



(22) Date de dépôt/Filing Date: 1993/12/16

(41) Mise à la disp. pub./Open to Public Insp.: 1994/07/07

(45) Date de délivrance/Issue Date: 2010/06/29

(62) Demande originale/Original Application: 2 408 532

(30) Priorité/Priority: 1992/12/21 (US07/994,402)

(51) Cl.Int./Int.Cl. *G06F 13/36* (2006.01),
G06F 13/40 (2006.01)

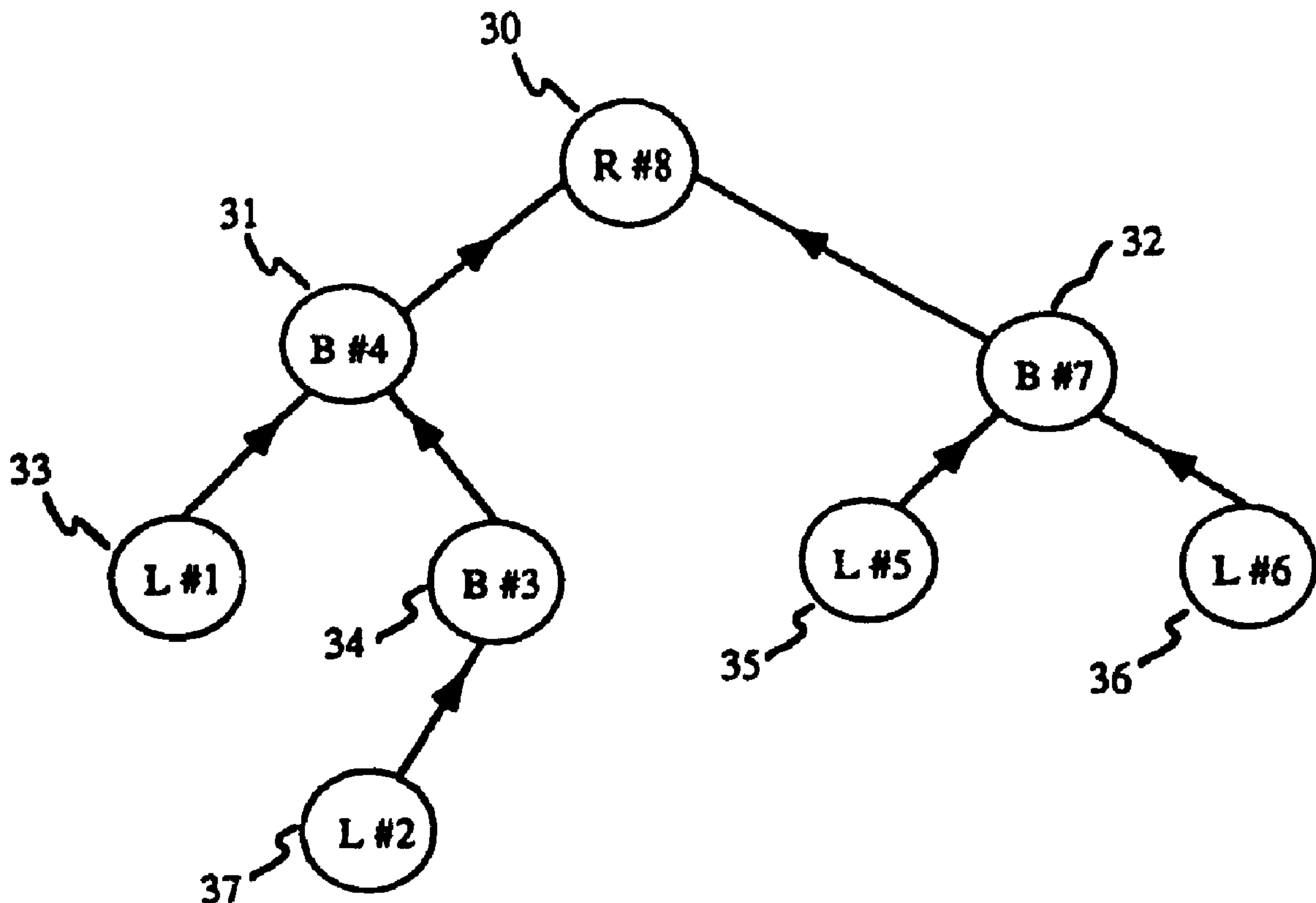
(72) Inventeur/Inventor:
OPRESCU, FLORIN, US

(73) Propriétaire/Owner:
APPLE INC., US

(74) Agent: RICHES, MCKENZIE & HERBERT LLP

(54) Titre : METHODE ET DISPOSITIF D'AFFECTATION D'ADRESSES UNIQUES, DE RECONNAISSANCE DE NOEUDS ET DE MAPPAGE TOPOLOGIQUE POUR GRAPHE ACYCLIQUE ORIENTE

(54) Title: METHOD AND APPARATUS FOR UNIQUE ADDRESS ASSIGNMENT, NODE SELF-IDENTIFICATION AND TOPOLOGY MAPPING FOR A DIRECTED ACYCLIC GRAPH



(57) Abrégé/Abstract:

A node identification system is described for use in a computer system in which the various components of the system are interconnected via nodes on a communications bus. Once the topology of the nodes has been resolved into an acyclic directed



(57) **Abrégé(suite)/Abstract(continued):**

graph, each node may be assigned a non-predetermined unique address. Each node having a plurality of ports has an a priori assigned priority for port selection. Each child node connected to a parent is allowed to respond in the predetermined sequence depending upon the port through which it is connected to its parent. Each node in the graph will announce its presence according to its location in the graph. Each receives an address incremented from the previous addresses assigned, thereby insuring uniqueness. The same mechanism may be implemented to allow each node in turn to broadcast information on the bus concerning the parameters of its local host. Likewise, additional information may be conveyed from each node concerning connections to other nodes thereby allowing a host system to generate a map of the resolved topology including any information about disabled links which may be used for redundancy purposes.

ABSTRACT

A node identification system is described for use in a computer system in which the various components of the system are interconnected via nodes on a communications bus. Once the topology of the nodes has been resolved into an acyclic directed graph, each node may be assigned a non-predetermined unique address. Each node having a plurality of ports has an a priori assigned priority for port selection. Each child node connected to a parent is allowed to respond in the predetermined sequence depending upon the port through which it is connected to its parent. Each node in the graph will announce its presence according to its location in the graph. Each receives an address incremented from the previous addresses assigned, thereby insuring uniqueness. The same mechanism may be implemented to allow each node in turn to broadcast information on the bus concerning the parameters of its local host. Likewise, additional information may be conveyed from each node concerning connections to other nodes thereby allowing a host system to generate a map of the resolved topology including any information about disabled links which may be used for redundancy purposes.

-1-

**METHOD AND APPARATUS FOR UNIQUE ADDRESS ASSIGNMENT,
NODE SELF-IDENTIFICATION AND TOPOLOGY MAPPING FOR A
DIRECTED ACYCLIC GRAPH**

This is a divisional application of Canadian patent application serial number 2,408,532, which in turn is a divisional application of Canadian Patent No. 2,151,368 which issued to patent on February 10, 2004 and was the national phase application of PCT International application PCT/US93/12314 filed December 16, 1993 (12.16.93) and published July 7, 1994 (07.07.94) under International publication no. WO 94/15303.

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates to computer systems. More particularly, the present invention relates to a method and apparatus for establishing and utilizing a

WO 94/15303

PCT/US93/12314

-2-

communications scheme between a plurality of arbitrarily assembled elements of a computer system.

Background

Components within a given computer system need the ability to convey signals amongst themselves. In very simple systems, it is possible to have each element of the system directly wired to all of the other parts of the system. However, in reality, in order to make computers expandable and to accommodate an unknown number of system parts, computer architects long ago developed the concept of a communications bus.

A bus is a communications path, such as a wire or wires, running throughout the computer system. Each component of the system need only plug into the bus to be theoretically connected to each of the other components in the system. Of course, each component cannot simultaneously communicate with other components because there may be only a single communications channel between the components. It is necessary when utilizing a communications bus to establish some form a sharing arrangement so that each component may use the bus to communicate with other components in an efficient manner

WO 94/15303

PCT/US93/12314

-3-

that does not leave critical pieces of information from one component hanging, waiting for bus access. The method by which components on the bus share the bus is generally referred to as a bus arbitration scheme.

In addition to the critical need to optimize the bus arbitration scheme so as to maximize the flow of important information, the physical (and logical/electrical) configuration of the bus itself can and should be optimized to minimize system delays while remaining as flexible as possible.

In order to communicate with other components attached to a bus, each component must be equipped with hardware such as transmitting and receiving circuitry compatible with the communications protocol implemented for the bus. One such communications standard is described in IEEE Standards Document P1394 entitled "High Performance Serial Bus", published October 14, 1992 in New York, New York, U.S.A. by the Institute of Electrical and Electronic Engineers, Inc. The standard described in P1394 is intended to provide a low cost interconnect between cards on the same backplane, cards on other backplanes, and external peripherals.

Prior art buses or networks required knowing what was being plugged in where. For example, the back of many

WO 94/15303

PCT/US93/12314

-4-

computers have specified ports for specific peripherals. Some computers implement several buses, such as the Macintosh which uses a bus referred to as ADB for components like a mouse and keyboard and SCSI bus for other peripherals. These types of buses provide for daisy chaining elements together but connections are of limited topology. Other known buses/networks require that the nodes of the network be arranged in a ring, a loop which must be closed in order to operate. Finally, star, or hub-and-spoke arrangements required that each node be directly linked to a central master. Each of the prior art systems lacks a desirable measure of flexibility.

It would be desirable, and is therefore an object of the present invention, to be able to arbitrarily assemble elements of a computer system onto a bus where the arbitrary topology can be resolved by the system into a functioning system without requiring a predetermined arrangement of components.

WO 94/15303

PCT/US93/12314

-5-

SUMMARY OF THE INVENTION

It is an object of the present invention to enhance the functionality of a computer system bus in which the nodes of the bus have been resolved into an acyclic directed graph.

It is also an object of the present invention to provide unique addresses to nodes on an acyclic directed graph.

It is another object of the present invention to provide a self identification mechanism for a collection of nodes on an acyclic directed graph.

It is still another object of the present invention to provide a mechanism for mapping the topology of a collection of nodes on a bus having an acyclic directed graph topology.

It is yet another object of the present invention to provide link redundancy information to a host system for an arbitrary collection of nodes resolved into an acyclic directed graph.

These and other objects of the present invention are implemented in a computer system in which the various

WO 94/15303

PCT/US93/12314

-6-

components of the system are interconnected via nodes on a communications bus. Once the topology of the nodes has been resolved into an acyclic directed graph, each node may be assigned a non-predetermined unique address. Each node having a plurality of ports has an apriori assigned priority for port selection. Each child node connected to a parent is allowed to respond in the predetermined sequence depending upon the port through which it is connected to its parent. Each node in the graph will announce its presence according to its location in the graph. Each receives an address incremented from the previous addresses assigned, thereby insuring uniqueness. In this way each node on the graph is assigned a unique address in a simple manner without any other differentiating characteristic required from the node than its topological location in the bus. The same mechanism may be implemented to allow each node in turn to broadcast information on the bus concerning the parameters of its local host. Likewise, additional information may be conveyed from each node concerning connections to other nodes thereby allowing a host system to generate a map of the resolved topology including any information about disabled links which may be used for redundancy purposes.

-6a-

Accordingly, in one aspect, the present invention provides a method for a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to the at least one adjacent nodes respectively, the method comprising: propagating a "You Are My Parent" (YAMP) signal from the first node to a second node of the at least one adjacent node; receiving, at the first node, a "You Are My Child" (YAMC) signal from the second node after said propagating the YAMP signal; establishing a parent-child relationship between the first node and the second node in response to said receiving the YAMC signal; and determining a round trip time for communication between the first node and the second node from a time gap between said propagating the YAMP signal and said receiving the YAMC signal.

In a further aspect, the present invention provides a method for a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to the at least one adjacent nodes respectively, the method comprising: propagating a first "You Are My Parent" (YAMP) signal from the first node to a second node of the at least one adjacent node; receiving, at the first node, a second YAMP signal from the second node after said propagating the first YAMP

-6b-

signal; determining a period in response to the second YAMP signal; propagating a third YAMP signal from the first node to the second node after expiration of the period if no YAMP signal from the second node is received at the first node during the period; receiving, at the first node, a first "You Are My Child" (YAMC) signal from the second node after said propagating the third YAMP signal; and establishing a parent-child relationship between the first node and the second node in response to the first YAMC signal.

In a still further aspect, the present invention provides a computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communication nodes and communications links comprising a bus of an acyclic directed graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the acyclic directed graph being designated branch nodes, all nodes initially having a status of being unidentified, said acyclic directed graph having established hierarchical parent-child relationships between all adjacent nodes

-6c-

proceeding from the root node down to any leaf nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other adjacent nodes, the root node being defined as having no parent node, to assign unique addresses to the nodes of the acyclic directed graph the computer system further comprising: means for each unidentified leaf node initially transmitting a "Bus Request" (BR) signal onto the bus; means for each branch node waiting until all unidentified adjacent child nodes each propagates a BR signal, then propagating a BR signal to a parent node; means for each node on the bus counting a number of address announcements that are broadcast on the bus; means for the root node waiting until all unidentified adjacent nodes each propagates a BR signal, then propagating a "Bus Grant" (BG) signal to one adjacent unidentified node based on a predetermined selection criterion of the root node for selecting adjacent nodes, said BG signal being propagated down through the acyclic directed graph through intervening nodes based on a predetermined selection criterion of each of the intervening nodes for selecting adjacent nodes until a node that initiated a BR signal receives the BG signal; means for the node receiving the

-6d-

BG signal then broadcasting an address announcement to all nodes on the bus; means for the node receiving the BG signal then setting a unique address to attaining a status of being identified, said unique address being a function of the number of address announcements counted at the node receiving the BG signal.

In a further aspect, the present invention provides a computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communications nodes and communications links comprising a bus of an acyclic directed graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the acyclic directed graph being designated branch nodes, all nodes initially having a status of being unidentified, said acyclic directed graph having established hierarchical parent-child relationships between all adjacent nodes proceeding from the root node down to any leaf nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other

-6e-

adjacent nodes, the root node being defined as having no parent node, to assign unique addresses to the nodes of the acyclic directed graph, the computer system further comprising: means for each unidentified leaf node initially transmitting a "Bus Request" (BR) signal onto the bus; means for each branch node waiting until all unidentified adjacent child nodes each propagates a BR signal, then propagating a BR signal to a parent node; means for the root node waiting until all unidentified adjacent nodes each propagates a BR signal, then propagating a "Bus Grant" (BG) signal to one adjacent unidentified node based on a predetermined selection criterion of the root node for selecting adjacent nodes, said BG signal being propagated down through the acyclic directed graph through intervening nodes based on a predetermined selection criterion of each of the intervening nodes for selecting adjacent nodes until a node that initiated a BR signal receives the BG signal; means for the node receiving the BG signal then broadcasting an address announcement to all nodes on the bus; means for the node receiving the BG signal then broadcasting, on the bus, information about a component associated with the node receiving the BG signal; means for each node on the bus counting, as a number, how many

-6f-

address announcements are broadcast by other nodes; means for each node setting a unique address as a function of the number of address announcements broadcast by other nodes until the node received a BG signal and subsequently broadcasts an address announcement, said node then attaining the status of being identified.

