

Network topology and routing algorithm based on 3-degree
node. ✓

BACKGROUND OF THE INVENTION

5 Field of the Invention

The present invention relates to a method of
arranging network nodes in a topology to improve the
routing capability of the nodes and the inter-arrival time
of data between source node and destination nodes. Since
10 the method according to the present invention is
independent of the network protocol, it can be used for
configuring and interconnecting existing network systems to
improve a delay between computers in different computer
networks.

15 The need for high-speed communication has been
the most influential factor in the design of a new network
topology and for a network topology dependent on a routing
algorithm for efficient data routing on a backbone network.

For a backbone network, the need to interconnect
20 network gateways in an efficient manner is imperative to a
good network performance in terms of end-to-end delay. The
network topology chosen to interconnect network gateways
can also be the determining factor in providing real-time
services between computer terminals of different gateways
25 as the number of gateways in the interconnected network

becomes larger.

Network topologies used to interconnect a larger number of network nodes will also determine the method used for switching technologies and topologies. The network topologies through the use of switching technologies are capable of delivering low-cost, highly reliable, scalable connectivity and service. The present invention can also be used as one of the user-defined network topologies that Fibre Channel switching topologies may support. Fibre Channel families of standards are the responsibilities of the standardization committee T11. T11 is a committee within the National Committee for Information Technology Standards (NCITS).

Description of Prior Art

There is a number of network topology already in existence and being implemented as a part of the international standards such as Token ring and Ethernet which are networks based on ring and bus topology, respectively. Interconnection network topologies can be broadly divided into 2 main classes of static type and dynamic type. Most of the static network topologies require node each having 2 and more links.

For measuring the performance of network topology, diameter of the network is of the paramount importance. Diameter of the network topology is a measure of the

shortest path between all possible of nodes in the network. The diameter of the network determines the number of links for routing a message from the source node to the destination. Diameter of the network topology determines
5 the shortest possible path and a routing algorithm can route a message from the source node to destination.. The 2 other important parameters include the degree of network nodes and the minimum expansion increment of nodes to the network. Degree of network node is a measure of number of
10 links that a network node must have. The minimum expansion increment is the minimum number of nodes to be added to the network.

A presently known network topology that uses 3-degree nodes is the cube-connected cycles network. It has
15 a large minimum expansion increment as the network size gets larger. The diameter of the network is very large as the number of nodes in the network increases. One of the most detrimental factor of this network topology is that it needs a complex routing algorithm and the algorithm is also
20 dependent on the number of nodes in the network topology. The network topology uses algorithms that must be changed as the size of network increases since a generic algorithm uses at each of network nodes becomes impossible.

Many of the existing routing algorithms and
25 network topologies are dependent on the type of network

protocol. These restrictions require implementation of new routing algorithms and network topologies with respect to legacy network systems that were already in place prohibitive both in terms of cost and technical
5 difficulties. Some of the routing algorithm requires information on the traversal path to be passed to the nodes between the source and the destination. The additional routing information puts a burden on the already scarce bandwidth resources.

10 Summary of the Invention

One object of the present invention is to improve the distance of a network topology comprising 3-degree nodes.

Another object of the present invention is to
15 improve the minimum expansion increment of nodes for a mesh type network topology.

A further object of the present invention is to reduce the complexity in the routing algorithm for the network topology, thereby enabling to implement a low cost
20 router on each of the 3-degree nodes in a network topology of a specific size.

A still further object of the present invention is to provide a highly scalable network topology capable of functioning as a backbone network architecture for allowing
25 a high-speed data delivery between arbitrary two nodes in

the network.

One further object of the present invention is to provide a network protocol independent of both network topology and routing algorithm for a point-to-point
5 connected network.

One still further object of the present invention is to eliminate passing of routing information from the source node to an intermediate traversal node and improve the band width utility in the routing management.

