Improving topological routing in N2R networks — **Source link**

Jose Manuel Gutierrez Lopez, Ruben Cuevas Rumin, Jens Myrup Pedersen, Ole Brun Madsen

Institutions: Aalborg University, Charles III University of Madrid

Related papers:

- A density-based general greedy channel routing algorithm in vlsi design automation

- Routing's algorithms and considerations

- A cost transformation based routing mechanism and its applications on constraint-routing problems

- Topological routing in large-scale networks

- Honeycomb networks: Topological properties and communication algorithms

Improving Topological Routing in N2R Networks

Jose M. Gutierrez Lopez[1], Ruben Cuevas Rumin[2], Jens M. Pedersen[1], and Ole B. Madsen[1]

[1] Network and Security Department, Aalborg University, Denmark
Niels Jernes Vej 12 9220 Aalborg
`jgl@kom.aau.dk, {jens,obm}@control.aau.dk`
[2] Departamento de Ingeniera Telemtica Escuela Politcnica Superior Universidad
Carlos III de Madrid
Av. Universidad, 30, Edif. Torres Quevedo. E-28911 Legans (Madrid)
`rcuevas@it.uc3m.es`

Abstract. Topological routing is basically table free, and allows for very fast restoration and thus a high level of reliability in communication. It has already been satisfactorily proposed for some regular structures such as Grid or Honeycomb. An initial proposal has also been developed for the N2R structures. This paper proposes a modification of this previous algorithm, and in addition two other alternatives. The three options are systematically analyzed in terms of executing time and path distances, showing that trade-offs are needed in order to determine which algorithm is best for a given case. Also, the possible practical applications the methods could have, are discussed for different traffic scenarios.

Keywords: Topological routing, N2R, QoS.

1 Introduction

Topological routing is an alternative to traditional routing methods, based on tables. It allows for very fast restoration, and is particularly well suited for large-scale communication where table updates can be time consuming and introduces significant overheads. It has been successfully proposed for a few regular structures such as Honeycomb and Grid [1]. Related to this topological routing idea there are Several studies about Small world networks (SWN) which demonstrate that with limited knowledge of the network, a greedy algorithm can construct short paths using only local information. Examples of these studies are Kleinberg's Small-world models [2] or the Watts and Strogatz Ring Model [3] which are examples of the huge number of publications that prove the existence of topological routing algorithms. The question is if an existing algorithm for N2R structures will perform efficiently.

An initial proposal for using it in N2R structures can be found in [4]; an N2R structure is a generalized Double Ring (DR) topology, where the inner ring links do not interconnect physically neighbour nodes. See Fig 1. A deeper introduction

to the N2R structures is given in 2. As there is already a base to work with, this paper intends to improve the initial proposal and make a more systematic analysis of the performance in terms of execution time and path distances, to decide which is the best option to continue working on, to obtain at least a method with similar properties as table routing techniques, if not better.

The future networks will demand better characteristics for the different services supported. Those characteristics can be related with reliability and a number of other performance metrics [5].

One of the motivations of this paper was the further work described at [4], where an idea of improving the path distances obtained, and especially the maximum distance in communication between nodes, was launched. The path length is directly related with the availability [5] of the communications, and the budget related with fiber placement [6]. The availability of a system is proportional to the possible failure points on a communication. Therefore, if the number of those failure points is reduced, the availability will be higher. The maximum path length is an important factor to be able to guarantee a certain level of QoS at a given network.

When building a network, the budget is always a conditioning factor. The highest investment on a fiber optic network is the fiber placement (about 70 %) [6]. Minimizing the paths reduces the length of ditches needed to dig up to install the fiber, hence, this reduction directly implies an important reduction on the budget.

Wired networks (such as FTTH) have the handicap of the fiber physical cuts which implies the loss of connectivity between nodes [7]. The topological routing allows a fast adaptation for a network in case of a failure [4]. If more than one path can be offered to establish a communication between any pair of nodes, the consequence is a higher reliable network, decreasing the probability of loss of communication.

When offering topological routing, the routers do not need routing tables reducing the number of operation to be developed. The question is if more complex routing algorithms are capable of obtaining better communication paths with no dramatic impact on the delay of the packet which can be a critical factor. Then, if this goal is achieved, the performance of the network can be improved.

The goal is to find a general algorithm that gives the opportunity of routing topologically a packet from any node to any node at any possible N2R configuration, improving the previous results. It is assumed that if there is a general algorithm for all the configurations, there is feasible topological routing solution for any specific case.

In case that the improvement is achieved, then more arguments can support the N2R structure to be considered as an option on the network structure design at any layer level.

The contribution in this paper forms the first step towards a useful algorithm for topological routing in N2R. We provide algorithms based on node addresses instead of tables, and evaluate their performance in terms of execution times

and path lengths. The adaptiveness, which is indeed crucial, must be dealt with in future research.

The structure of the rest of the document is as follows. Section 2 treats the definitions and the proper notation to understand the network structure under study. Section 3 introduces the modification of the previous algorithm, and the two new proposals implemented for the topological routing. In Section 4 the proposed algorithms are explained with formulas and examples, and in section 5 the algorithms are simulated and the results are commented. Finally, Section 6 exposes the conclusions extracted from this paper, and it introduces the further work to be developed in Section 7.