In a still further aspect, the present invention provides in a computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communications nodes and communications links comprising a bus of an acyclic directed graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the acyclic directed graph being designated branch nodes, all nodes initially having a status of being unidentified, said acyclic directed graph having established hierarchical parent-child relationships between all adjacent nodes proceeding from the root node down to any lead nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other

-6g-

adjacent nodes, the root node being defined as having no parent node, to assign unique addresses to the nodes of the acyclic directed graph the computer system further comprising: means for each unidentified leaf node initially transmitting a "Bus Request" (BR) signal onto the bus; means for each branch node waiting until all unidentified adjacent child nodes each propagates a BR signal, then propagating a BR signal to a parent node; means for each node on the bus counting a number of address announcements that are broadcast on the bus; means for the root node waiting until all unidentified adjacent nodes each propagates a BR signal, then propagating a "Bus Grant" (BG) signal to one adjacent unidentified node based on a predetermined selection criterion of the root node for selecting adjacent nodes, said BG signal being propagated down through the acyclic directed graph through intervening nodes based on a predetermined selection criterion of each of the intervening nodes for selecting adjacent nodes until a node that initiated a BR signal receives the BG signal; means for the node receiving the BG signal then broadcasting an address announcement to all nodes on the bus; means for the node receiving the BG signal then broadcasting, on the bus, topology information concerning the node receiving the BG signal; means for the

-6h-

node receiving the BG signal then setting a unique address to attaining a status of being identified, said unique address being a function of the number of address announcements counted at the node receiving the BG signal.

In a further aspect, the present invention provides in a computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communications nodes and communications links comprising a bus of a directed acyclic graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the directed acyclic graph being designated branch nodes, said directed acyclic graph having established hierarchical parent-child relationships between all adjacent nodes proceeding from the root node down to any leaf nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other adjacent nodes, the root node being defined as having no parent node, a first node connected to at least one adjacent node on the bus in a node identification process, the first node comprising:

-6i-

means for counting, since a starting point of a node identification process, a number of address announcements broadcast on the bus; when the first node obtains access to broadcast; means for setting a unique address for the first node, the unique address being as a function of the number of address announcements counted at the first node; and means for broadcasting an address announcement from the first node to all nodes on the bus; when the first node has at least one adjacent child node; means for receiving a first "Bus Request" (BR) signal from a second node, the second node being one of the at least one adjacent child node of the first node; when the first node has a parent node: means for forwarding the first BR signal to the parent node of the first node after receiving the first BR signal; means for forwarding a first "Bus Grant" (BG) signal for the first BR signal from the parent node of the first node to the second node; means for transmitting a second BR signal for the first node to the parent node of the first node; and means for receiving a second BG signal for the second BR signal from the parent node of the first node; when the first node has no parent node: means for transmitting a third BG signal for the first BR signal to the second node; and means for granting the first node access to broadcast.

-6j-

In a still further aspect, the present invention provides a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to the at least one adjacent node respectively, the first node comprising: means for counting, since a starting point of a node identification process, a number of announcements received at the first node; means for broadcasting an announcement from the first node when the first node obtains access to broadcast; and means for setting a unique address for the first node based on the number of announcements received by the first node since the starting point to when the first node obtains the access to broadcast.

In a further aspect, the present invention provides a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to the at least one adjacent nodes respectively, the first node comprising: means for propagating a "You Are My Parent" (YAMP) signal from the first node to a second node of the at least one adjacent node; means for receiving, at the first node, a "You Are My Child" (YAMC) signal from the second node after propagating the YAMP signal; means for establishing a parent-child relationship between the first node and the

-6k-

second node in response to receiving the YAMC signal; and means for determining a round trip time for communication between the first node and the second node from a time gap between propagating the YAMP signal and receiving the YAMC signal.

In a still further aspect, the present invention provides a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to the at least one adjacent nodes respectively, the first node comprising: means for receiving, at the first node, a "You Are My Parent" (YAMP) signal from a second node of the at least one adjacent node; means for propagating a "You Are My Child" (YAMC) signal from the first node to the second node in response to receiving the YAMP signal; means for establishing a parent-child relationship between the first node and the second node in response to receiving the YAMP signal; means for receiving, at the first node, a "You Are My Child Acknowledged" (YAMCA) signal from the second node after propagating the YAMC signal; and means for determining a round trip time for communication between the first node and the second node from a time gap between propagating the YAMC signal and receiving the YAMCA signal.

-61-

In a further aspect, the present invention provides a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to the at least one adjacent nodes respectively, the first node comprising: means for propagating a first "You Are My Parent" (YAMP) signal from the first node to a second node of the at least one adjacent node; means for receiving, at the first node, a second YAMP signal from the second node after propagating the first YAMP signal; means for determining a period in response to the second YAMP signal; means for propagating a third YAMP signal from the first node to the second node after expiration of the period if no YAMP signal from the second node is received at the first node during the period; means for receiving, at the first node, a first "You Are My Child" (YAMC) signal from the second node after propagating the third YAMP signal; and means for establishing a parent-child relationship between the first node and the second node in response to the first YAMC signal.

In a still further aspect, the present invention provides a machine readable medium containing executable computer program instructions which when executed by a data processing system cause said system to perform a

-6m-

method performed by a first node connected to at least one adjacent node on a bus in a node identification process in a computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communications nodes and communications links comprising the bus of a directed acyclic graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the directed acyclic graph being designated branch nodes, said directed acyclic graph having established hierarchical parent-child relationships between all adjacent nodes proceeding from the root node down to any leaf nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other adjacent nodes, the root node being defined as having no parent node, the method comprising: counting, since a starting point of a node identification process, a number of address announcements broadcast on the bus; when the first node obtains access to broadcast; setting a unique address for the first node, the unique address being as a function of

-6n-

the number of address announcements counted at the first node; and broadcasting an address announcement from the first node to all nodes on the bus; when the first node has at least one adjacent child node: receiving a first "Bus Request" (BR) signal from a second node, the second node being one of the at least one adjacent child node of the first node; when the first node has a parent node: forwarding the first BR signal to the parent node of the first node after receiving the first BR signal; forwarding a first "Bus Grant" (BG) signal for the first BR signal from the parent node of the first node to the second node; transmitting a second BR signal for the first node to the parent node of the first node; and receiving a second BG signal for the second BR signal from the parent node of the first node; when the first node has no parent node: transmitting a third BG signal for the first BR signal to the second node; and granting the first node access to broadcast.

In a further aspect, the present invention provides a machine readable medium containing executable computer program instructions which when executed by a data processing system cause said system to perform a method performed by a first node connected to at least one adjacent node on a bus in a node identification process in

-60-

a computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communications nodes and communications links comprising the bus of a directed acyclic graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the directed acyclic graph being designated branch nodes, said directed acyclic graph having established hierarchical parent-child relationships between all adjacent nodes proceeding from the root node down to any leaf nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other adjacent nodes, the root node being defined as having no parent node, the method comprising: counting, since a starting point of a node identification process, a number of address announcements broadcast on the bus; when the first node obtains access to broadcast; setting a unique address for the first node, the unique address being as a function of the number of address announcements counted at the first node; and broadcasting an address announcement from the

-6p-

first node to all nodes on the bus; when the first node has at least one adjacent child node: receiving a first "Bus Request" (BR) signal from a second node, the second node being one of the at least one adjacent child node of the first node; when the first node has a parent node: forwarding the first BR signal to the parent node of the first node after receiving the first BR signal; forwarding a first "Bus Grant" (BG) signal for the first BR signal from the parent node of the first node to the second node; transmitting a second BR signal for the first node to the parent node of the first node; and receiving a second BG signal for the second BR signal from the parent node of the first node; when the first node has no parent node: transmitting a third BG signal for the first BR signal to the second node; and granting the first node access to broadcast.

In still a further aspect, the present invention provides a machine readable medium containing executable computer program instructions which when executed by a data processing system cause said system to perform a method for a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to the at least one adjacent nodes respectively, the method comprising: propagating a

-6q-

"You Are My Parent" (YAMP) signal from the first node to a second node of the at least one adjacent node; receiving, at the first node, a "You Are My Child" (YAMC) signal from the second node after said propagating the YAMP signal; establishing a parent-child relationship between the first node and the second node in response to said receiving the YAMC signal; and determining a round trip time for communication between the first node and the second node from a time gap between said propagating the YAMP signal and said receiving the YAMC signal.

In a further aspect, the present invention provides a machine readable medium containing executable computer program instructions which when executed by a data processing system cause said system to perform a method for a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to the at least one adjacent nodes respectively, the method comprising: receiving, at the first node, a "You Are My Parent" (YAMP) signal from a second node of the at least one adjacent node; propagating a "You Are My Child" (YAMC) signal from the first node to the second node in response to said receiving the YAMP signal; establishing a parent-child relationship between the first node and the second node in response to said

-6r-

receiving the YAMP signal; receiving, at the first node, a "You Are My Child Acknowledged" (YAMCA) signal from the second node after said propagating the YAMC signal; determining a round trip time for communication between the first node and the second node from a time gap between said propagating the YAMC signal and said receiving the YAMCA signal.

In a still further aspect, the present invention provides a machine readable medium containing executable computer program instructions which when executed by a data processing system cause said system to perform a method for a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to the at least one adjacent nodes respectively, the method comprising: propagating a first "You Are My Parent" (YAMP) signal from the first node to a second node of the at least one adjacent node; receiving, at the first node, a second YAMP signal from the second node after said propagating the first YAMP signal; determining a period in response to the second YAMP signal; propagating a third YAMP signal from the first node to the second node after expiration of the period if no YAMP signal from the second node is received at the first node during the period; receiving, at the

-6s-

first node, a first "You Are My Child" (YAMC) signal from the second node after said propagating the third YAMP signal; and establishing a parent-child relationship between the first node and the second node in response to the first YAMC signal.

In still a further aspect, the present invention resides in a machine readable medium containing executable computer program instructions which when executed by a data processing system cause said system to perform a method for a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to that at least one adjacent node respectfully, the method comprising: counting, since a starting point of a node identification process, a number of announcements received at the first node; broadcasting an announcement from the first node when the first node obtains access to broadcast; and setting a unique address for the first node based on the number of announcements received by the first node since the starting point to when the first node obtains the access to broadcast.

WO 94/15303

PCT/US93/12314

-7-

BRIEF DESCRIPTION OF THE DRAWINGS

The objects, features and advantages of the present invention will be apparent from the following detailed description in which:

Figure 1 illustrates a block diagram of the hardware layer implementation utilized in accordance with the present invention.

Figures 2(a)-2(b) illustrate arbitrarily assembled collection of nodes, one being acyclic and the other including multiple cycles.

Figure 3(a) is the arbitrarily assembled collection of nodes of **Figure 2(a)** undergoing the graph transformation process in accordance with the present invention.

Figures 3(b)-3(d) illustrate alternative communications exchanges between nodes in implementing the present invention.

Figure 3(e) graphically illustrates the directed graph resulting from the arbitrarily assembled network of nodes of **Figure 2(a)**.

Figure 4 illustrates a symmetrical graph arrangement which requires resolving a root contention.

WO 94/15303

PCT/US93/12314

-8-

Figure 5 illustrates a directed acyclic graph with a possible unique address assignment order indicated.

Figure 6(a)-6(e) illustrate the process flow for carrying out the graph transformation procedure in accordance with the preferred embodiment of the present invention.

WO 94/15303

PCT/US93/12314

-9-

DETAILED DESCRIPTION OF THE INVENTION

A method and apparatus for utilizing a bus having an arbitrary topology are described. In the following description, many specific details are set forth such as various computer components in order to provide a thorough understanding of the present invention. It will be obvious, however, to one skilled in the art that the present invention may be practiced without such specific details. In other instances, well-known control structures and coding techniques have not been described in detail in order not to obscure unnecessarily the present invention.

Throughout this detailed description, numerous descriptive terms are introduced to provide metaphorical clarity to the description. For example, frequent references will be made to parent-child relationships between nodes in a given topology. The purpose of this is to provide the concept of "direction" to the finally resolved graph. As will be described, once an arbitrary topology has been reduced to an acyclic directed graph, there will be one node identified as the "root" node. The root node will not have a parent node, all nodes logically immediately adjacent to the root node are the child nodes

WO 94/15303

PCT/US93/12314

-10-

of the root. The "tree" metaphor is completed by the inclusion of nodes referred to as "branches" and "leaves".

The bus architecture described herein, though described with reference to components for a single computer, in general has a broader scope. The present invention for defining the bus topology may be applied to any arbitrarily assembled collection of nodes linked together as in a network of devices. One point that must be noted is that it is necessary to distinguish a node from a physical computer component. Each component to reside on the bus will have associated with it at least one node physical layer controller. In certain circumstance, a given component may advantageously be associated with multiple nodes but in the usual case there will be a one-to-one correspondence between devices or components on the bus and nodes.

Referring now to Figure 1 a block diagram of a node 10 is illustrated. The physical implementation of a node is somewhat arbitrary. In the preferred embodiment implementation of the present invention, the nodes are designed to comply with the IEEE P1394 High Performance Serial Bus communications protocol which was published October 14, 1992 in New York, New York, U.S.A. by the Institute of Electrical and Electronic Engineers, Inc. The node 10 includes arbitration state machine logic 11. This Arbitration state machine logic

WO 94/15303

PCT/US93/12314

-11-

incorporates all the logic circuitry for carrying out the methodologies and algorithms to be described herein. The circuitry may comprise a programmable logic array (PLA) or be uniquely designed to carry out the functions described herein. Those skilled in the art, once described the functions to be carried out by the node logic will be able to implement the present invention without undue experimentation. The node, by means of its logic, shall implement the minimum arbitration protocol including the bus initialization, tree identification, self identification, and the bus arbitration functions, all to be described in detail further herein.