10 In order to achieve these objects, according to the present invention, there is provided a method of increasing the bandwidth and reducing a distance between network nodes in a mesh connected network which includes the following steps of:

15 a. converting nodes in the mesh connected network being a representative of a network sub-system such as routers, bridges, network-switching sub-systems into a 2-dimensional network topology consisting of 3-links nodes,

b. generating addresses for the 3-links nodes using
20 a naming scheme based on 2-tuple radix-n scheme to ensure a unique address for each of the 3-links nodes in a specified network using a mathematical formula; and

c. interconnecting each 3-links node to an adjacent 3-links nodes linking via point-to-point connections using
25 a mathematical criterion based on a numerical

representation for each 3-links node to form a network topology of a specified configuration.

According to the present invention, there is also provided a system including the following means:

5 means for configuring network sub-systems and interconnecting the network sub-systems into a network topology based on 3-links nodes;

 means for configuring each network sub-system using invertedly oriented and upright oriented 3-links
10 nodes; and

 means for generating a unique address or name for each of the 3-links nodes;

 means for configuring and interconnecting the network sub-systems into a network topology;

15 means for routing a data message between a source node and a destination node in the network topology using a routing algorithm for selecting a shortest path from the destination node to the destination node.

 The network sub-system such as bridges, routers
20 and gateways to be interconnected are represented by 3-links network nodes. Each of these sub-systems is implemented with 3 duplex links for interconnecting to adjacent sub-systems, thereby making 3-links nodes a representative of the network sub-systems. Each of these
25 nodes is either in upright or inverted orientation.

Interconnections between 3-links nodes are performed based on a mathematical formula. Each of the interconnected nodes is provided with a unique address or naming convention based on 2-tuple radix-n number system. The
5 addressing scheme adopted in naming the nodes is used by the routing algorithm for finding the shortest path between any 2 nodes within the network configuration.

Brief Description of the Drawings

These and other objects and features of the
10 present invention will become clear from the following description taken in conjunction with the preferred embodiments thereof with reference to the accompanying drawings throughout which like parts are designated by like reference numerals, and in which:

15 Fig. 1A and Fig. 1B are diagrams showing upright oriented node and invertedly oriented node, respectively;

Fig. 1C is a diagram showing a connection of full-duplex links between adjacent nodes;

20 Fig. 1D is a diagram showing a network topology composed of 3-links nodes;

Figs. 2A and 2B when combined as shown in Fig. 2 show a flowchart showing a network topology construction according to the present invention; and

25 Figs. 3A and 3B when combined as shown in Fig. 3 show a flowchart showing a shortest path routing algorithm

according to the present invention.

Detailed Description of the Preferred Embodiments

The invention will best be understood by reference to the following detailed description of an illustrative embodiment when read in conjunction with the accompanying drawings.

In Figs. 1A and 1B, there are shown an upright 3-links node 1.1 and an inverted 3-links node 1.2, respectively.

A unique numeric representation (value) is assigned to each link of the 3-link nodes, as shown therein. The assignment of numeric representations to respective links of the node is determined so that a correct connection between an upright node 1.1 and an inverted node 1.2 is established if a sum of the numeric representations assigned to each one link of the upright and inverted nodes is zero.

In Fig. 1C, there is illustrated a connection of a full duplex link between two nodes 1.3 and 1.4.

Referring to Fig. 1D, addressing or naming rule for each node is explained according to the following mathematical formula:

$$[1] \quad X(MST_N, LST_N) = X((y/n) \% m, y \% n);$$

For $y=0, 1, 2, 3 \dots (m \times n) - 1$.

MST and LST represent most and least significant

tuples, % represents modulus operation, m represents number of rows in the network, n represents number of columns in the network, y is +ve integer number, and (y/n) is an integer division.

5 Connections of two links of nodes adjacent to each other are established according to the following criteria.

[2] Upright Oriented 3-Links Node

X(x2,x1).+ve_Link is connected to X(x2,[(x1+1)%n]).-ve_Link
 10 X(x2,x1).-ve_Link is connected to X(x2,[(x1-1)%n]).+ve_Link
 X(x2,x1).-ve_RadixM_Link is connected to X([(x2-1)%m], x1).-ve_RadixM_Link

[3] Inverted Node Oriented 3-Links Node

X(x2,x1).+ve_Link is connected to X(x2,[(x1+1)%n]).-ve_Link
 15 X(x2,x1).-ve_Link is connected to X(x2,[(x1-1)%n]).+ve_Link
 X(x2,x1).+ve_RadixM_Link is connected to X([(x2+1)%m], x1).-ve_RadixM_Link

Referring to Fig. 2, construction of the network topology based on 3-links nodes can be performed based on
 20 the algorithm depicted in a flowchart of Fig. 2.