2   Definitions and Notation

The first of the definitions are given to understand the whole concept of this paper, topological routing. Topological routing is understood as follows:

At a given address scheme, from any node, any packet can be routed given only knowledge of the addresses of the current node and the destination node, with no routing tables involved [1].

The number of nodes in the N2R structure is any positive even integer, larger or equal to 6. These rings each contain the same number of nodes (p). Links in the outer ring and the links interconnecting the two rings can be described in the same way as the DR structure, but links in the inner ring are interconnecting node I_i and node $I_{(i+p)modq}$, where q is a positive integer. To avoid forming two separated networks in the inner ring, q must fulfil $gcd(p, q) = 1$ (Greatest Common Divisor), also q is evaluated from 1 to p/2 [8].

These N2R structures have been studied and compared to other degree three topologies. The results show that N2R structures are preferable regarding low delays, high bandwidth and reliability, and also when errors occur [8]. Therefore, the effort of improving the topological routing might end up in possible solutions that, in the future, can be compared to the table routing techniques at a real scenario. At the moment, the goal is to solve the problems identified at the first algorithm proposed at [4], to be able to implement an efficient routing algorithm in terms of path distances and delay of communications.

The addresses of the nodes are given in a certain way to make the algorithm as easy and fast as possible. The outer ring nodes addresses vary from *0* to *p-1* counterclockwise and the inner ring addresses vary from *p* to *2*p-1*. The relation between the outer and the inner nodes, is in the way that the outer node *X* is connected to the inner node *X+p*.

This address system allows simple operations at the programming of the algorithm, and the addresses of the neighbours are well known by every node.

Each node is connected to the neighbours with three links. In each of the cases there is a possibility of naming them as left, right and center *(L, R, C)*. Every node knows the address of the neighbors by looking at the name of the link it is connected to. At a given address *X* of a outer node, to follow the link *L* means to reach the node *X+1*, link *R* means *X-1* and link *C* means *X+p*. In
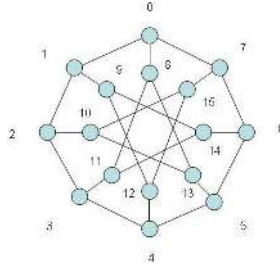
Fig. 1. N2R(8,3) notation

the same way with the inner nodes, to follow the link L means to reach the node $X+q$, link R means X-q and link C means X-p (considering *mod p* for all the neighbours addresses at the same ring). Fig. 1 illustrates this idea.

The node mechanism to treat a packet is to identify the destination, run the proper algorithm, and then forward it using one of the three links it is connected to. Before and after this operation, the node has no task to work on such as neighbours recognition or distance vectors. Therefore the node tolerance to saturation is higher due to the less number of operations to execute.

3   Node Decision

When routing a packet, the nodes have to make the decision of which of their three links should be used. This decision is based on an internal algorithm that calculates the values of the different possibilities (distance in number of hops).

In this section the methodology and the different options available to obtain the different algorithms to support topological routing over N2R structures are analyzed. This analysis is theoretical, and based on the discussion of the different options, at Section 4 the algorithms implementation is explained.

The assumed information implicit at every node is:

– p and q of the N2R structure
– Node Id (address): N_x

Each of the nodes has the possibility of knowing exactly which link to forward the packet, without knowing the source node.

For the purpose of the packet routing, three possibilities are studied and compared to obtain the feasibility of the use of topological routing at a N2R structure. The subsections 3.1, 3.3 and 3.2 describe the three algorithms proposed: Fast Response Algorithm, Optimal Path Algorithm and Balanced Algorithm.

3.1   Fast Response Algorithm (*FRA*)

The algorithm proposed is based on the previous algorithm treated in [4]. The mathematical properties used are the same, but it is implemented with some slight differences.

The previous algorithm made a difference between the source and the middle nodes. The procedure was to identify at the source node the kind of communication, (using the outer ring or the inner ring), and then add flags to the header of the packet to identify the type of communication at the middle nodes. These flags are helpful when the current and destination nodes are to reduce the number of instructions to execute:

– When current and destination is both at the outer ring, and the source node was also at the outer ring.
– When current and destination is both at the inner ring.

At the rest of the cases the middle nodes will have to execute the whole algorithm, since this flags are not helpful. Further explanation at [4]. In our implementation, however, we do not use these flags. Thus, in the modified algorithm it does not matter if nodes are source or middle nodes. The motivation of the modification was also to try to have a header as small as possible to complete a communication and to try to adapt it to the theoretical definition of topological routing (just current and destination addresses).

The rest of the mechanism is the same as the previous algorithm, but in any case it is described and commented to give a better understanding of the two other proposals. The attribute *"Fast Response"* is given due to the goal of a low execution time algorithm by using simple operations and a relatively short script.

A main reason to implement this algorithm was to obtain values for the executing time, (at the previous work it was not considered,) in order to be able to include another testing factor in the analysis of the three possibilities. Hence, even though there will be executed fewer instructions at some occasions, (the cases when the flags can be used), the executing time will be assumed to be the same due to their similarity.