The node 10 shown in **Figure 1** also includes transmission multiplexers 12 and 13 and data transmitter, receiver and resynchronizer 14. The node illustrated in **Figure 1** is coupled to local host 15. Local host 15 may be any device one wishes to attach to the bus such as a disk drive, CPU, keyboard or any other component which needs to communicate with other components in the system. The node 10 communicates with other nodes through communications links. A link is a connection between two ports and in immediate practical terms is a cable segment but in general it may be implemented as any physical communication channel. A link shall be able, at minimum,

WO 94/15303

PCT/US93/12314

-12-

to provide a half duplex communication channel between the two ports which it connects. A port is the interface between a node and a link. In accordance with the present invention, a port must have the ability to transmit and receive data and arbitration signaling. A port needs to be able to determine whether or not it is connected to another port through a link. One method of facilitating this is by having connected ports apply a biasing voltage through the link which is detectable by the port at the other end of the link. Thus, if a port has a link attached which is not connected to a port at the other end, a naked link, the port will determine that it is not a connected port. In **Figure 1**, the illustrated node 10 has three external ports 21, 22 and 23 with connecting links 17, 18 and 19, respectively.

Some of the rules of implementation for nodes in order to implement the present invention are that a node may have one or more ports. A node shall be able to transmit and receive data on any one of its ports. A node shall be able to receive data on one and only one of its enabled ports at a time and be able to retransmit this data on all remaining enabled ports. A node shall be able to receive and transmit signaling messages through all of its ports simultaneously and independently. Separate signaling

WO 94/15303

PCT/US93/12314

-13-

transceivers, encoders and decoders are required for each port of a node. A minimum implementation node does not require a local host device. For example, such a node may function as a cable extension. From hereon devices and local hosts will be ignored and all references to bus topology will refer to nodes and node connections through various ports.

Graph Transformation

Figures 2(a) and 2(b) illustrate arbitrarily assembled collections of nodes. From hereon, nodes will be illustrated merely as circles, but are deemed to each incorporate elements equivalent to those described with respect to Figure 1. Note, however, that each node may have more or less than the three external ports shown in that figure. The lines illustrated connecting each of the nodes are the method by which links are shown. Ports are not illustrated, but are impliedly the interface where a link and a node connect.

The bus arbitration methodology to be described herein requires that the arbitrary topology be resolved into an acyclic directed graph. In an arbitrary topology graph a collection of nodes and links may form a cycle. A

WO 94/15303

PCT/US93/12314

-14-

cycle exists if starting from a specific node in the graph it is possible to return to the same node by traversing links and nodes without any link being traversed twice.

Figure 2(a) illustrates an acyclic graph because none of the nodes illustrated are connected within a loop. **Figure 2(b)**, however, is not an acyclic graph because the region in bounding box 25 contains a collection of nodes, 40-47 which form multiple cycles. The bus arbitration methodology to be described requires that there be no cycles, so a method of user intervention to resolve cycles will also be described further herein.

In addition to the requirement that a graph be acyclic, it must also be directed. A directed graph is one in which a hierarchical arrangement has been established between adjacent nodes. Initially, there are no established parent-child relationships between nodes. That is, for example, node 31 may be the "parent node" for node 34, or be the "child node" for node 34. Thus, it is necessary to take a given arbitrary topology graph and transform it into an acyclic and directed graph. The methods described herein will work to perform this transformation for any give arbitrary topology, regardless of the number of nodes or how they are physically linked

WO 94/15303

PCT/US93/12314

-15-

and regardless of the signaling propagation time along the links.

Node Communication

Initially, the process of transforming an acyclic arbitrary topology graph into a directed graph will be described. The case where cycle resolution is required will follow. **Figure 3(a)** shows the arbitrary graph of **Figure 2(a)** wherein the nodes and links have status labels and communicated signals are indicated for the graph transformation process for directing a graph. It is instructive at this point to describe signal communications between nodes. **Figure 3(b)** illustrates two nodes 50 and 51 (hereinafter node A and node B, respectively) coupled by link 52. As described, the link is the communications channel coupling transceiver ports of the respective nodes as described above with reference to **Figure 1**. During the graph transformation process, it becomes necessary for nodes to establish parent-child relationships with adjacent nodes. Two nodes are said to be adjacent nodes if there is at least one link connected between a port of the first node and a port of the second node. In **Figures 3(b)-3(d)** it will be assumed that the relationship to be

WO 94/15303

PCT/US93/12314

-16-

resolved is that node B is the parent of node A and that it is appropriate for the nodes to establish that relationship.

Prior to a direction being established, when it become appropriate for node A to establish node B as its parent, node A will transmit from its port to which link 52 is coupled the signal "You Are My Parent" (YAMP). This message content may take any form, so long as node A knows that it is signaling YAMP and node B is capable of understanding that the received message is YAMP. When YAMP signal 53 is received by node B, node B will respond to node A by sending "You Are My Child" (YAMC) through link 52 to node A. The arbitration state machine logic 11 of node A will keep track of the time delay between sending YAMP signal 53 and receiving YAMC signal 54. The time measured signifies twice the propagation delay between nodes A and B. Upon receiving the YAMC signal, node A will respond with a "You Are My Child Acknowledged" (YAMCA) signal 55. This provides node B with the ability to also determine the propagation time delay between the nodes equal to the time delay between sending YAMC and receiving YAMCA. For half duplex communication links, the YAMCA message also has the effect of properly orienting the communication channel.

WO 94/15303

PCT/US93/12314

-17-

For full duplex communications channels the three logical messages, YAMP, YAMC and YAMCA can alternatively be relayed by only two signal transmissions. In Figure 3(c) this situation is illustrated where node A asserts the YAMP signal 56 continuously until it receives the return YAMC signal 57. The YAMCA signal is logically transmitted to node B when the YAMP signal is detected as no longer arriving.

The use of this described triple asynchronous message exchange provides a mechanism by which both nodes involved in the message exchange can determine the propagation time delay through the link. This delay value is used in resolving contention events to be described further herein as well as during normal bus arbitration to optimize bus performance. The dynamic extraction of this parameter is not mandatory. As an alternative a maximum propagation time delay can be apriory defined at the expense of optimum bus performance.

Once nodes A and B have exchanged messages signifying that node B is the parent of node A, the link can be said to be directed. Node A within its logic labels its port to which link 52 is coupled as a parent port (it talks to a parent node) and node B labels its port to which link 52 is coupled a child port (it talks to a child node). It is

WO 94/15303

PCT/US93/12314

-18-

important to maintain the labels that ports achieve because the methods to be described below will be in terms of the labels assigned to nodes and ports at a given time. A short hand graphical notation is illustrated in **Figure 3(d)** where the direction arrow 58 indicates that node B is established as the parent of node A and the link is directed.

Direction Determination

Referring back now to **Figure 3(a)** and to process **Figures 6(a)-6(e)**, the process of directing the overall arbitrary topology will now be described. It is necessary to introduce a few more colorful definitions to aid in explaining the topology transformation process. First, a "leaf" node is defined as a node with only one connected port. A node recognizes its status as a leaf node as soon as it is initialized after power-up or other bus initialization. A "branch" node is a node which has a least two connected ports. Through all but one of the connected ports a branch node will have received the YAMP signal and have acknowledged it. Through its remaining port, a branch node has sent the YAMP signal thus establishing that it has a parent node. A node does not

WO 94/15303

PCT/US93/12314

-19-

achieve branch status until it has established that it has one parent (a node can have only one parent node) and all its other ports are connected to child nodes. Prior to achieving branch status, a node is considered a "cycle" node because until it is determined to be a branch the possibility exists that the node is part of a cycle which makes establishing direction impossible.

The graph transformation procedure begins at step 60 upon bus initialization (power-up or instigated) at which time the leaf nodes in the arbitrary topology recognize at step 61 and label themselves at step 68 as leaf nodes by determining that they have only one connected port at decision box 66. In the graph depicted in **Figure 3(a)**, nodes 33, 35, 36 and 37 are leaf nodes which, once initialized, at step 69 each transmits the YAMP signal through its only connected port to its adjacent node. The nodes receiving these signals will then propagate the YAMC signals back to the leaf nodes at step 70, thus establishing a direction for the given link between respective parent-child pairs when the YAMCA communication is completed. At step 71 each leaf node labels its one connected port as a parent port and each receiving port on the parent node is labeled a child port.

WO 94/15303

PCT/US93/12314

-20-

The nodes on the graph which are not initially leaf nodes are initially considered "cycle nodes" for the reason described above and proceed according to the Cycle Node Procedure 63. Any cycle node which has labeled all but one of its connected ports as child ports then propagates the YAMP signal from its remaining unlabeled port at step 85. When that direction is established for the link, the cycle node then becomes labeled a branch node. Thus, after leaf node 37 establishes that node 34 is its parent, node 34 has only one unlabeled port (having labeled the link connection to node 37 as being through a child port) so node 34 broadcasts the YAMP signal to node 31, resulting in node 34 becoming a branch node. Likewise, once node 31 has identified that nodes 33 and 34 are its children, node 31 broadcasts the YAMP signal to node 30. When one node has received through all of its ports the YAMP signal at decision box 75, that node becomes the root node. In **Figure 3(a)** after node 30 has received the YAMP signals from nodes 31 and 32, its label changes from being a cycle node to being the root node. In the graph of **Figure 3(a)**, it is not necessarily the case that node 30 would become the root. If some of the links in the tree provided long propagation delays, node 30 might have received a YAMP signal on one port and then transmitted a YAMP signal through its other port. Any of the nodes may become the

WO 94/15303

PCT/US93/12314

-21-

root, even a leaf, the root property taking precedence. **Figure 3(e)** shows the resulting directed graph in response to the communicated signals shown in **Figure 3(a)** with each node labeled and the directions indicated by dark arrows.

Root Contention

In certain circumstances a root contention situation may arise. This may happen, for example in the case where the arbitrary topology has a symmetrical arrangement to it such as that shown in **Figure 4**. In the arbitrary graph illustrated in **Figure 4**, nodes 160 and 161 have each established that it is a parent to the two leaf nodes to which it is coupled. Then, each has propagated the YAMP signal to the other at nearly the same time. The root contention situation is recognized by both nodes involved at decision box 86. Each node is receiving a signal which designates it as a parent while it has sent the same signal out through the same port. Each of the contending nodes responds to the other with the YAMC signal at step 91 which allows each to determine the "decision time period" which is equal to twice the propagation time between the nodes.

WO 94/15303

PCT/US93/12314

-22-

The root contention situation is resolved by utilizing a random decision mechanism incorporated on each arbitration state machine logic unit 11 of each node. For every "decision time period" that elapses, each node will randomly decide at step 92 (with a 50% probability) whether to again transmit the YAMP signal to the other. Almost certainly within a finite number of the cycles, one node will decide to designate the other its parent without that one reciprocating. The one that is designated the parent becomes the root at step 95. Alternatively, predetermined selection criteria values may be assigned to nodes, the larger or smaller determining which dominates in a contention event. The dynamic determination of the "decision time period", while it offers optimum performance is not essential in implementing the present invention. As an alternative an apriory defined "decision time period" may be used as long as it is greater than the worst case link propagation that can be encountered in any bus using this algorithm. The same method used to resolve root contentions will also be used to resolve other contention events to be described further herein.

Root Assignment

WO 94/15303

PCT/US93/12314

-23-

As described above, the result of the graph transformation process is the assignment of the root attribute to one and only one node in the graph. The root node will have the ultimate decision in the bus arbitration scheme to be described and therefore can access the bus with maximum priority without the use of special priority time gaps. It is often desirable to be able to assign the root property to a predetermined node either when it is manufactured or dynamically (during run time) to optimize a given system. A given bus may include a node which requires isochronous data transfer. Isochronous data is data that must be transmitted at a given time to be of any value. For example, music from a compact disk needs to be transferred and output in the order in which it is to be heard and with no significant delays, unlike data files which may be transferred piecemeal and not necessarily in order.

Nodes can be classified into three categories with respect to root designations. These designations may be applied during manufacturing by hard-wiring the designation into the device, programming the arbitration state machine logic or by higher level software making the decision then initiating a reboot while preserving that decision. The three designations that a node may be assigned with respect

WO 94/15303

PCT/US93/12314

-24-

to being designated a root are: nodes that do not want to be root, nodes that may (should) be root and nodes that shall be root. These designations are tested for at steps 81 and 83. A node designated in the first category will begin the graph transformation procedure as soon as it is directed to do so. This will usually be immediately following the completion of the bus initialization procedure. A node from the second category will delay the beginning of the graph transformation procedure for a predetermined amount of time after it is directed to begin the procedure at step 84. By this delay, the node increases its chance of becoming the root. (The YAMP signals are more likely to propagate to it due to the delay.) Despite the added delay, it is still possible that a "may be root" node will not wind up being designated the root. This will depend on the given topology and message propagation delays. The amount of delay can be defined during design to be greater than a reasonable worst case propagation delay through a fairly complex graph.

A node from the third category of root designation possibilities may only recognize the fact that it must be the root after the graph has already been transformed and all nodes have identified themselves. The arbitration state machine logic may make this determination or software

WO 94/15303

PCT/US93/12314

-25-

running on the host system. When this occurs, the node that has to be root agrees with all other nodes on the bus that it is going to be the one and only root and restarts the graph transformation process by signaling a preemptive bus initialization signal which is described further herein. The node then waits at step 82 to become the root and does not participate in the graph transformation until it has received the YAMP signal on all of its ports, thus forcing it to be designated the root.

Once the root has been determined, the graph can be said to be directed. There is a defined relationship existing between all adjacent nodes on the graph.

Cycle Resolution

The procedures described above for directing a graph will only work for an acyclic graph. If there are cycles in the arbitrary topology, they must be broken by the procedure beginning at step 80. The existence of a cycle is detected at step 79 when, after a predetermined time-out period has elapsed, a node is still labeled a cycle node rather than a leaf, branch or root. The "cycle detect" timing starts immediately after the end of the bus initialization function. The time-out period need be no

WO 94/15303

PCT/US93/12314

-26-

longer than the worst case duration of the graph transformation process (adding in delay time for a "may be root" node and a possible root contention event).

The "cycle detect" time-out event does not have to occur synchronous for all nodes of a graph as all message exchanges are asynchronous events. As such, it is possible for a node which has not yet reached its "cycle detect" time-out event to receive a message indicating that cycle resolution is ongoing. Such a node will terminate its cycle detect time-out interval and begin the appropriate cycle resolution process.

The method of cycle resolution in accordance with the present invention requires the user of the assembled collection of nodes to intervene. When a node encounters the "cycle detect" time out the system user may be notified at step 100 of **Figure 6(a)** through an output device that a cycle exists and which nodes are then involved. The user will then be instructed to disconnect links to eliminate whatever cycles are present. The user will then return control to the graph transformation procedure.