The construction of the network topology may be divided into two main phases, namely naming phase and connecting phase.

In the naming phase, nodes in an unconfigured
 25 network are mapped to either upright or invertedly oriented

3-links nodes. All 3-link nodes in the mapped network topology are named. The naming process consists of steps 2.1, 2.2, 2.3 and 2.4 together with unnumbered steps.

In step 2.1, network parameters are initialized
5 by assigning the number of rows m and that of column n to configure a desired network. Further, a node count $Node_Cnt$ is reset to zero. Then, in step 2.2, generation of unique names or representatives of the names is performed for all 3-links nodes contained in the network based on the
10 above equation [1].

When an unique node address is assigned to one 3-links node, it is judged if the $Node_Cnt$ is an even number. If the $Node_Cnt$ is an even number, then the 3-links node having the unique node address is assigned as an upright
15 oriented 3-links node at step 2.3. If it is an odd number, the corresponding 3-links node is assigned as an invertedly oriented node at step 2.4. When generation of node addresses and assignment to either upright or invertedly oriented 3-links nodes are completed for all 3-links nodes
20 contained in the network, the process enters into the second phase for connecting 3-links nodes after resetting node counter $Node_Cnt$, row counter m_Cnt and column counter n_Cnt to zero, respectively.

With reference to Fig. 2, the second phase of the
25 network construction provides connections of receiving and

transmitting ports of 3 links for each of the 3-links nodes in the network to be configured. This connection operation is made based on the criteria cited in [2] and [3].

In step 2.6, the receiving port and transmission
5 port of link (+ve_Link) between two 3-links nodes are connected based on the criteria of [2] and [3]. Also, in step 2.7, a connection for the receiving port and transmission port of link (-ve_Link) between two 3-links nodes is established according to the criteria [2] and [3].
10 Further, in step 2.8, with respect to link (-ve_RadixM_Link), the receiving and transmission ports are connected between two 3-links nodes according to the criteria [2] and [3] for upright oriented nodes and in step 2.9, with respect to link (+ve_RadixM_Link), the receiving and transmission
15 ports are connected according to the criteria [2] and [3] for invertedly oriented nodes. By repeating the above steps for all 3-links nodes, a network topology is constructed as shown in Fig. 1D.

Fig. 3 is a flowchart showing a routing algorithm
20 to be implemented in every 3-links node for routing messages between arbitrary two 3-links nodes in the network topology. For simplicity, the 3-links node is referred to as node hereinafter.

This routing algorithm is to guarantee the
25 shortest path between a source node to a destination node

in the constructed network topology. The algorithm requires values of m and n to know the size of the network topology. In addition, addresses of the source node and destination nodes are required for searching the shortest possible path.

In step 3.1, the number of virtual rows in the network is calculated. The "virtual row" is defined as a row consisting of only either upright oriented nodes or invertedly oriented node having the same MST_M value of the address field $X(MST_M, LST_N)$.

In the next step 3.2, the column distance between the source node and destination node is calculated according to equations given in the box of 3.2. The column distance is calculated using the LST_N values of the source and destination nodes in the address field $X(MST_M, LST_N)$. The column distance in a positive direction ($n_pos-dist$) gives a distance for which a message should traverse traversal nodes starting from the $+ve_Link$'s transmitting port of the source node and before reaching the destination node. Traversal nodes are nodes through which the message is transmitted to the destination node. The column distance in a negative direction ($n_neg-dist$) gives the number of links through which the message should traverse traversal nodes starting from the $-ve_Link$'s transmitting port of the source node and before reaching the destination

node. The unit for both `n_pos-dist` and `n_neg-dist` is same.