This algorithm gives as a result the same paths as the previous one. Hence, it is already known that the result of the paths used in the communications using this method will not always give an optimal solution in terms of distance (number of hops). This result will be compared with the *"Optimal Path"* and *"Balanced"* solution at Section 5 to find the balance between total path length and path completion time.

The algorithm is theoretically described below and the implementation information is explained in Section 4.

A source node is selected to start the communication, then the relative position of the destination node is calculated to determine the orientation of the communication. The structure is virtually split in half from the source node, if the relative position is at the right half of the structure, the communication will be established in the right direction, and vice versa with the left half. The same method is applied for all the possible communications, without mattering the position of the nodes (at the inner or outer ring). This method simplifies the implementation of the algorithm to a few instructions. Fig. 2 illustrates this method with an example of the communication obtained.
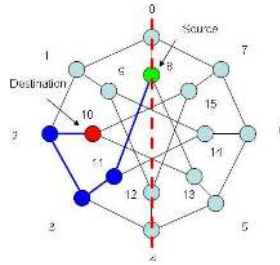
Fig. 2. N2R(8,3) Fast Response Path calculation

Then, two of the three links are left as possible gateways. The decision between these two is made by comparing the distances using the outer ring, the inner ring, or both for the communication. Each type of communication is directly related with a link, hence the link can be easily recognized.

The best result is used to forward the packet, and at the next node the same operations are realized, until the destination node is reached. Thus, in every node the next hop is found by a qualified guessing on which potential hop is closest to the destination.

3.2   Balanced Algorithm (*BA*)

The name *"Balanced Algorithm"* is given because it is a priori expected that the values of path length and execution time will be in between the values of the other two algorithms.

This method considers to use the three link to forward the packet (the *FRA* only considers two from the very beginning, the third one is discarded at the side recognition task). The idea is to find the values of the different ways the communication can be established, using the outer ring, the inner ring or both. The procedure is the same as at the *FRA* but also including the opposite orientation of the communication (if the source and destination are at the right half of the network, the communication oriented to the left is also considered).

Including the third possible gateway has the consequence that more instructions and conditions are needed for running the algorithm. Hence, the difference between the executing time and the path length obtained must be analyzed, to see if the costs make it a real possibility.

The Fig. 3 illustrates the same example as at Section 3.1 obtaining a better result (in terms of distance) at the end of the communication.

3.3   Optimal Path Algorithm (*OPA*)

In this case the algorithm implemented obtains the optimal solution in terms of number of hops. The results obtained were compared to the results of previous works, to verify the attribute *"Optimal Path"* [9]. This algorithm includes a loop to be able to find the best solution, which makes the execution time much
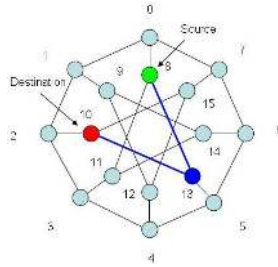
Fig. 3. N2R(8,3) Balanced Algorithm calculation

longer, directly related with a longer delay at the nodes. Therefore, the time a node will be treating a packet, is obviously longer than the *FRA* or the *BA* and the discussion must be focused on the trade off between execution time-path distance.

In case that the paths are shorter, less nodes are involved in the communication between a pair of nodes. Hence, the algorithm needs to be executed less times to reach the destination. At Section 5 this discussion will give the parameters involved in the three cases.

4   Algorithm

At this section the algorithms used to prove the theoretical statements formulated at Section 3 are explained in detail. From the theoretical point of view of the topological routing methods are quite simple, but at the time of programming an unified algorithm for all the possible configurations, it turned to have some difficulties for very specific cases (the *OPA*).

Next, the basic mechanism of the algorithms is deeply commented and explained with graphic examples. This explanation focuses on the mathematical properties of the N2R to define the general rules that must be followed to work on the topological routing issue. One of the bases of the algorithms is that a packet is never routed to the same node twice. Therefore there is no possibility of loops, count to infinite problems, or suboptimal results.

4.1   Fast Response Algorithm (*FRA*)

The algorithm implemented in this case, following the rules defined at Section 3.1, is described as follows:

– Side Recognition: Being S and D any source and destination nodes and their id N_S and N_D, the orientation of the communication is easily calculated by formula (1).

$$D_{out} = N_D'' - N_S \tag{1}$$

For this calculation the ring that the nodes belong to must be considered. The following operations, formulas (2) and (3), are required to obtain a result in the necessary range to make the comparison [1]:

$$N'_D = \begin{cases} N_D & \text{if } N_S < p \text{ \& } N_D < p \text{ (O - O)} \\ N_D & \text{if } N_S \geq p \text{ \& } N_D \geq p \text{ (I - I)} \\ N_D - p & \text{if } N_S < p \text{ \& } N_D \geq p \text{ (O - I)} \\ N_D + p & \text{if } N_S \geq p \text{ \& } N_D < p \text{ (I - O)} \end{cases} \tag{2}$$

$$N''_D = \begin{cases} N'_D & \text{if } N'_D > N_S \\ N'_D + p & \text{if } N'_D < N_S \end{cases} \tag{3}$$