Once each of the loops is broken, and no cycles remain, the procedure for transforming the graph as

WO 94/15303

PCT/US93/12314

-27-

described in the earlier sections may proceed until the entire graph is both acyclic and directed.

Unique Physical Address Assignment

Once a directed acyclic graph has been established from the original arbitrary topology, it is then possible to assign unique physical addresses to each node on the graph. This process begins with all leaf nodes requesting the bus by transmitting through their single connected ports the Bus Request (BR) signal. The parent node receiving the signal will wait until it has received the BR signal from all of its child ports and then will propagate the BR signal to its parent. The BR signals will propagate through the graph until the root has received the BR signal from all of its children. Once the root has received a bus request through all of its child ports, it will make a decision for granting the bus through one port and propagating a Bus Denial (BD) signal through its remaining child ports. The method for selecting which bus request to grant may be an apriory decision such as that described above where, for example, ports are selected from left to right or based on port numbering, etc. The Bus Grant (BG) signal will be transmitted from the root to its requesting

WO 94/15303

PCT/US93/12314

-28-

child. If that requesting child is itself a parent node which has propagated the bus request from one of its children, it will send the bus denial signal through all but one of its child ports in the same predetermined fashion as described above. Eventually one leaf node will receive the bus grant signal to which it will reply with a Bus Grant Acknowledged (BGA) signal which will be propagated back to the root node. The propagation of the BD and BGA signals serve to orient the communication links which may be necessary for the case of half duplex communications channels. All of the denied nodes will then wait for activity by the node which finally receives the BG signal.

The node which is finally granted access to the bus will transmit an address assignment packet. The node will transmit this packet on the bus and it will be received by all other nodes, each of which will count the number of address packets they receive. The transmitted address packet may have any arbitrary information. A node's unique physical address will be based on the number of address packets a node has counted before it transmits an address packet. Thus, no two nodes will acquire the same physical address despite not having address information assigned in advance. The actual composition of the address packet is

WO 94/15303

PCT/US93/12314

-29-

arbitrary and may be any bit stream efficiently utilizable by the system. After transmitting the physical address assignment packet, a node will transmit a "Child ID Completed" signal (CIC) signal. The parent node receiving this on its child port will then transmit the "Child Identification Completed Acknowledgment" (CICA) signal and label the port as an identified child port. In response to the next BR signal propagation, the parent of the node which has just identified itself will then select its next child to transmit the physical address packet. Once all the child nodes of a parent node have identified themselves, the parent node will request the bus and, when granted the bus, will propagate its physical address assignment packet. This procedure will continue following the predetermined selection criteria until all nodes determine a unique physical address assignment by counting. **Figure 5** illustrates the graph of **Figure 3(e)** in which a left-to-right predefined selection criteria is implemented. The nodes are uniquely assigned addresses where node 33 receives the first and as described, the root node 30 receives the eighth and last address.

When this procedure is completed, each node in the graph will have a unique physical address, which need not

WO 94/15303

PCT/US93/12314

-30-

have been determined in advance and which may be utilized for system management or other purposes.

Node Self-Identification

The process of node self identification essentially follows the same routine as the physical address assignment procedure described above. As each node transmits its physical address assignment packet, that packet may include further information such as identification of the physical device comprising the identification of the local host related to the node, how much power it requires, and, for example, whether it supports a "soft power-on" attribute, etc. In fact, the node self-identification information may serve as the physical address assignment packet because the practice of sending any information at all provides the basis for counting to yield unique physical addresses.

With respect to the node self-identification packet, the particular information concerning the node need only be "listened" to by those nodes affected by the nature of the announcing node. This procedure, as with the above, proceeds until all nodes have transmitted their node self identification information.

WO 94/15303

PCT/US93/12314

-31-

Topology Mapping

The method of topology mapping follows along the same lines as physical address assignment and node self-identification. This procedure thus has each node, when it is going through the address assignment or node self-identification process, further transmit information concerning all of its ports such as the number of child ports it has and whether or not it has any disabled ports. With respect to disabled ports, it may be desirable to implement a communication protocol between ports that are disabling so that they can identify from whom they are disabled. Thus, when a port identifies a disabled port it will give an identifier indicating its own ID as well as the port ID from which it has become disabled.

By assembling all the topology information about all the ports received during the topology mapping procedure, the bus server, host or any software level application may logically reconstruct the resolved bus topology. This is useful for many purposes including implementing redundancy where if a link unexpectedly goes down, previously disabled links may serve to prevent the loss of communication channels to any nodes.

WO 94/15303

PCT/US93/12314

-32-

Fair Bus Access Arbitration

Once the topology mapping, node self identification or physical address assignment routines have completed, the bus can be considered up and running. One arbitration scheme implemented in accordance with the present invention is that of fair bus access. When a node desires access to the bus, it sends through its parent port (unless it is the root) a bus request (BR) signal. The parent, upon receiving the BR signal from one child sends a bus denied signal (BD) through all its other child ports. The parent then propagates the BR signal upward through its parent until the signal reaches the root. The root issues a bus grant signal (BG) responsive to the first BR signal it receives and sends the BD signal through all of its other child ports which propagate downward thereby orienting the links. The BG signal propagates downward through the graph until it reaches the requesting node which then sends Bus Acknowledge (BA) signal followed by the packet of information that the node needed to send on the bus. When the packet is completed, all nodes return or enter into an idle state.

WO 94/15303

PCT/US93/12314

-33-

In the case where the root receives nearly simultaneous requests for the bus, the predetermined selection criteria for the root node will be used for granting to one of the nodes bus access. This may be the same predetermined priority selecting criteria as described above.

A further aspect of fair bus access arbitration is that a parent node has priority over its children. Thus, when a parent node wants the bus, it sends the BD signal through all of its child ports, then propagates the BR signal up toward the root. One potential problem with this mechanism is that if the parent has a large quantity of information to transmit on the bus a child node may have trouble getting adequate bus access. There is therefore introduced a gap system which is widely used and well-known in the art. After a node has utilized the bus, the node must wait for one gap period before it can again request the bus. This gives equal chance of being granted the bus to every node on the bus independent of its topological placement on the bus. In order to guarantee a fair arbitration protocol the length of the gap has to be greater than the worst case signal propagation delay through the bus. The gap value can be predetermined and hard-wired into the node logic but such an approach will

WO 94/15303

PCT/US93/12314

-34-

result in all but the most extreme case in a sub optimal utilization of the bus. The topology mapping capability together with the measurement of the propagation delay between adjacent nodes performed during the graph transformation phase enables the calculation of an optimal fair gap that will maximize the bus performance for any specific implementation.

Priority Bus Arbitration

In the bus arbitration scheme implemented in accordance with the above fair bus access arbitration, it may be desirable that the root always have bus priority. When this is implemented, the root node may grant the bus to itself at any time. This is done by first sending the BD signal down through all of the nodes in the graph. Priority bus access for the root is very useful for the case where the root node is required to perform isochronous data transfer.

Token Passing Bus Arbitration

As an alternative to the fair and priority bus access arbitrations schemes described above, the present invention

WO 94/15303

PCT/US93/12314

-35-

may be utilized in implementing a token passing bus arbitration scheme. Metaphorically speaking, token passing bus access refers to the notion that a node may communicate on a bus when it is in possession of a token that is passed between nodes. The token is passed from node to node in a cyclic fashion so that each node receives the bus in a predetermined point in the cycle. Token passing is implemented in the present invention in following the same manner as the physical address assignment routine described above. The predetermined selection mechanisms implemented are used to select the order in which the token will be passed from node to node. This order resembles the order as shown in **Figure 5** which dictates the order of unique address assignment. Each node, when it is assigned the token will propagate its information packet on the bus while the remaining nodes listen. The node will then pass the token to the next logical node based on the predetermined sequencing method as described above.

Preemptive Bus Initialization

An important feature that may be implemented in accordance with the present invention is the notion of preemptive bus initialization. The state machine logic

WO 94/15303

PCT/US93/12314

-36-

incorporated on each node is capable of triggering a bus initialization (BI) signal to be propagated from the node through all of its ports upon certain conditions. When a node has determined it is necessary to signal a bus initialization condition, it will propagate the BI signal out through all of its ports for a length of time sufficient to guarantee that all adjacent nodes have received it and then released. A node will then go into the initiating procedures which then lead to the graph transformation process in the above described procedures.

There are a number of situations which may make it necessary or desirable to trigger a preemptive bus initialization. First, this may be a node response to an unforeseen error. Additionally, at the host level, it may be determined that a different node should acquire the root attribute, for example, an isochronous data transfer node. This assignment will be preserved throughout the bus initialization routine thereby causing the desired node to wait during the transform procedure until it receives the root designation. Another condition leading to a preemptive bus initialization may be the breakage of a link, in which case it may be necessary to calculate a new acyclic directed graph for the attached nodes. Finally, an important situation in which a preemptive bus

WO 94/15303

PCT/US93/12314

-37-

initialization should occur is when a device is added to the network, referred to as "hot addition" of peripherals. The port to which a new device is connected will detect the presence of a new node and trigger a bus initialization which will be transparent to the user of the system but which allows the addition and subtraction of peripherals, for example, without having to shut down and repower. A new acyclic directed graph is calculated which includes the presence of the added node. It is possible that upon removing certain nodes, it will not be necessary to trigger a bus initialization, for example, when a leaf node is removed, there is no harm to the network. However, if a branch node is disengaged from an operating bus, it is likely to be necessary to reconfigure the graph.

Although the present invention has been described in terms of preferred embodiments, it will be appreciated that various modifications and alterations might be made by those skilled in the art without departing from the spirit and scope of the invention. The invention should, therefore, be measured in terms of the claims which follow.

-38-

We claim:

1. A computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communication nodes and communications links comprising a bus of an acyclic directed graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the acyclic directed graph being designated branch nodes, all nodes initially having a status of being unidentified, said acyclic directed graph having established hierarchical parent-child relationships between all adjacent nodes proceeding from the root node down to any leaf nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other adjacent nodes, the root node being defined as having no parent node, to assign unique addresses to the nodes of the acyclic directed graph the computer system further comprising:

means for each unidentified leaf node initially transmitting a "Bus Request" (BR) signal onto the bus;

means for each branch node waiting until all unidentified adjacent child nodes each propagates a BR signal, then propagating a BR signal to a parent node;

means for each node on the bus counting a number of address announcements that are broadcast on the bus;

means for the root node waiting until all unidentified adjacent nodes each propagates a BR signal,

-39-

then propagating a "Bus Grant" (BG) signal to one adjacent unidentified node based on a predetermined selection criterion of the root node for selecting adjacent nodes, said BG signal being propagated down through the acyclic directed graph through intervening nodes based on a predetermined selection criterion of each of the intervening nodes for selecting adjacent nodes until a node that initiated a BR signal receives the BG signal;

means for the node receiving the BG signal then broadcasting an address announcement to all nodes on the bus;

means for the node receiving the BG signal then setting a unique address to attaining a status of being identified, said unique address being a function of the number of address announcements counted at the node receiving the BG signal.

2. The computer system of claim 1, further comprising:

means for the node receiving the BG signal then broadcasting, on the bus, information about a component associated with the node receiving the BG signal.

3. The computer system of claim 2, further comprising:

means for the node receiving the BG signal then broadcasting, on the bus, topology information concerning the node receiving the BG signal.

4. The computer system of claim 3, wherein said topology information includes information about adjacent nodes and link status of the node receiving the BG signal.

-40-

5. A computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communications nodes and communications links comprising a bus of an acyclic directed graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the acyclic directed graph being designated branch nodes, all nodes initially having a status of being unidentified, said acyclic directed graph having established hierarchical parent-child relationships between all adjacent nodes proceeding from the root node down to any leaf nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other adjacent nodes, the root node being defined as having no parent node, to assign unique addresses to the nodes of the acyclic directed graph, the computer system further comprising:

means for each unidentified leaf node initially transmitting a "Bus Request" (BR) signal onto the bus;

means for each branch node waiting until all unidentified adjacent child nodes each propagates a BR signal, then propagating a BR signal to a parent node;

means for the root node waiting until all unidentified adjacent nodes each propagates a BR signal, then propagating a "Bus Grant" (BG) signal to one adjacent unidentified node based on a predetermined selection criterion of the root node for selecting adjacent nodes, said BG signal being propagated down through the acyclic

-41-

directed graph through intervening nodes based on a predetermined selection criterion of each of the intervening nodes for selecting adjacent nodes until a node that initiated a BR signal receives the BG signal;

means for the node receiving the BG signal then broadcasting an address announcement to all nodes on the bus;

means for the node receiving the BG signal then broadcasting, on the bus, information about a component associated with the node receiving the BG signal;

means for each node on the bus counting, as a number, how many address announcements are broadcast by other nodes;

means for each node setting a unique address as a function of the number of address announcements broadcast by other nodes until the node received a BG signal and subsequently broadcasts an address announcement, said node then attaining the status of being identified.

6. In a computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communications nodes and communications links comprising a bus of an acyclic directed graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the acyclic directed graph being designated branch nodes, all nodes initially having a status of being unidentified, said acyclic directed graph having established hierarchical parent-child relationships between all adjacent nodes

-42-

proceeding from the root node down to any leaf nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other adjacent nodes, the root node being defined as having no parent node, to assign unique addresses to the nodes of the acyclic directed graph the computer system further comprising:

means for each unidentified leaf node initially transmitting a "Bus Request" (BR) signal onto the bus;

means for each branch node waiting until all unidentified adjacent child nodes each propagates a BR signal, then propagating a BR signal to a parent node;

means for each node on the bus counting a number of address announcements that are broadcast on the bus;

means for the root node waiting until all unidentified adjacent nodes each propagates a BR signal, then propagating a "Bus Grant" (BG) signal to one adjacent unidentified node based on a predetermined selection criterion of the root node for selecting adjacent nodes, said BG signal being propagated down through the acyclic directed graph through intervening nodes based on a predetermined selection criterion of each of the intervening nodes for selecting adjacent nodes until a node that initiated a BR signal receives the BG signal;

means for the node receiving the BG signal then broadcasting an address announcement to all nodes on the bus;

means for the node receiving the BG signal then broadcasting, on the bus, topology information concerning the node receiving the BG signal;

-43-

means for the node receiving the BG signal then setting a unique address to attaining a status of being identified, said unique address being a function of the number of address announcements counted at the node receiving the BG signal.

7. The computer system of claim 6, wherein said topology information includes information about adjacent nodes and link status receiving the BG signal.