In step 3.3, virtual row distances are calculated using the virtual row calculated in the step 3.1. The virtual row distance in a positive direction (`m_pos-dist`) is a distance for which a message should traverse traversal nodes starting from the `+ve_RadixM_Link`'s transmitting port of the source node and before reaching the destination node. The virtual row distance in a negative direction (`m_neg-dist`) is a distance for which a message should traverse traversal nodes starting from the `+ve_RadixM_Link`'s transmitting port of the source node and before reaching the destination node. Calculations of these two row distances are made according to the conditions given in the box of step 3.3.

After the calculation of column and row distances in steps 3.2 and 3.3, the message routing is determined in the horizontal and vertical directions.

In steps 3.5 and 3.4, three possible ways of the message routing in the horizontal direction are determined. If `n_pos-dist < n_neg-dist`(step 3.5), then `+ve_Link` is selected as the message routing direction `n_direct` for the current node and if `n_pos-dist > n_neg-dist`(step 3.4), then `-ve_Link` is selected as the `n_direct`. The third possibility is `n_direct=LEVEL` in the case `n_pos-dist=n_neg-dist`. "LEVEL" means that no routing of message in the

horizontal direction is necessary. It is needed in the case that the current node has no link for the vertical direction routing due to the orientation of the current node.

5 In steps 3.6 and 3.7, three possible ways of the message routing in the vertical direction are determined. If $m(\text{dest.})/2 = m(\text{source})/2$ (step 3.6), then m_direct is set equal to LEVEL. If $m_pos\text{-}dist < m_neg\text{-}dist$, then m_direct is set to DOWN indicative of $+ve_RadixM_Link$ of the current
10 node. If $m_pos\text{-}dist > m_neg\text{-}dist$, then m_direct is set to UP indicative of $-ve_RadixM_Link$ of the current node. "LEVEL" indicates that the destination node is in the same row as the current node and prevents the routing algorithm from selecting the input port of the message as an output
15 port of the message. When m_direct is DOWN, $+ve_RadixM_Link$ is selected and when m_direct is UP, $-ve_RadixM_Link$ is selected for routing the message.

 In step 3.8, "Entry point" of message is checked to classify into four cases 3.9, 3.10, 3.11 and 3.12
20 depending on "Entry point" indicative of the origin of the message to be routed.

 If Entry point = host mode, m_direct is set according to conditions defined in the box of 3.9. If it is $-ve_Link$, m_direct and/or n_direct are set according to
25 conditions defined in the box of 3.10. In the case of

Entry point = +ve_Link, m_direct and/or n_direct are set according to conditions defined in the box of 3.11. If Entry point is either of -ve or +ve_RadixM_Link, then m_direct is set equal to LEVEL in step 3.12.

5 The variables m_direct and n_direct determined in one of four steps 3.9 to 3.12 are final directional pointers in the vertical and horizontal directions, respectively.

Step 3.13 performs an alternate routing procedure
10 according to the conditions defined in the box of 3.13. In step 3.14, the final routing output link is determined.

The source node and the traversal nodes performed the routing algorithm once until the destination node has been reached. The routing algorithm would cease to process
15 if the destination address were the same as the address of the current node. In this case, the current node is the destination node. The routing algorithm routes message out of one of the 2 possible links for traversal nodes and 3 possible links if it is the source node. This routing
20 algorithm is able to find the shortest path for message between the source and the destination nodes.

Performance criteria for the network are given as follows:

The minimum distance between 2 nodes in the
25 network is reflected by the diameter of the network.

$$\text{Diameter} = M/2 + N/2 \quad \text{for } N > M$$

Diameter = M for $N \leq M$

Where M is number of rows and N is the number of columns of the network

5 The minimum incremental number of nodes that can
be added to the existing network topology is reflected in
the minimal expansion increment. This property of the
network topology provides an indicator for the scalability
of the network topology. There is 2 ways of adding node to
10 existing network; row and column addition.

Case 1: Rows Addition

$$N_{\text{new}} = M_{\text{old}} + 2;$$

Additional Nodes, $N = 2 \times N_{old}$

Case 2: Columns Addition

$$15 \quad N_{\text{new}} = N_{\text{old}} + 2;$$

Additional Nodes, $N = 2 \times M_{old}$

Advantages of the routing algorithm for network topology using 3-links nodes are summarized as follows:

1. No routing table needs to be maintained by
20 each node or a central node.

2. Routing information of traversal nodes such as address or representative of the traversal node name need not be passed to traversal nodes and destination nodes. This cuts down the bandwidth utility during messages routing.