Applying these properties, the value of D_{out} gives the orientation of the communication as it is explained at (4):

$$D_{out} \begin{cases} \leq p/2 & \text{Left orientated} \\ > p/2 & \text{Right orientated} \end{cases} \tag{4}$$
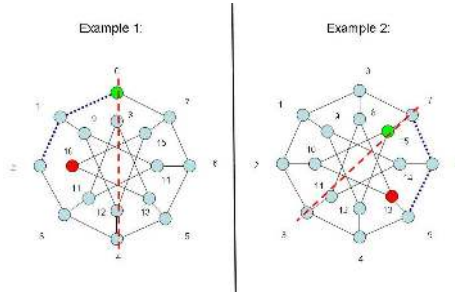

Fig. 4. Examples of Side Recognition

In Fig. 4 there are two examples illustrating this idea. Example1 (N2R(8,3), p=8 ,q=3):

- $N_S = 0$ (outer ring) and $N_D = 10$ (inner ring)
- The nodes are located at different rings:$N'_D = N_D - p = 10 - 8 = 2 = N''_D$.
- $D_{out} = N''_D - N_S = 2 - 0 = 2 \leq p/2$
- Left side orientated communication.

Example 2 (N2R(8,3), p=8 ,q=3):

- $N_S = 15$ and $N_D = 13$. Both at the inner ring.
- $N'_D = N_D$ and $N'_D < N_S$. Then $N''_D = N'_D + p = 13 + 8 = 21$.
- $D_{out} = N''_D - N_S = 21 - 15 = 6 > p/2$
- Right side orientated communication.

---

[1] O=Outer Ring, I=Inner Ring.

– Link Decision: The link decision method is to find the best option between the two links available for the communication (the third one was already discarded at the side recognition step). Hence, the criterion is to find the shortest distance (using the outer ring, the inner ring or both). The value of the distance using the outer ring (DT_{out}) is related with the previous value of D_{out} (1). There are two considerations related with this value:

• Depending on the orientation of the communication, the value of D_{out} must be converted as it is shown in (5):

$$D'_{out} = \begin{cases} p - D_{out} & \text{if } D_{out} > \text{p}/2 \\ D_{out} & \text{if } D_{out} \leq \text{p}/2 \end{cases} \quad (5)$$

The reason is that the value obtained was the distance orienting the communication to the left. Hence, since the number of nodes at the outer ring is p, this operation will give the distance using a right oriented communication in the case of being shorter. At Fig. 4 the hops taken at the outer ring (D'_{out}) are represented with a blue dotted line at both examples.

• To obtain the total value of the distance using the outer ring (DT_{out}), the hops in between rings must be considered (D_{rout}). The total distance is given by formula (6).

$$DT_{out} = D'_{out} + D_{rout} \quad (6)$$

The value of D_{rout} depends on in which ring the source and the destination are located, the possible values are presented at (7) :

$$D_{rout} = \begin{cases} 0 \text{ if } N_S < p \text{ \& } N_D < p \text{ (O - O)} \\ 1 \text{ if } N_S \geq p \text{ \& } N_D < p \text{ (I - O)} \\ 1 \text{ if } N_S < p \text{ \& } N_D \geq p \text{ (O - I)} \\ 2 \text{ if } N_S \geq p \text{ \& } N_D \geq p \text{ (I - I)} \end{cases} \quad (7)$$

The inner ring distance (DT_{in}) is related with the number of hops at this ring to reach the destination (D_{in}). This value is easily calculated by formula (8)[2].

$$D_{in} = Round(D'_{out}/q) \quad (8)$$

There are two possibilities for (DT_{in}), using the inner ring or both, it is explained at (9):

• If D'_{out}/q is an integer, the communication will only need to use the inner ring.

• If D'_{out}/q is not an integer, the communication will use both of the rings.

$$DT_{in} = \begin{cases} D_{in} + D_{rin} & \text{if } D'_{out}/q \in \mathbb{I} \\ D_{in} + D_{rin} + Dq_{out} & \text{if } D'_{out}/q \notin \mathbb{I} \end{cases} \quad (9)$$

The values in these cases for D_{rin} depending on the ring location of the nodes are presented at (10).

---

[2] The Round function converts a real number to the closest integer.

$$D_{rin} = \begin{cases} 2 \text{ if } N_S < p \text{ \& } N_D < p \text{ (O - O)} \\ 1 \text{ if } N_S \geq p \text{ \& } N_D < p \text{ (I - O)} \\ 1 \text{ if } N_S < p \text{ \& } N_D \geq p \text{ (O - I)} \\ 0 \text{ if } N_S \geq p \text{ \& } N_D \geq p \text{ \& } D'_{out}/q \in \mathbb{I} \text{ (I - I)} \\ 2 \text{ if } N_S \geq p \text{ \& } N_D \geq p \text{ \& } D'_{out}/q \notin \mathbb{I} \text{ (I - I)} \end{cases} \tag{10}$$

The value of Dq_{out} is the distance from the node where the communication jumps from the inner ring to the outer ring (N'_q), to the destination node. This value is calculated as D_{out} by formula (1) and realizing the same necessary conversions. The value of N_q is calculated by formula (11), "+" for left oriented and "-" for right oriented:

$$N_q = N_S \pm D_{in} * q \tag{11}$$

The same as in formula (1) the values must be given in the proper range. Therefore, assuming that $0 \leq N_S < p$ (value converted if necessary) there are two conversions depending on the orientation, see (12):

$$N'_q = \begin{cases} N_q - p \text{ if } N_q \geq p \text{ (Left orientated)} \\ N_q + p \text{ if } N_q < 0 \text{ (Right orientated)} \end{cases} \tag{12}$$

In this way the values of N'_q are in the proper range to work with.
– Comparison: At the end, the comparison between the values of DT_{in} and DT_{out} (the shortest one is chosen) determines how the communication will be established, hence, the link to use to forward the packet. The algorithm does not calculate the paths, it calculates the possible distances to be able to decide which link to use to forward the packet.