8. In a computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communications nodes and communications links comprising a bus of a directed acyclic graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the directed acyclic graph being designated branch nodes, said directed acyclic graph having established hierarchical parent-child relationships between all adjacent nodes proceeding from the root node down to any leaf nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other adjacent nodes, the root node being defined as having no parent node, a first node connected to at least one adjacent node on the bus in a node identification process, the first node comprising:

means for counting, since a starting point of a node identification process, a number of address announcements broadcast on the bus;

-44-

when the first node obtains access to broadcast;
means for setting a unique address for the first node, the unique address being as a function of the number of address announcements counted at the first node; and
means for broadcasting an address announcement from the first node to all nodes on the bus;
when the first node has at least one adjacent child node;
means for receiving a first "Bus Request" (BR) signal from a second node, the second node being one of the at least one adjacent child node of the first node;
when the first node has a parent node:
means for forwarding the first BR signal to the parent node of the first node after receiving the first BR signal;
means for forwarding a first "Bus Grant" (BG) signal for the first BR signal from the parent node of the first node to the second node;
means for transmitting a second BR signal for the first node to the parent node of the first node; and
means for receiving a second BG signal for the second BR signal from the parent node of the first node;
when the first node has no parent node:
means for transmitting a third BG signal for the first BR signal to the second node; and
means for granting the first node access to broadcast.

9. The first node of claim 8, wherein the address announcement from the first node includes information about a component associated with the first node.

-45-

10. The first node of claim 9, wherein the address announcement from the first node includes topology information concerning the first node.

11. The first node of claim 10, wherein said topology information includes information about adjacent nodes of the first node and link status.

12. The first node of claim 8, wherein transmitting the second BR signal for the first node is performed after forwarding the first BG signal.

13. The first node of claim 12, further comprising:
means for waiting until receiving BR signals from all unidentified nodes in the at least one adjacent child node before forwarding the first BR signal;

wherein the first node forwards BG signals to the at least one adjacent child node in an order based on a predetermined adjacent node selection criterion.

14. The first node of claim 8, wherein the first node grants the first node access to broadcast after transmitting the third BG signal.

15. The first node of claim 14, wherein the first node is the root node.

16. A first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to the at least one adjacent node respectively, the first node comprising:

-46-

means for counting, since a starting point of a node identification process, a number of announcements received at the first node;

means for broadcasting an announcement from the first node when the first node obtains access to broadcast; and

means for setting a unique address for the first node based on the number of announcements received by the first node since the starting point to when the first node obtains the access to broadcast.

17. A first node as in claim 16, wherein the first node obtains the access to broadcast after the first node receives a "Bus Grant" (BG) signal for a "Bus Request" (BR) signal initiated from the first node.

18. A first node as in claim 17, wherein the first node initiates the BR signal after each of child nodes of the first node obtains access to broadcast.

19. A first node as in claim 16, wherein the first node obtains the access to broadcast after each of child nodes of the first node obtains access to broadcast.

20. A first node as in claim 19, wherein the first node grants access to broadcast to the child nodes of the first node in an order based on a predetermined selection criterion for selecting adjacent nodes.

21. A first node as in claim 16, wherein the announcement comprises topology information concerning the first node.

-47-

22. A first node as in claim 21, wherein the topology information includes information about at least one adjacent node and link status.

23. A machine readable medium containing executable computer program instructions which when executed by a data processing system cause said system to perform a method performed by a first node connected to at least one adjacent node on a bus in a node identification process in a computer system comprising a plurality of components, said plurality of components each having at least one communications node wherein said communications nodes of said plurality of components are interconnected by communications links, said communications nodes and communications links comprising the bus of a directed acyclic graph wherein one node is designated a root node, all nodes coupled to only one adjacent node being designated leaf nodes, all other nodes in the directed acyclic graph being designated branch nodes, said directed acyclic graph having established hierarchical parent-child relationships between all adjacent nodes proceeding from the root node down to any leaf nodes wherein a leaf node has only one parent node and all nodes adjacent to the root node are child nodes with respect to the root node but parent nodes with respect to other adjacent nodes, the root node being defined as having no parent node, the method comprising:

counting, since a starting point of a node identification process, a number of address announcements broadcast on the bus;

when the first node obtains access to broadcast;

-48-

setting a unique address for the first node, the unique address being as a function of the number of address announcements counted at the first node; and

broadcasting an address announcement from the first node to all nodes on the bus;

when the first node has at least one adjacent child node:

receiving a first "Bus Request" (BR) signal from a second node, the second node being one of the at least one adjacent child node of the first node;

when the first node has a parent node:

forwarding the first BR signal to the parent node of the first node after receiving the first BR signal;

forwarding a first "Bus Grant" (BG) signal for the first BR signal from the parent node of the first node to the second node;

transmitting a second BR signal for the first node to the parent node of the first node; and

receiving a second BG signal for the second BR signal from the parent node of the first node;

when the first node has no parent node:

transmitting a third BG signal for the first BR signal to the second node; and

granting the first node access to broadcast.

24. The medium of claim 23, wherein the address announcement from the first node includes information about a component associated with the first node.

25. The medium of claim 24, wherein the address announcement from the first node includes topology information concerning the first node.

-49-

26. The medium of claim 25, wherein said topology information includes information about adjacent nodes of the first node and link status.

27. The medium of claim 23, wherein said transmitting the second BR signal for the first node is performed after said forwarding the first BG signal.

28. The medium of claim 27, wherein the method further comprises:

waiting until receiving BR signals from all unidentified nodes in the at least one adjacent child node before said forwarding the first BR signal;

wherein the first node forwards BG signals to the at least one adjacent child node in an order based on a predetermined adjacent node selection criterion.

29. The method of claim 23, wherein the first node grants the first node access to broadcast after said transmitting the third BG signal.

30. The medium of claim 29, wherein the first node is the root node.

31. A machine readable medium containing executable computer program instructions which when executed by a data processing system cause said system to perform a method for a first node to communicate with at least one adjacent node through at least one point-to-point link connected from the first node to that at least one adjacent node respectfully, the method comprising:

-50-

counting, since a starting point of a node identification process, a number of announcements received at the first node;

broadcasting an announcement from the first node when the first node obtains access to broadcast; and

setting a unique address for the first node based on the number of announcements received by the first node since the starting point to when the first node obtains the access to broadcast.

32. A medium as in claim 31, wherein the first node obtains the access to broadcast after the first node receives a "Bus Grant" (BG) signal for a "Bus Request" (BR) signal initiated from the first node.

33. A medium as in claim 32, wherein the first node initiates the BR signal after each of child nodes of the first node obtains access to broadcast.

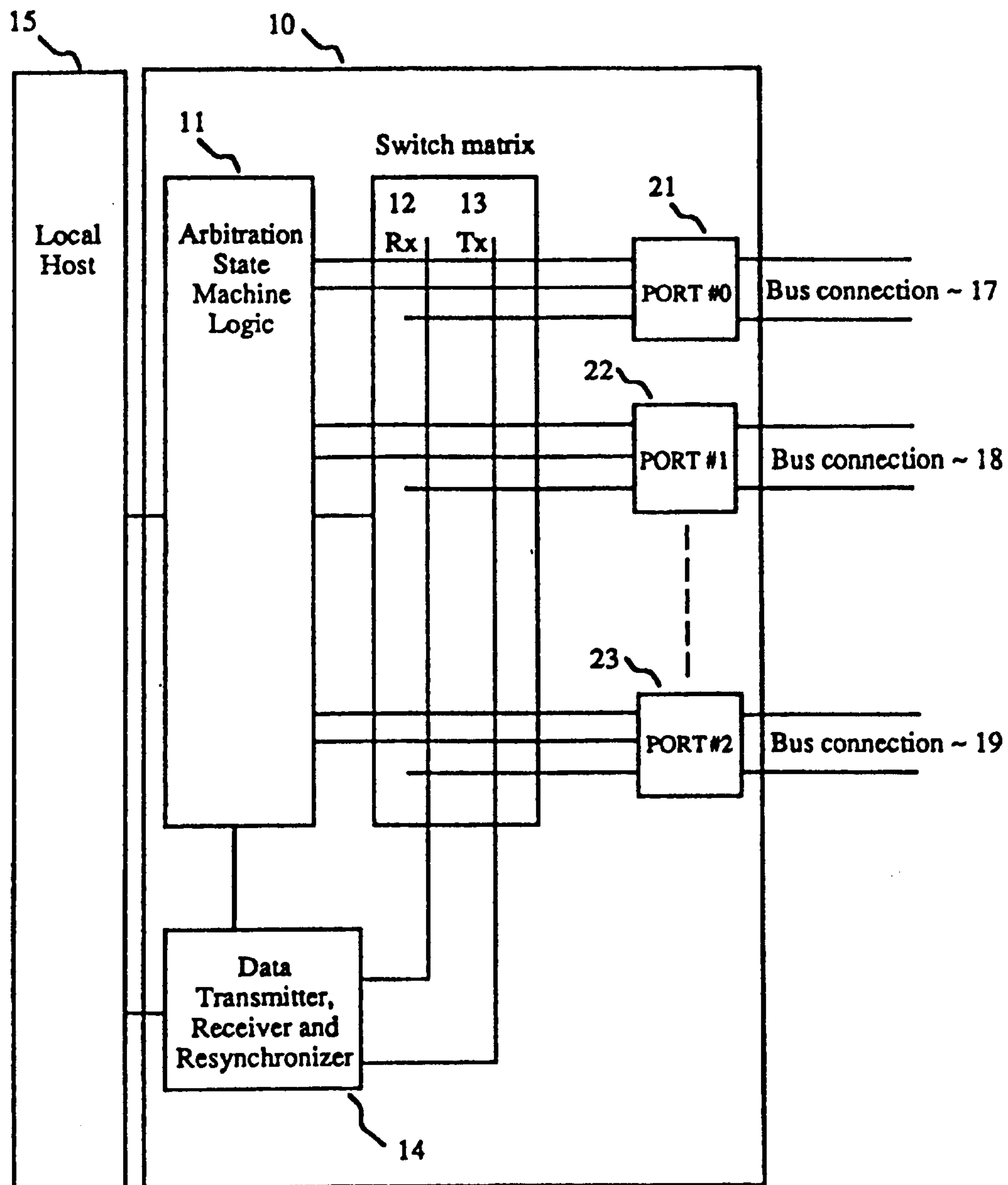
34. A medium as in claim 31, wherein the first node obtains the access to broadcast after each of child nodes of the first node obtains access to broadcast.

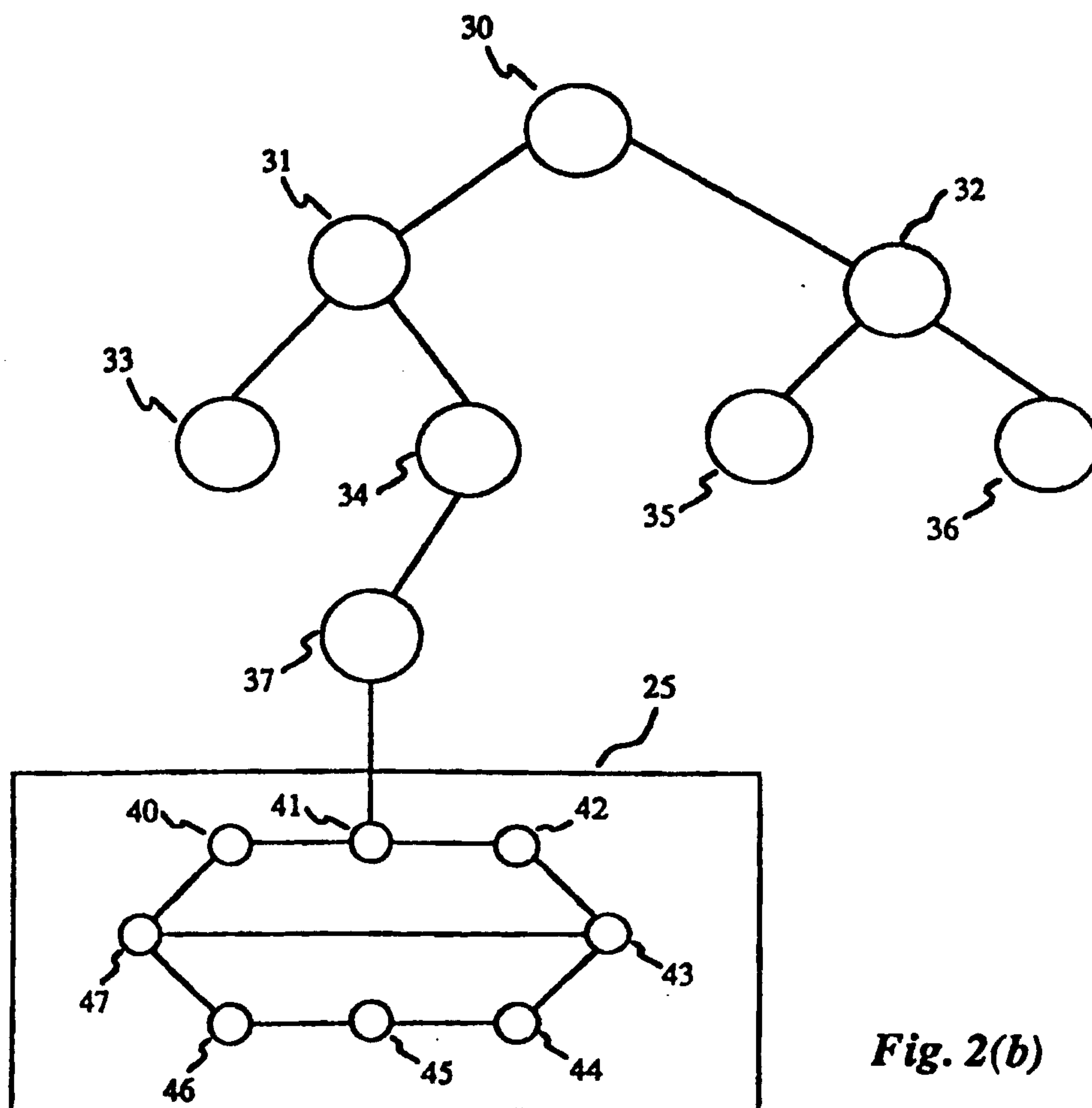
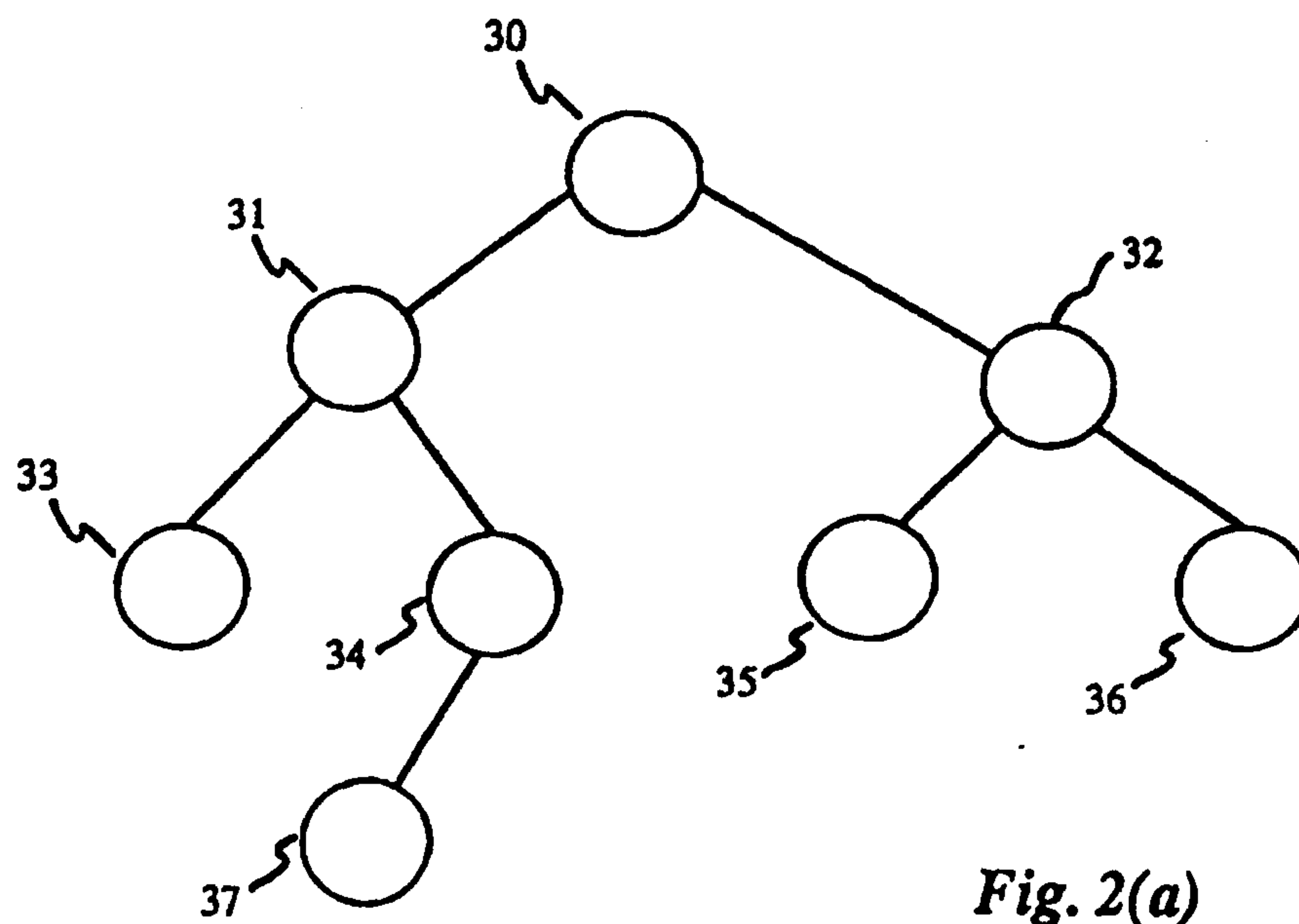
35. A medium as in claim 34, wherein the first node grants access to broadcast to the child nodes of the first node in an order based on a predetermined selection criterion for selecting adjacent nodes.