25

3. Computational and memory resources required for the routing algorithm is minimal and an efficient implementation can be achieved. This is mainly due to the architecture of the network topology where the routing task is distributed among the nodes connected in the network. Each node is only responsible for routing data out of the 3 possible links and the distributed routing effort carried out by the nodes in the network guarantees the theoretical shortest path are determined.

4. No high performance computational power is necessary for one node to do the routing for the whole network.

5. Because of the distributed properties of the routing algorithm, the emulation of the 3-links node can be performed easily. This emulation is required if the number of nodes to be added to the existing network is less than minimal expansion increment. The physical nodes can implement the virtual nodes easily with minimal computational and memory resources.

Although the present invention has been fully described in connection with the preferred embodiments thereof with reference to the accompanying drawings, it is to be noted that various changes and modifications are apparent to those skilled in the art. Such changes and modifications are to be understood as included within the

scope of the present invention as defined by the appended claims unless they depart therefrom.



WHAT IS CLAIMED IS:

1. A method of increasing the bandwidth and reducing a distance between network nodes in a mesh connected network comprising steps of:

5 a. converting nodes in the mesh connected network being a representative of a network sub-system into a 2-dimensional network topology consisting of 3-links nodes;

b. generating addresses for the 3-links nodes using a naming scheme based on 2-tuple radix-n scheme to ensure a
10 unique address for each of the 3-links nodes in a specified network using a mathematical formula; and

c. interconnecting each 3-links node to an adjacent 3-links nodes linking via a point-to-point connection using mathematical criteria based on a numerical representation
15 for each 3-links node to form a network topology of a specified configuration.

2. The method according to claim 1 wherein the method of generating the address for each 3-links node in the network is based on a 2-tuple radix-n number, $X(MST, LST)$, where the MST is most significant tuple and LST is
20 the least significant tuple and the mathematical formula generating the 2-tuple number for each of the 3-links nodes in the specified network is based on the number of columns n and number of rows m of the specified network.

25 3. The method according to claim 1 wherein the



mathematical criteria determine the connection between any 2 different nodes in the network based on the numerical representation of the 3-links of upright and inverted nodes and a correct connection of the links between two nodes results in a sum of zero based on the mathematical criteria for each of the 3 links for both the upright and inverted nodes.

4. A method of finding a shortest path for routing a data message from a source node to a destination node in a network composed of 3-links nodes connected in a mesh comprising steps of:

introducing a generic routing algorithm to each 3-links node of the network, said generic routing algorithm determining a link for outputting the data message from a current node to an adjacent node so that a path from the source node to the destination node through traversal node becomes shortest;

performing the generic routing algorithm at each node to which the data message is transmitted;

outputting the data message via a link determined by the generic routing algorithm to an adjacent node; and

repeating the performing step and outputting step until the data message reaches the destination node;

thereby enabling to transmit the data message from the source node to the destination node through the



shortest path.

5. The method according to claim 4 wherein the generic routing algorithm comprises steps of:

a. computing row and column distances between
5 current and destination nodes based on addresses of these nodes in 2-tuple radix-n representation;

b. determining a point of input for the data message to be routed; and

c. determining an output of the data message based
10 on the type of node and computed shortest distance between the current node and destination node.

6. A method of configuring a network topology using 3-links nodes comprising steps of:

a. determining the numbers of rows and columns of
15 the network to be configured based on the total number of nodes in the network;

b. generating addresses for all nodes of the network;

c. interconnecting 3 links of the nodes in the
20 network using mathematical criteria said mathematical criteria determining the connection between any two different nodes in the network based on the numerical representation of three links of upright and inverted nodes and correct connection of the links between two nodes
25 results in a sum of zero based on the mathematical criteria



for each of the three links for the upright and inverted nodes.

7. A network topological method of reforming an existing network to a new network by adding two rows or
5 columns of nodes to the existing network to reduce the minimum expansion increment of the network configuration and thereby enabling to reduce a distance between any two nodes and improve the scalability of the network.