Fig. 5 illustrates two examples to better understand the real meaning of the values calculated. At the Example 3 (same procedure for example 4):
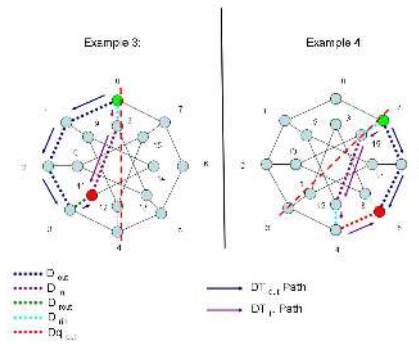

Fig. 5. Examples Node Decision

- N2R(8,3), p=8 ,q=3
- DT_{out}:
  $N_S = 0$ and $N_D = 11$, different rings: Convert $N'_D = N_D - p = 11 - 8 = 3 = N''_D$
  $D_{out} = N''_D - N_S = 3 - 0 = 3 = D'_{out}$ (Left orientated)
  $D_{rout} = 1$ (Outer-Inner)
  $DT_{out} = D'_{out} + D_{rout} = 3 + 1 = 4$
- DT_{in}:
  $D_{in} = D'_{out}/q$=3/3=1, integer. Only inner ring used.
  $D_{rin} = 1$ (Outer-Inner)
  $DT_{in} = D_{in} + D_{rin} = 1 + 1 = 2$
- $DT_{in} < DT_{out}$, therefore the link is used C (center). Start all over from the next node.

4.2   Balanced Algorithm (BA)

The basic idea is the same, to find the DT_{out} and DT_{in} to make a decision. In this case there are two distances related with the inner ring to calculate, one for a clockwise communication and the other for counterclockwise.

Being D_{in1} and D_{in2} the hops at the inner ring to establish a connection in both directions, D_{in1} is the value obtained using the already commented formula (8). D_{in2} corresponds to formula (13).

$$D_{in2} = Round((p - D'_{out})/q) \tag{13}$$

The procedure to calculate the values of DT_{out}, DT_{in1} and DT_{in2} are exactly the same as in Section 4.1, applying the same proper conversions. These three values are compared and the best option is used to forward the packet. The procedure is the same at the next node until the destination is reached.

4.3   Optimal Path Algorithm (OPA)

This algorithm obtains at the end of any communication the shortest path possible for all the situations and N2R configurations. In order to achieve this goal the algorithm must be much more complex than the FRA and the BA.

The problem identified is that the BA did not consider to take a whole loop or more at the inner ring to establish the communications. Hence, when $q \approx p/2$ the distances obtained were not the optimal ones, this problem was already identified at [4]. To solve this problem it is unavoidable to try out this looping around when implementing the algorithm. Instead of calculating directly the values of D_{in1} and D_{in2} it tries all the possible values to find the best in each direction. This loop range is proportional to p.

Then when the loop has found all the values of D_{in1} and D_{in2} and the best option is considered, the values of the rest of the necessary parameters are calculated to obtain the best DT_{in1} and DT_{in2} to compare them with the value of DT_{out}, just as the previous cases.

At the time that the three best values are calculated, compared and the optimal is chosen, the link related to that option is selected to forward the packet. Then the procedure starts all over again from the next node until the destination is reached.

To better understand the effect of using the loop at Fig. 6 the path obtained for the *FRA*, *BA*, and *OPA* are represented. The configuration chosen is N2R (12,5) since it was one of the problematic situations found.


Fig. 6. N2R(12,5) Communication Example

5   Analysis

In this section the comparison of the results of the different cases is discussed. The different natures of traffic demand different QoS levels and different requirements [4]. Therefore, two parameters, among others, are analyzed to study the feasibility of each algorithm depending on the traffic nature:

– Path Length: The fact of obtaining shortest paths as possible is related with the availability of the network, the BW cost of the packet routing and the network installation cost. Minimizing the paths, reduces as much as possible the probability of failure caused by links or nodes failures, there are less failure points. If the paths are shorter, in connection with the installation, it will be necessary to use less fibre to build the network (less fiber at each link to support the traffic). In the case of taking minimum routes reduces the cost in BW of the communications, the less cost the more communications are allow with the same link capacity. All these properties are explained in detail in [5].