36. A medium as in claim 31, wherein the announcement comprises topology information concerning the first node.

-51-

37. A medium as in claim 36, wherein the topology information includes information about at least one adjacent node and link status.

*Fig. 1*



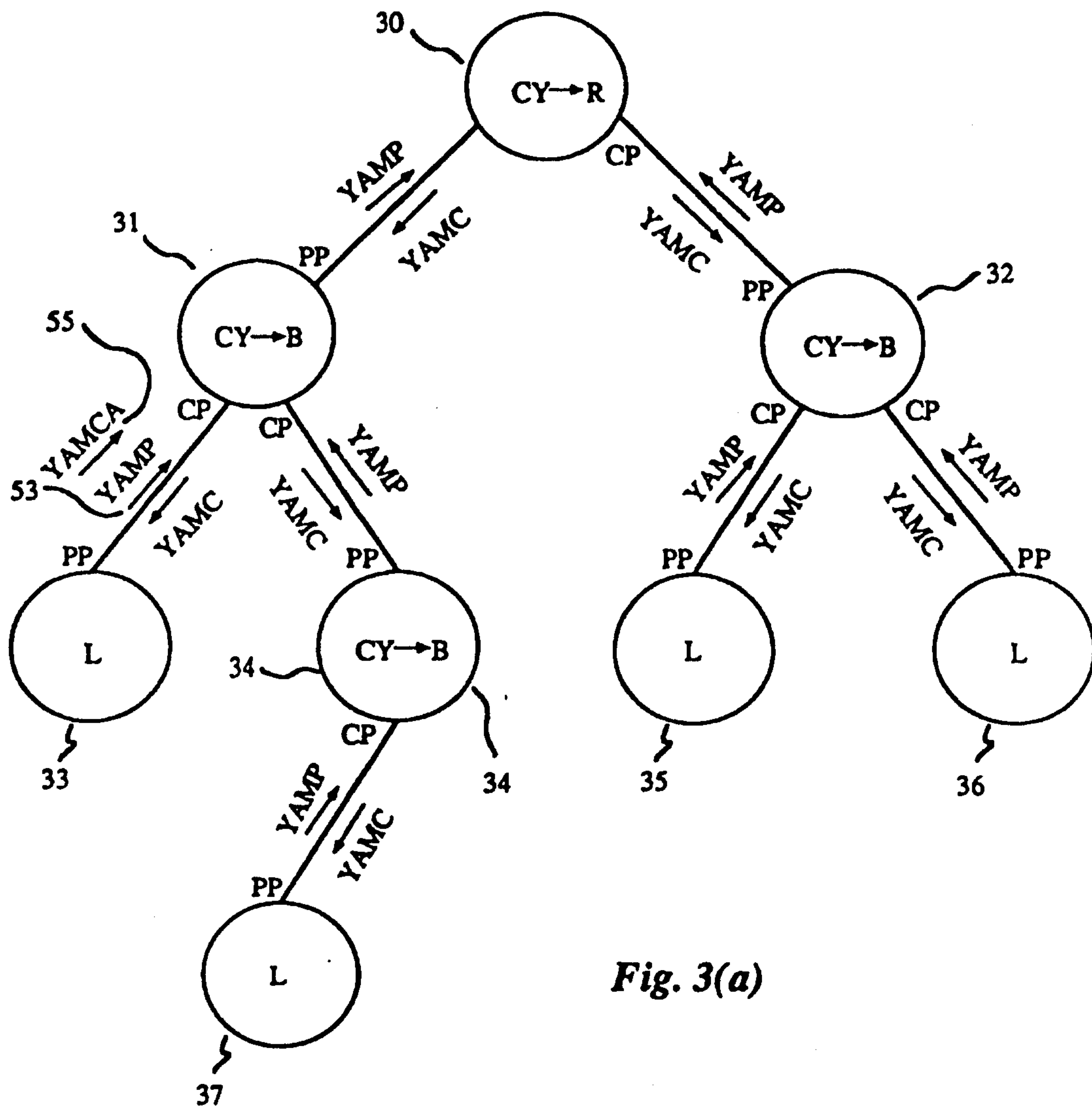


Fig. 3(a)

