B), or mass storage and to support new telecommunications services.

1.5. User Interface System (UIS)

The UIS will replace the SPS as concentrators and access systems. They will perform the general functions of traffic and circuit concentration, local switching, and will offer new protocol interfaces, such as, BSC/SDLC, X25. They will also support service enhancements like multi-host access.

2. SITA ROUTING ALGORITHM

2.1. Addressing

The following SITA systems : DIS, HLS, SPS, TAC, UIS, MSS, RCS, MSC etc..., as well as Airline Computer Systems (ACS) are each given a two-octect address called High Level Designator (HLD).

This address identifies the system and is network independent : the address of a system remains unchanged in case of network configuration change.

2.2. Packet Switching

Packets are transmitted over the network with the HLDs of the origin system, respectively designated as HEN and HEX.

In each DIS, packet switching is based on the HEX information.

A table (called Access Handling Table) contains one entry per HLD (i.e. one entry per system part of/or connected to the SITA Network).

This entry gives the output link for that HLD in the switching node.

Access to that table is done by a Hash-coding technique.

The AH table contains two different types of information, depending on the HLDs :

I. Local Delivery Information

This corresponds to HLD of systems belonging to the local network of the corresponding DIS.

II. Dynamically Updated Output Links

These are associated to distant systems and thus allow traffic routing over the DTN.

The updating mechanism is exposed in the next paragraph. However, one important conclusion may be drawn immediately : routing is destination dependant only, which means that for a given destination, whatever the state of the network may be, routes always have a tree pattern.

2.3. Route Updating Mechanism

2.3.1. Network Status Messages (NSMs)

Every link of the DTN is given a number between 0 and 255.

This number is used to identify the link networkwide. As every link may only be in two different states, up or down, the state of the network* may be represented by a vector, where the ith bit represents the state of link number i and is set to 1 if the link is up and 0 if it is down, this vector is called the Network Status Vector, and every DIS has one copy of it.

This vector is updated each time there is a change in the network, by high priority service messages, called Network Status Messages (NSMs).

Each time a change occurs in Network Status Vector, each DIS starts updating its A.H.Table.

2.3.2. Best Routes Tables

For routing purposes, the local network of each DIS may be divided into several destination groups. For each destination group of a DIS, in every other DIS, a table of Alternate Routes called a Best Route Table is defined. This table contains a sequential list of route descriptions giving the following information :

- The outgoing link of that route for that destination group.

- A "Link-up condition" vector associated to that route. This link-up condition vector is a binary vector with bit positions, except those corresponding to the links of the route, set to 0.

Each time a change occurs in the Network State Vector, a sequential examing of each Best Route Table is triggered and the first route, for which associated link-up conditions are verified, is selected.

The corresponding entries in the A.H. Tables are then updated with the new outgoing link. If no route has been found satisfying the link-up conditions, (important degradation of the network) the systems belonging to the destination group, are declared isolated.

2.3.3. Functional Constraints

A. Symmetric Isolation

Due to the interactive nature of type A traffic, and end-to-end protocol on type B traffic, isolation states between systems must always be symmetrical.

This implies that all Best Routes Tables in DIS A for destination groups of DIS B, and Best Routes Tables in DIS B for destination groups of DIS A, must contain the same set of routes. Of course, in the different tables the routes do not appear in the same order, so that there is no symmetry in actual traffic routing.

This rule implies that same routes are introduced in the tables just to insure this symmetric isolation constraint.

B. Bounds on Best Routes Tables Length

The maximum length of a Best Route Table, in terms of alternate routes is 12.

There is also a requirement for a minimum number of 4 alternate routes, to insure a sufficient availability of routes between nodes.

C. Precedence Constraints

To insure that the routing pattern for every destination group of every node has a tree pattern, some constraints have to be applied on the Best Routes Tables. Let us see an example of how these constraints, called precedence constraints, work.

Consider the following Best Routes Tables :

---

* No distinction is made between isolated nodes and nodes which are down.