   The P2P file sharing traffic creates, in most countries, more than 50% of the total traffic. The delay is not a critical factor for this kind of traffic, therefore it would be possible to use the longer algorithm (execution time) [10]. In this way the amount of fiber required to support these services can be minimized, due to the shortest distances. This also applies to similar kinds of traffic such as ftp or SAN (Storage Area Network).
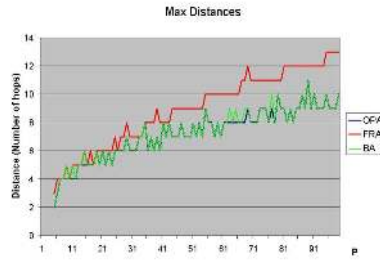
– Delay: Obtaining minimum delay at the nodes to complete a communication improves the performance of the network by reducing working tasks at the nodes [5]. A short delay is needed for the upcoming real time networks [1]. For real time traffic, such as video conference or TVoIP, the delay becomes
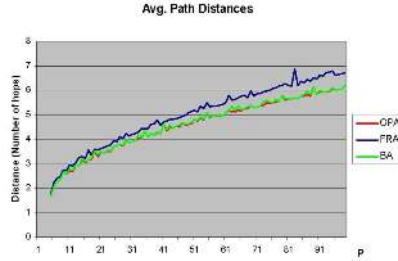
(a) Average Path Distances
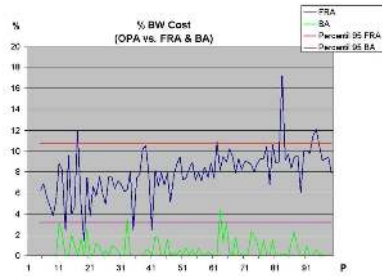
(b) % BW Cost

(c) Maximum Path Distances

Fig. 7. Distance Graphs

a critical factor. Therefore, the method used must be able to minimize this, but with the consequence of more installation budget and more BW cost per packet if the distances are longer. The path length can impose delays as well; even though an algorithm has a fast execution, if the paths are too long, the algorithm must be executed too many times (delaying the packet more at each execution).

Hence, the three algorithms were run to obtain the average distance for a communication, diameter (maximum distance), average path completion time (the time spent at the nodes to reach the destination) and maximum path completion time[3] varying the value of p from 5 to 100 (200 nodes). The q selected for the graphics at each value of p is the one that obtains the shortest average path length. These values were obtained simulating all the possible communications (from all the nodes to all the nodes), and then the average was calculated.

---

[3] The execution time was obtained under the same conditions, since depending on the machine where these algorithms are executed the result will vary. It is assumed though that the proportion will be maintained. The machine used is a Genuine Intel(R) CPU T2050 @1.60 GHz (2 CPU) and 1GB of RAM.

Fig. 7(a) illustrates the average path length for the p, for the ranges of the three algorithms studied. The best option as expected is the *OPA*. The *BA* obtains very small differences for the average values, hence it could be also a possibility to consider depending on the rest of the factors. In the result for the *FRA* there is a significant difference, increasing at the same time as p.
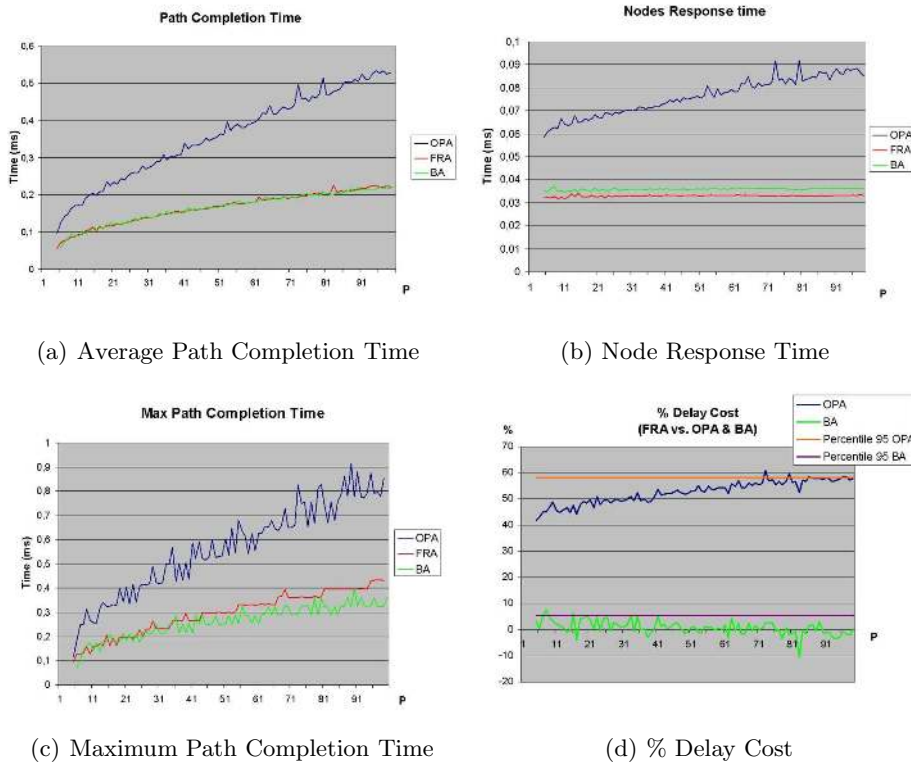