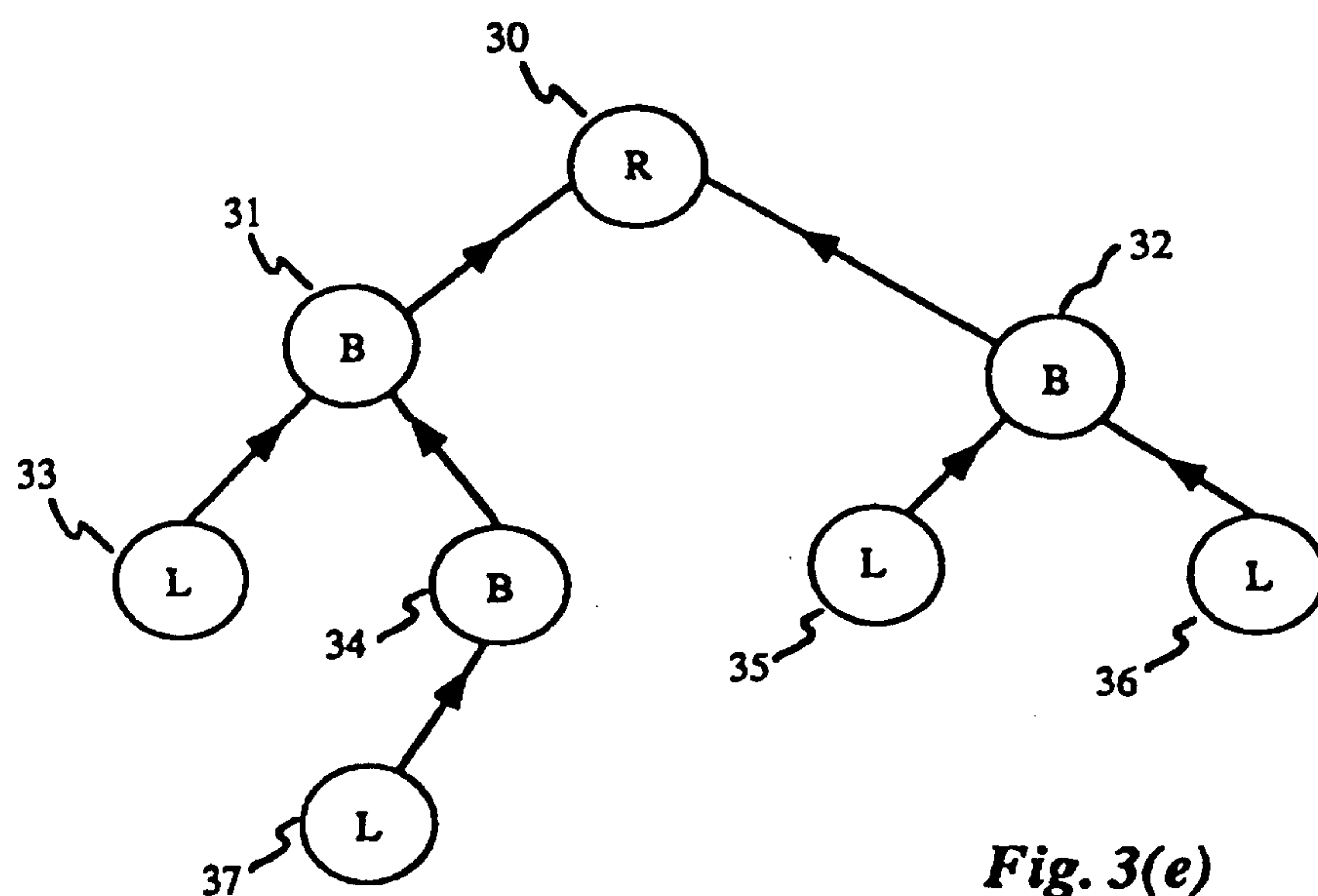
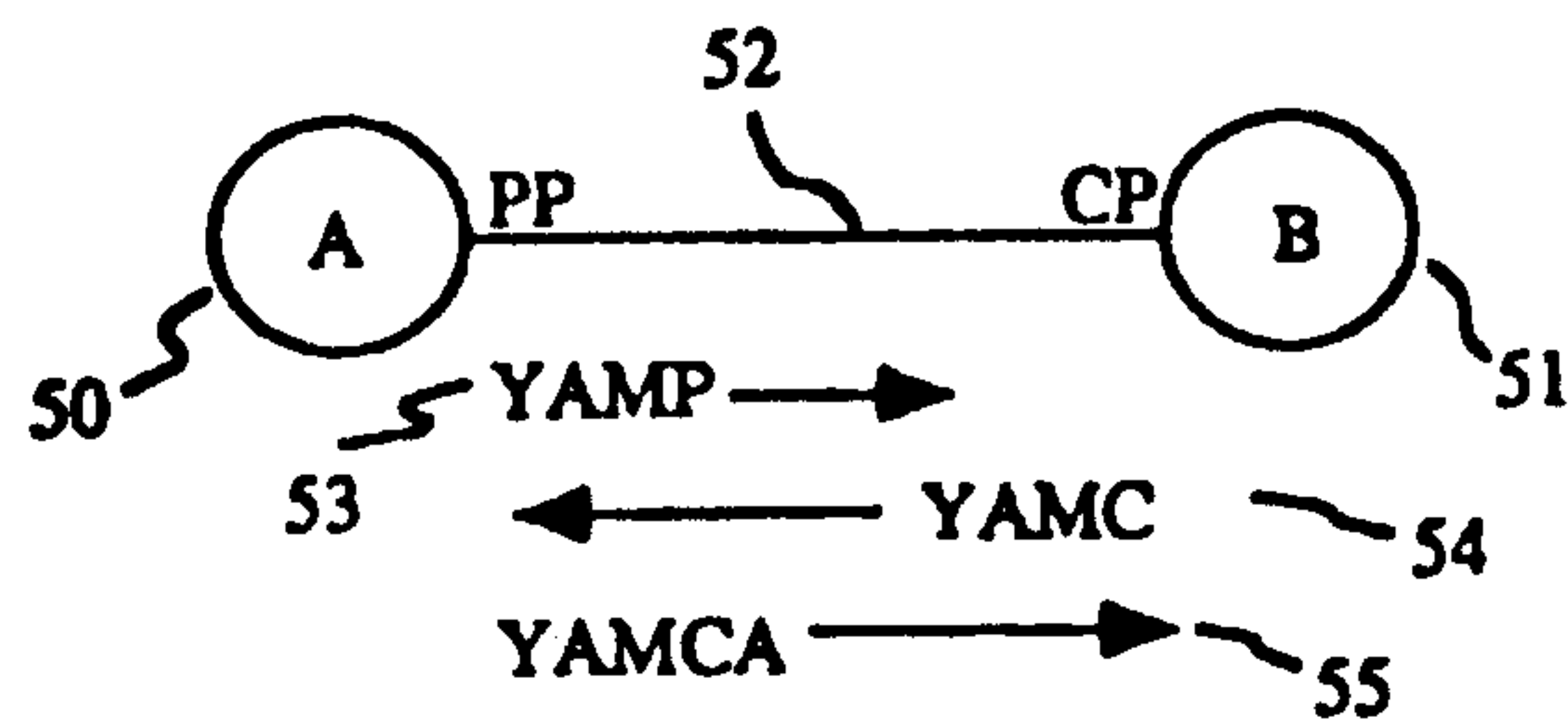
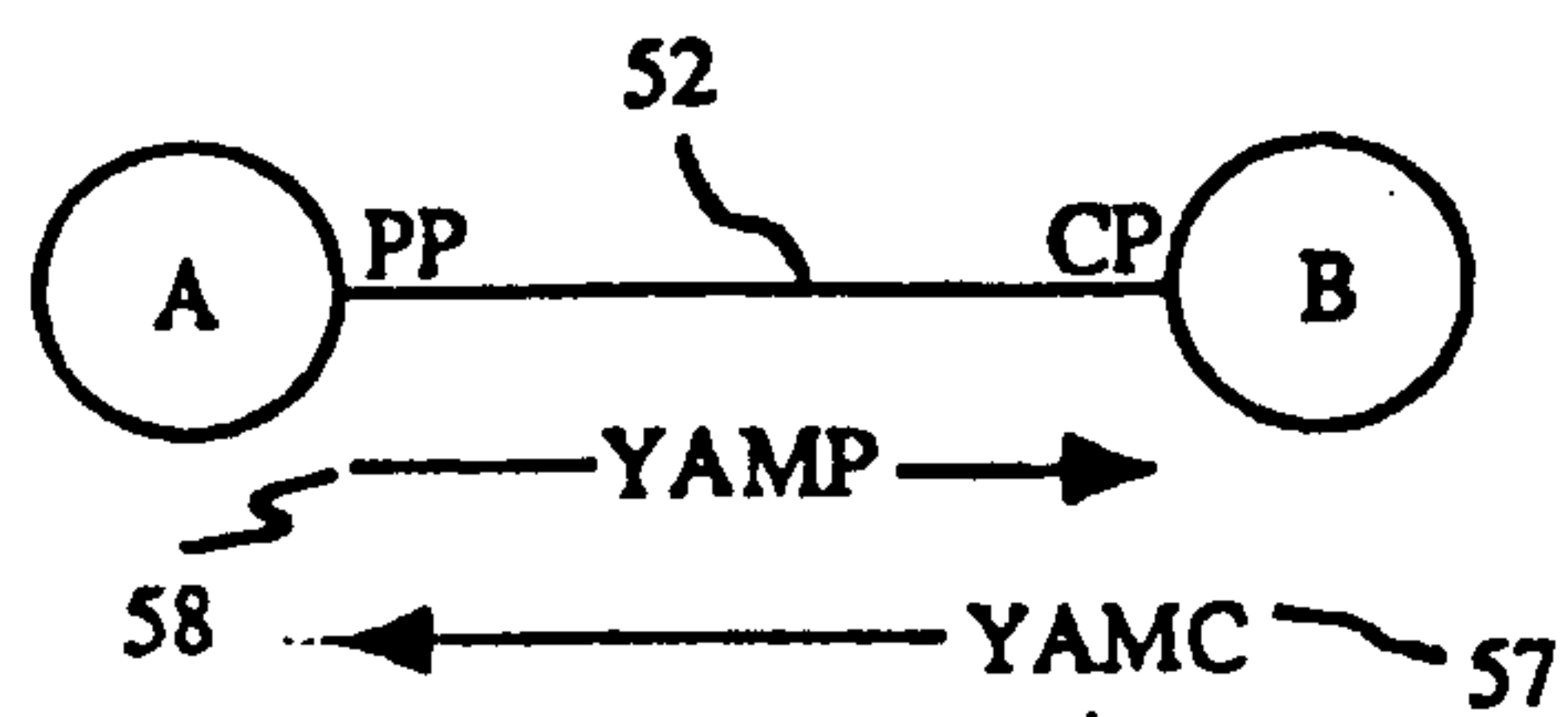
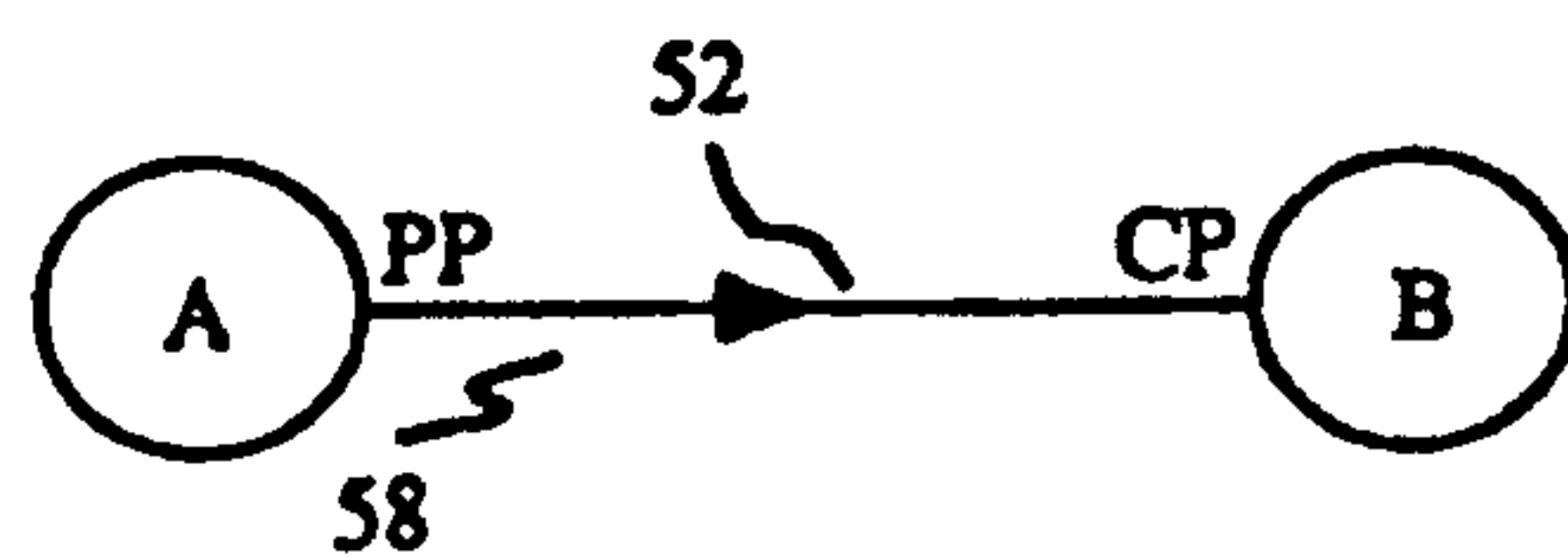
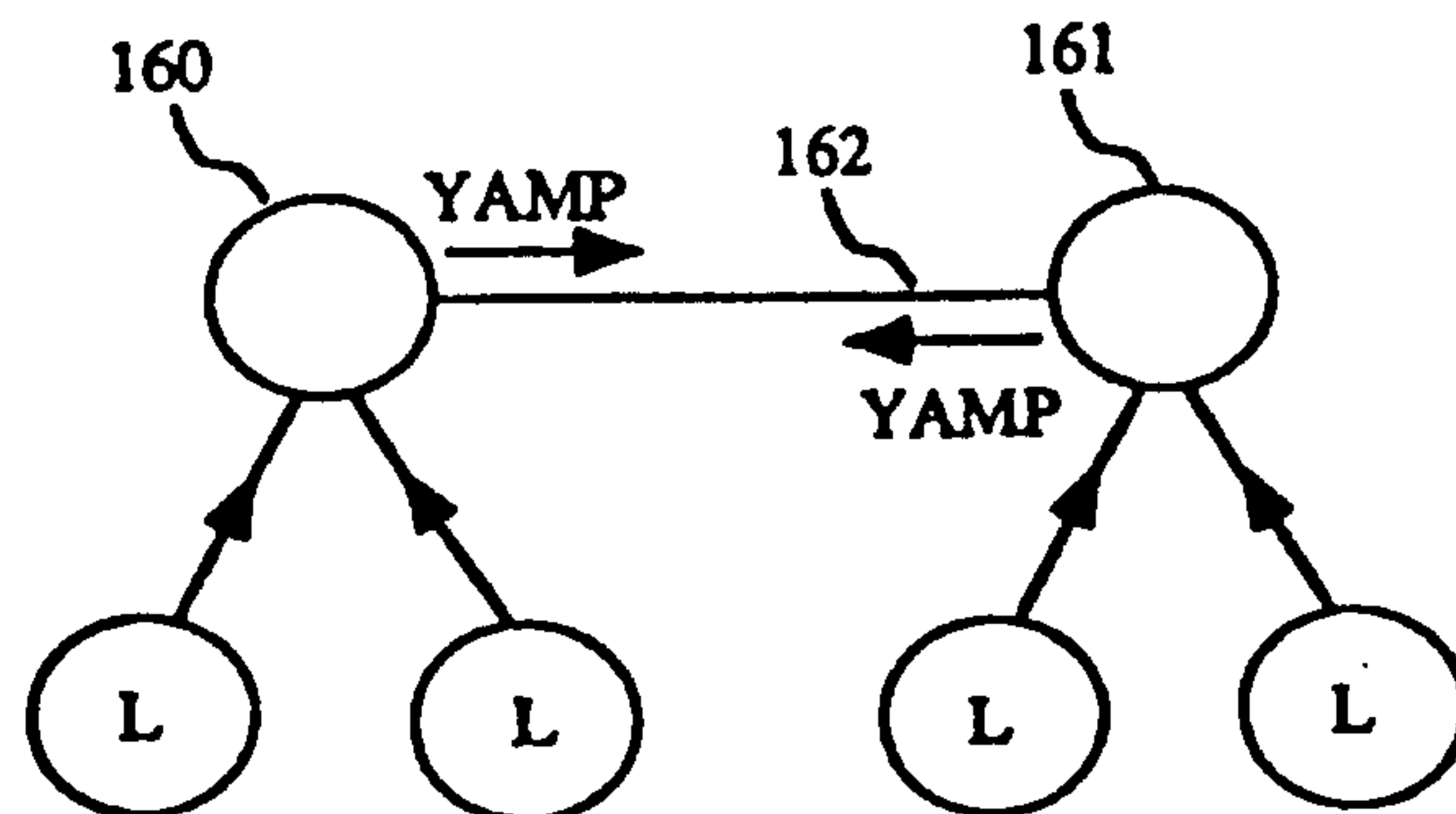


Fig. 3(e)