ORIGIN NODE : HHP
DESTINATION
NODE : TYO DESTINATION GROUP : 1

01 . HHP-2-TYO
02 . HHP-1-TYO
03 . HHP-2-FRF- -TYO
04 . HHP-1-FRF- -TYO
05 . HHP- -ISP- -TYO
06 . HHP- -LON- -ISP- -TYO
07 . HHP- -ISP-2-FRF- -TYO
08 . HHP- -ISP-1-FRF- -TYO
09 . HHP-2-FRF-2-ISP- -TYO
10 . HHP-1-FRF-2-ISP- -TYO
11 . HHP-2-FRF-1-ISP- -TYO
12 . HHP-1-FRF-1-ISP- -TYO

TABLE 1 A

ORIGIN NODE : FRF
DESTINATION
NODE : TYO DESTINATION GROUP : 1

01 . FRF- -TYO
02 . FRF-2-HHP-2-TYO
03 . FRF-1-HHP-2-TYO
04 . FRF-2-HHP-1-TYO
05 . FRF-1-HHP-1-TYO
06 . FRF-2-ISP- -TYO
07 . FRF-1-ISP- -TYO

TABLE 1 B

Nodes are given three-letter codes, links are identified by their extreme nodes, plus a number, if we have multiple links between the nodes.

As the destination nodes and the destination group of the two tables are the same, precedence constraints apply.

We call a sub-route of a route, the position of that route from the first transit node to the destination node. For instance, the sub-route of :

HHP-2-FRF-2-ISP-TYO is FRF-2-ISP-TYO

Precedence constraints exist between routes having the same first link (i.e. same outgoing link), and express the fact that one route must precede the other in the Best Routes Tables.

For example, we have precedence constraints between :

HHP-2-FRF-2-ISP-TYO and HHP-2-FRF-1-ISP-TYO

But there is no precedence constraint between :

HHP-2-FRF-TYO and HHP-1-FRF-TYO

Rule

Routes having precedence constraints between them must appear in the Best Routes Tables, in the same order as their respective sub-routes appear in their own table.

Example :

HHP-2-FRF-1-ISP-TYO cannot be introduced in the the table before HHP-2-FRF-2-ISP-TYO.

Consequence :

A route cannot be introduced in a table, if the sub-route does not appear in its corresponding table.

REMARK 1 : Routes having precedence constraints between them can be interleaved with others having no precedence between them.

Example :

HHP-ISP-TYO can appear between HHP-1-FRF-TYO and HHP-1-FRF-2-ISP-TYO.

REMARK 2 : Routes, such as, FRF-2-HHP-2-TYO, FRF-2-HHP-1-TYO, etc... cannot create precedence constraints in the HHP-TYO table, because they have HHP as a transit node.

REMARK 3 : Precedence constraints only apply to longer routes (those having a sub-route).

The demonstration that this rule is a necessary and sufficient condition to have tree patterns, is given in the attached annex.

3. BEST ROUTES TABLES GENERATION ALGORITHM

3.1. General Approach

As precedence constraints are created by a relative position in the Best Routes Tables of shorter routes, on longer ones, the Best Routes Tables will be generated with the following (small) restriction.

Shorter routes (in terms of Hop-Length) are always put in the tables before the larger ones.

This restriction limits the complexity of the problem and also has the advantage of minimizing network resource utilization, although it is not always optimal (in terms of response time).

The general approach is then the following :

- Define a block of routes as the set of routes, between a given pair of nodes, having the same Hop-Length.

- Then divide Best Routes Tables in blocks of routes.

a) Create in the Best Routes Tables all blocks of routes of length 1, then of length 2, then of length 3, etc.

b) If the length of current Best Routes Tables allows it, add entire block of routes to it.

c) Otherwise, search for the "Best" subset of the block satisfying precedence constraints and links on table length.

d) Each time a set of routes is to be added to a table, order it by applying simultaneously precedence constraints and an optimality criteria.

e) Stop when each Best Routes Table has the desired number of routes.