(a) Average Path Completion Time

(b) Node Response Time

(c) Maximum Path Completion Time

(d) % Delay Cost

Fig. 8. Delay Graphs

Fig. 7(b) illustrates the relative percentage of the BW cost of using the longer solutions under the same conditions (the same network capacity), in other words, the percentage of pps (packets per second) reduction. The value of the *"Percentile 95"* [4] is represented for both cases. This value gives a pseudo maximum avoiding the maximum peaks at very special situations. The use of *BA* implies around maximum 2.3% reduction of the pps and the use of the *FRA* implies a higher reduction, around 10.7% of the maximum. These values can be applied in case of analyzing the cost of having the same throughput for the different cases. To

---

[4] The pth percentile is a value such that at most (100p)% of the observations are less than this value and that at most 100(1 - p)% are greater. (p is a value between 0 and 1).

obtain the same performance, the percentages given are the extra amounts of fibre needed to be installed at the links. Therefore, the increment of the budget concerning the civilian construction when implanting this kind of networks.

This analysis assumes symmetric traffic (same traffic from all the nodes to all the nodes) to obtain a general overview of the behaviour of the algorithms. With asymmetric traffic models the results will vary, e. g. if there is a high traffic between nodes close to each other, the cost of the BA and FRA will be lower since they obtain optimal, or close to optimal, paths. But if there is a higher traffic between pairs of nodes further away, then it can be assumed that the cost would be higher due to the differences on the distances to the optimal solution.

To be able to guarantee a certain QoS level for any communication the maximum values should also be considered. The Fig. 7(c) illustrates the maximum distance for the p given range. As in Fig. 7(a) the best option is the OPA, but noticing the small difference with the BA. For 93.7% of the p values the maximum value is the same. Only in very specific cases (the rest 6.3%) this value changes only by one hop. The difference with FRA is very significant. Only for 9.5% of the p values the maximum value is the same.

This small difference between the OPA and BA is due to the q values represented. It is assumed that when implementing a network, the best configuration is used for any number of nodes required. At the most of the N2R(p,q) there are two optimal q values. One is $q \approx p/2$, and the other is much smaller [4].

In Section 4.3 it was already identified that the problem of the BA with high q values. Hence, for equal results for the OPA distances, the smallest q value (in case there are two optimal) was used for the analysis and the graphics. If the two algorithms are executed over configurations with larger q values, the differences would be much more significant.

A priori an idea could be to add a condition to the OPA to decide when it has to use the loop or not. It is clear that for values of $q \ll p/2$ it will not be necessary. The problem is that there are apparently no mathematical relation between p and q to be able to give a limit for the condition. Therefore it was not implemented.

Fig. 8(a) illustrates the average path completion time for the three algorithms. The conclusion is the opposite than the path length analysis. The cost in delay of obtaining the optimal paths *(OPA)* is very high. The completion time for the other two cases is very similar; depending on the number of nodes the BA completes the path even faster than the FRA. This fact is due to the shorter paths obtained, and despite the node response time always being smaller for the FRA. Fig. 8(b) illustrates the Node Response Time for the three cases as a function of p.

The nodes response time at the FRA is independent of p and always being the FRA time (around 32 μs) shorter than the BA time(around 36 μs). The opposite situation is the OPA, the time is much longer than the other two cases, and increases with p which makes it scale badly.

Fig. 8(c) illustrates the maximum path completion time. Obviously, the maximum time of the OPA is much longer than the maximum time for the other

two cases. Out of the comparison between the *FRA* and the *BA* an interesting conclusion is obtained. For 88% of the p values at the *BA* the maximum delay for a communication is the minimum among the possibilities. For $p < 50$ (100 nodes) the maximum times are similar, being in some situations shorter than the *BA* and in some situations longer. But for $p > 50$ the maximum time is always longer with the *FRA*, and in addition to the fact that it always obtains the longest paths, this algorithm can be discarded for these p values.

The last result obtained is shown in Fig. 8(d) which represents the cost in delay of using the *BA* or *OPA* versus the *FRA*. This cost is represented as a percentage, depending on the value of the *FRA* delay, the values for the other two cases is "$(1 + 0.a) * FRA_{delay}$" being a percentage. As for the distance analysis the percentile 95 is represented for both cases. The *OPA* values, as in the rest of the execution time analysis, are extremely high compared with the other two cases. At the *BA* in some cases the percentage is even negative, (no cost, delay reduction) since it takes less time to reach the destination.

6   Conclusion

This paper introduced and compared different algorithms for topological routing in N2R networks, which can increase the reliability and availability compared to traditional routing schemes. At this first stage the schemes are still not adaptive, and as such further research is needed before an implementation can take place.

Three different algorithms were proposed, a modification of a previous algorithm and two new proposals. They all achieved the goal of obtaining paths for any communication between nodes. Those different algorithms have much different characteristics, but none of them obtains the optimal result for all the parameters tested. Thus, it is a matter of trade-offs between parameters.

There is a possibility of an algorithm which, at the end of the communication between any pair of nodes, obtains the shortest path for all the possible N2R configurations, Optimal Path Algorithm *(OPA)*. But the cost in delay for this case is the highest among the possibilities, probably at some circumstances unacceptable.

On the other hand there is no possibility of an algorithm that obtains the smallest delay in all situations. Related to this factor the Fast Repose Algorithm *(FRA)* and the Balanced Algorithm *(BA)*obtain the best result, depending on the number of nodes. The BW cost for the *FRA* is higher than the cost for the *BA*, and the maximum value for the path distances makes it an improbable solution. In the same way the previous algorithm would have the same BW cost, hence even if there is a reduction of a few instructions on the algorithm of the middle nodes, due to the flags use to identify the path at some situation it could be discarded.

The *BA* does not obtain the best results possible for delay and path distance in all the situations, but the difference with those minimum values is not very costly (the most around 2.3% in BW cost and and 5.3% in delay cost[5]).Another

---

[5] Percentile 95 values.

important characteristic is the maximum values for the two factors (paths length and completion time). Both are very similar to the optimal, therefore almost the same QoS can be guaranteed (only in 12% of the cases for the delay and 6.3% of the cases for the path distances, the minimum values are not obtained).

These results demonstrate the availability improvement over the previous studies about the N2R topological routing. The average path lengths and the maximum path lengths have been reduced. The budget for the installation of the fiber can be reduced if this new method is applied over the previous algorithm.

But there is still work to do in order to improve the reliability of the network. A number of tasks mentioned in [4] are not yet solved. There might be some possible methods to be able to route a packet using a second path (in case of a failure of any element of the first one) to provide a reliable network system. This idea is explained at Section 7.

7 Further Work

As an introduction to further work three theoretical solutions are proposed to be simulated, tested and proved in the next step of this topological routing issue, in order to offer adaptive path redundancy and fault tolerance.

- Link Restoration: The packets are routed normally through the network. In case a next hop in the communication is not reachable, the packet is sent directly using the only link available (one of the other two is the failure and the last one is the arrival link). At the next steps the shortest path is followed considering that the packet cannot return to a previous node. This method is fast but at the same time, some of the communications could take too long unnecessary paths.
- Path Restoration: When a failure occurs, a node detects it when it tries to send a packet through that failure. At that moment the node sends a warning message to the source node and immediately the source uses another algorithm that can route the packet using an independent (disjoint) second path. This solution would increase the routing time at the nodes due to the complexity of the algorithm, but with shorter solutions (in distance). For this method additional information about the source is needed. Another problem is the packet loss while the source does not receive the warning message, therefore this method must be improved.
- Combination: Theoretically the best option found is the combination of both of the previous proposals. At the transition time, while the source does not receive the warning message, the solution used must be the "Link Restoration" and at the moment that the network is stabilized the "Path Restoration" solution.

These proposals must be analyzed as the first path methods, the distance and delays, and simulated to obtain clear results of the possible solutions. These analyses should take into account that different failure characteristics may impose different requirements to the restoration schemes used.

References

1. Pedersen, J.M., Knudsen, T.P., Madsen, O.B.: Topological Routing in Large-Scale Networks. In: IEEE/ICACT 2004, Korea, February 2004, p. 912 (2004), Available: `http://ieeexplore.ieee.org/iel5/9073/28787/01293001.pdf`
2. Martel, C., Van Nguyen: Analyzing Kleinberg's (and other) small-world Models. In: Annual ACM Symposium on Principles of Distributed Computing archive Proceedings of the twenty-third annual ACM symposium on Principles of distributed computing, St. John's, Newfoundland, Canada, pp. 179–188 (2004), ISBN:1-58113-802-4
3. Watts, D., Strogatz, S.: Collective dynamics of small-world networks. Nature 393, 440–442 (1998)
4. Pedersen, J.M., Riaz, M.T., Madsen, O.B.: A Simple, Efficient Routing Scheme for N2R Network Structures. In: IT&T Annual Conference 2005, Cork, Ireland, pp. 69–80 (2005)
5. Grover, W.D.: Mesh-Based Survivable Networks, Options and Strategies for Optical, MPLS, SONET and ATM Network, 1st edn. Prentice Hall PTR, Englewood Cliffs (2003)
6. Sanchez, M.G.S., Rumin, R.C., Gutierrez, J.M.: Triple Play Network Modelling for Spain. MsC Thesis. Control Department Aalborg University (June 2006)
7. Pedersen, J.M.: Structural Quality of Service in Large-Scale Network. PhD Thesis. Control Departament, Aalborg University (April 2005)
8. Jorgensen, T., Pedersen, L., Pedersen, J.M.: Reliability in single, double and N2R ring network structures. In: CIC 2005. The International Conference on Communications in Computing, Las Vegas, Nevada, United States, June 2005, pp. 2–4 (2005)
9. Madsen, O.B., Knudsen, T.P., Pedersen, J.M.: SQoS as the Base for Next Generation Global Infrastructure. In: Proc. of IT&T 2003, Information Technology and Telecommunications. Annual Conference 2003, Letterkenny, Ireland, October 2003, pp. 127–136 (2003), Available: `http://www.kommunikation.aau.dk/ddn/Filertildownload/SQoSastheBaseforNextGenerationGlobal.pdf`
10. Karagiannis, T., Papagiannaki, K., Faloutsos, M.: BLINC: Multilevel traffic classification in the dark. ACM SIGCOMM 35(4), 4 (2005), Available: `http://www.acm.org/sigs/sigcomm/sigcomm2005/paper-KarPap.pdf`