3.2. Analysis of Constraints

Generating the Best Routes Tables with only 2 of the 3 sets of constraints (symmetric isolation, limits on table length, precedence constraints) would be quite straight forward. However, applying all of them raises some problems, as we shall see in the following example :

Let us suppose that all routes of Hop-Length 2 have already been put in tables, and that we have the following set of routes between HHP and LON with associated precedence relation (local network of LON has been divided in two destination groups).

ROUTES	PRECEDENCE CONSTRAINTS		
	Direction HHP-LON		Direction LON-HHP
	Destination		
	Group 1	Group 2	
1. HHP-FRA-NYC-LON	-	2	-
2. HHP-FRA-PAR-LON	1	-	-
3. HHP-FRA-ROM-LON	2	1	10
4. HHP-FRA-AMS-LON	3	3	-
5. HHP-NYC-FRA-LON	-	6	-
6. HHP-NYC-PAR-LON	5	2	-
7. HHP-NYC-AMS-LON	6	5	8
8. HHP-PAR-AMS-LON	-	10	4
9. HHP-PAR-FRA-LON	8	8	5
10. HHP-PAR-ROM-LON	9	-	-
11. HHP-PAR-BEY-LON	10	12	-
12. HHP-PAR-NYC-LON	11	9	11

TABLE 2

One should read this table as follows :

The presence of a 3 in column Destination Group 2 and for route N° 4, means that in the Best Route Table in HHP corresponding to destination group 2 of LON, route N° 4 (i.e. HHP-FRA-AMS-LON) must be preceeded by route N° 3 (i.e. HHP-FRA-ROM-LON). (Note that route N° 4 must also be preceeded by route N° 1, which must preceed route N° 3).

From this table let us construct the following graph :

- The nodes are the routes of table 2.

- There is an oriented arc from node i to node j, if, and only if node j must preceed node i in one of three Best Routes Tables.

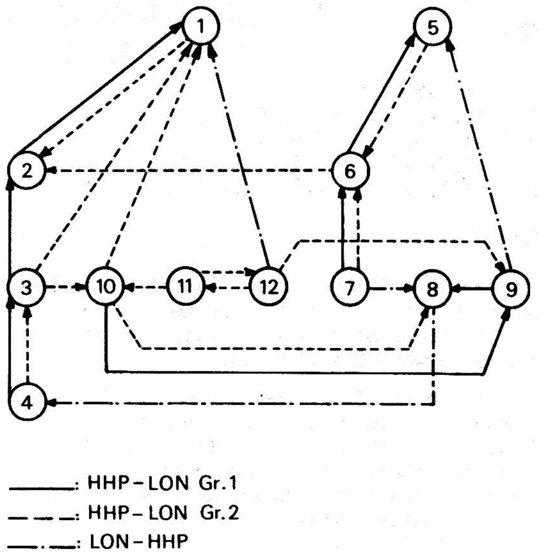

FIGURE 2 :
Graph of precedence relations of Table 2

The graph contains circuits (i.e. oriented loops) like (1,2) or (3,10,9,8,4).

Any route added to a table, must be added to every table between the same pair of nodes, hence, the meaning of circuits in the graph is the following : if we want to add to a table, a route whose corresponding node is in a circuit, we must add simultaneously all routes with the corresponding nodes of which belong to the same circuit.

This implies that it is not always possible to add one route at a time and thus, it is not straight forward to respect limits on table length, while applying the two other constraints.

Let us consider a reduced graph obtained by the following rules :

a) Merge in one node, all nodes belonging to the same circuit.

b) Suppress redundant arcs, (multiple ones).

c) Suppress the arc from node i to node j, only if an oriented path exists from i to j.

We obtain the graph in Figure 3.

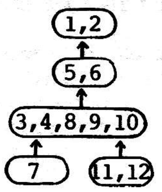

FIGURE 3 : Reduced Graph

The graph in Figure 3 gives a clear view of which subset of routes may be added to the table, if we want to respect symmetric isolation and precedence constraints. These subsets are listed below :

SUBSET		TOTAL N° OF ROUTES
A	1,2	2
B	1,2,5,6	4
C	1,2,3,4,5,6,8,9,10	9
D	1,2,3,4,5,6,7,8,9,10	10
E	1,2,3,4,5,6,8,9,10,11,12	11
F	1,2,3,4,5,6,7,8,9,10,11,12	12

TABLE 3

For instance, if the current length of tables between HHP and LON is 4, then we can select subset B and we shall end up with tables of length 8 between HHP and LON.

Once the subset of routes is chosen, the corresponding routes are successively added to each of the 3 tables, after performing each time an ordering, taking into account precedence constraints specific to each table, and an optimality criteria.

3.3. Optimality Criteria

As noted previously, this criteria can never override precedence constraints.

As the first route of the tables is the most often used (more than 60% of the time only first routes are used all over the network, they are optimized separately).

Switching nodes and links are modelled as single server M/M/1/ queues and traffic flows.

The heuristic used tries to minimize transmission delay of each flow rather than average delay over the network (which may lead to important delays for smaller traffic flows).

Propagation delays for satellite links are not taken into account in this routing optimization, because load balance of the network is important for rerouting purposes.

Once first routes have been computed, the generation of tables is done. Route evaluation is based on a bottleneck analysis.

Each route is evaluated by the load of its most loaded element.

Then the optimization routine selects the routes with a minimal value of the criteria.

Once a route is added to the table, the load of its elements (nodes and links) is updated with the traffic flow by a coefficient which decreases as the rank of the route in the table increases.

4. PROGRAMS

Two programs have been developped. The first one, named CARDIN, computes first routes of the tables and network load and transmission delays in nominal case (i.e. all links and nodes are up).

The second program named DIOR generates all the Best Routes Tables which are then written in magnetic tapes and loaded in the DISs.

Both programs are written in PL 1 (1800 instructions for CARDIN and 3000 for DIOR) and are run on an IBM 4341/2 computer. Average execution time for DIOR is 100 seconds CPU, for generating a set of 300 to 400 Best Routes Tables.

These programs have been used for 3 years now and give satisfactory results to their users.

5. CONCLUSIONS

We have exposed the present routing algorithm and the problems it raises. To the best of our knowledge, equivalent problems have not been published in the literature.

As pointed out earlier, solving the routing problem of the SITA Network can be seen as a way of generating trees in a graph. (Routing patterns in the DTN are trees). Even more, if we do not restrict the length of the Best Routes Tables, the algorithm of paragraph 3, together with the NSM mechanism, may be seen as a way of efficiently generating and storing all the trees of a graph (it suffices to construct tables including the bounded set of all non-looping routes between every pair of nodes and when applying the route update mechanism to eliminate the cases corresponding to an isolation between any two nodes).

ANNEX I. NECESSARY CONDITION

We suppose that the first route from node 2 to node 3 is (see Figure 4) : 2-4-5 and that the first route from node 1 to node 3 is : 1-2-5-3

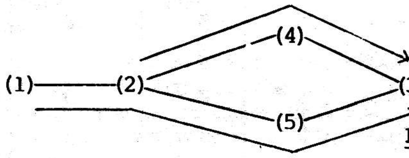

FIGURE 4

in that case the routing pattern corresponding to destination node 3 is not a tree, hence the Necessary Condition.

II. SUFFICIENT CONDITION

We want to prove that if we have a routing pattern which is not a tree, this necessarily implies that the rule is broken.

Consider Figure 4 supposing that for a given state the route from node 2 to node 3 is : 2-4-3 and that the route from node 1 to node 3 is : 1-2-5-3. This implies that links 1-2, 2-4, 4-3 are up and thus that route 1-2-4-3 is up too. Thus, either this route does not appear in the table from node 1 to node 3, or it appears after route 1-2-5-3. In each case the rule is not respected, hence the Sufficient Condition.