WO 94/15303

PCT/US93/12314

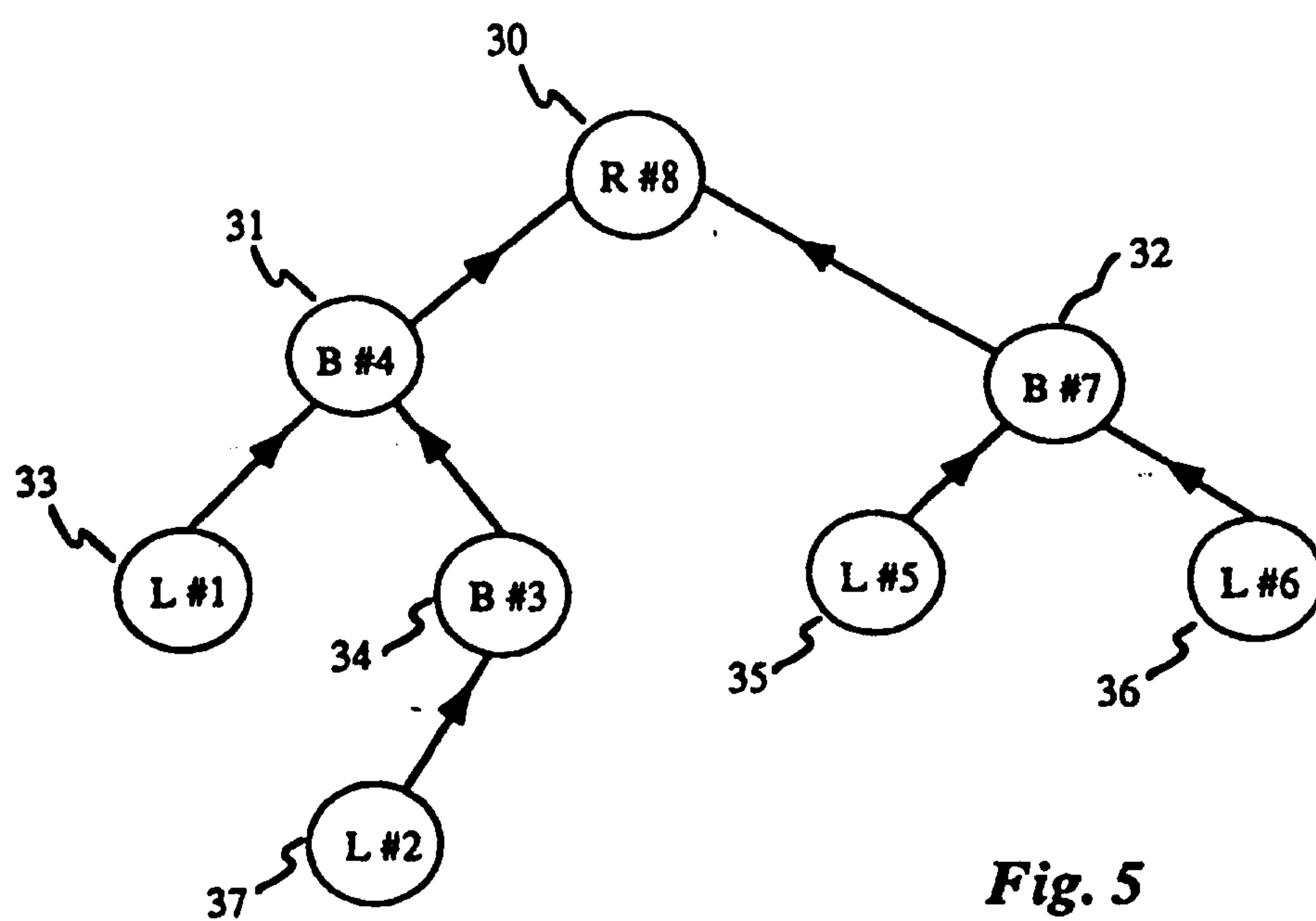
4 / 1 0

*Fig. 3(b)**Fig. 3(c)**Fig. 3(d)**Fig. 4*

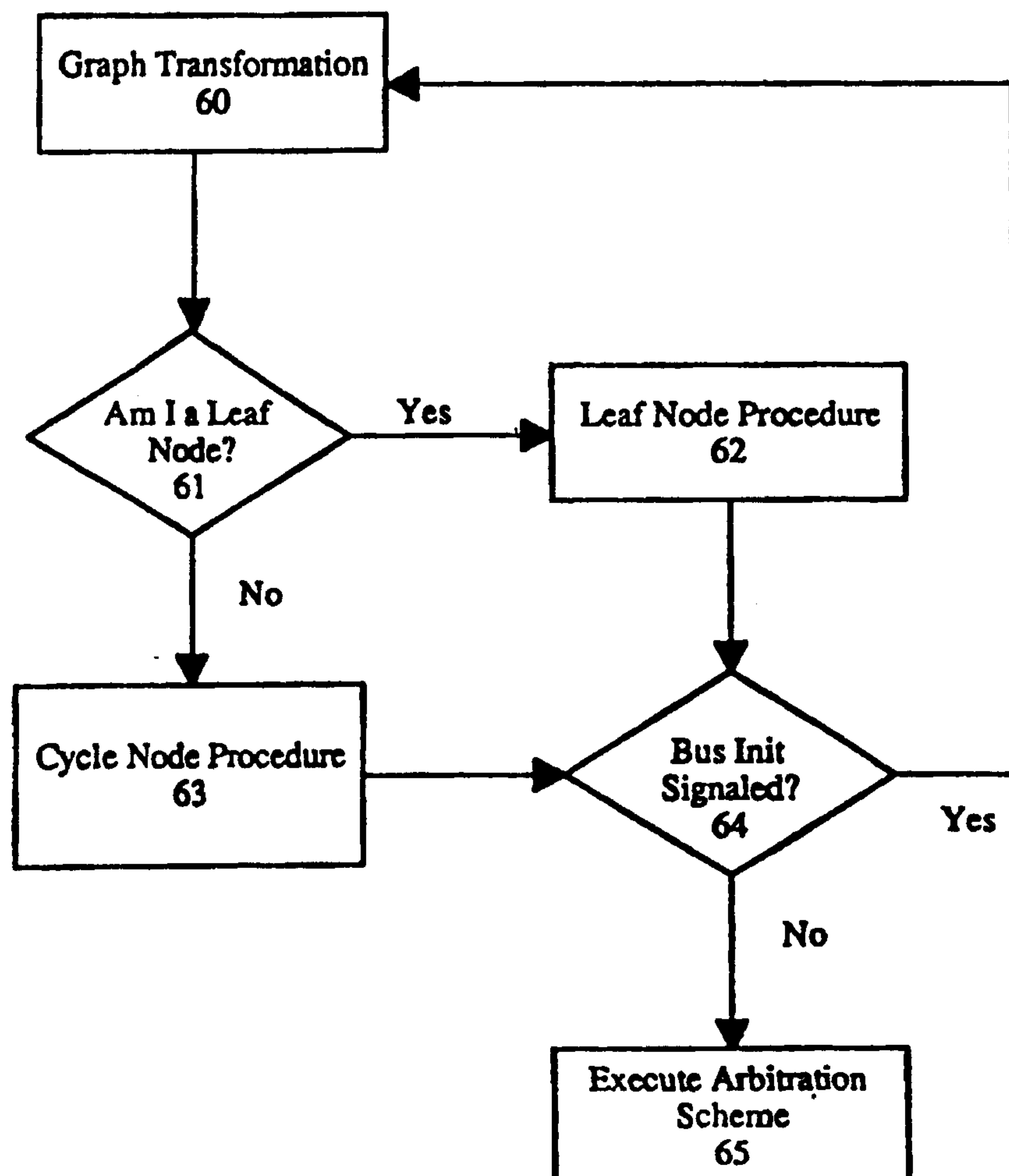
WO 94/15303

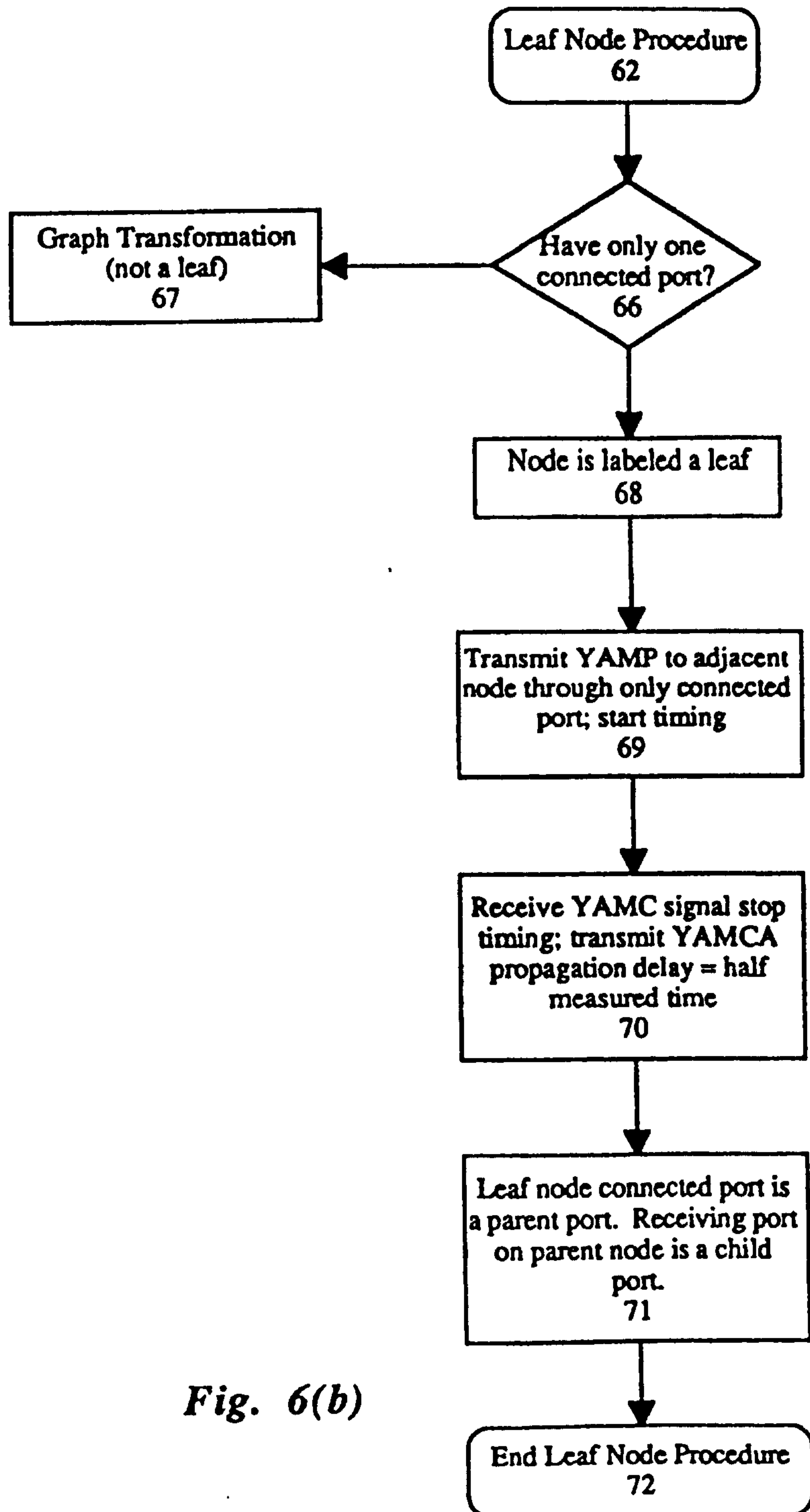
PCT/US93/12314

5 / 1 0

**Fig. 5**

6 / 1 0

*Fig. 6(a)*

*Fig. 6(b)*

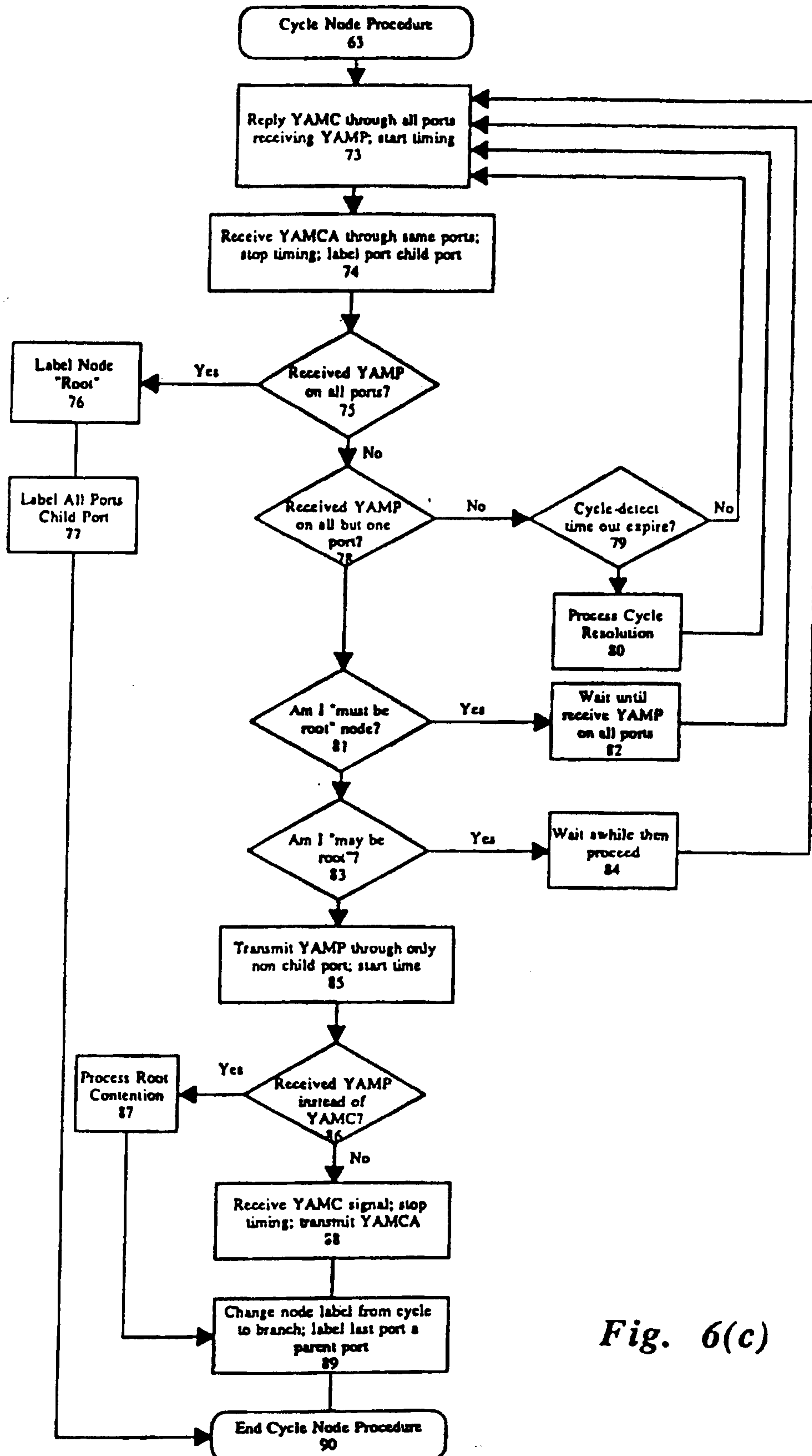


Fig. 6(c)

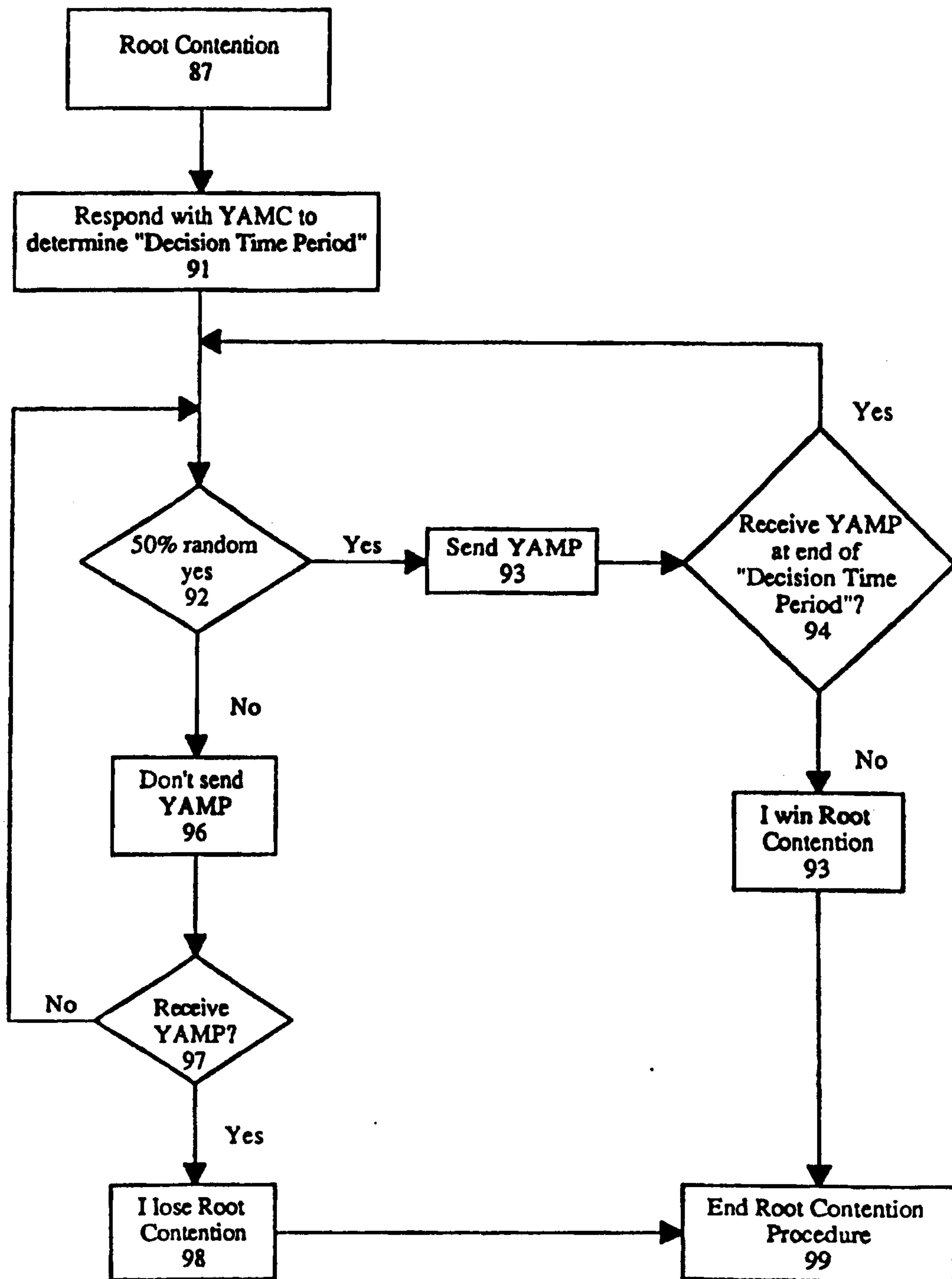
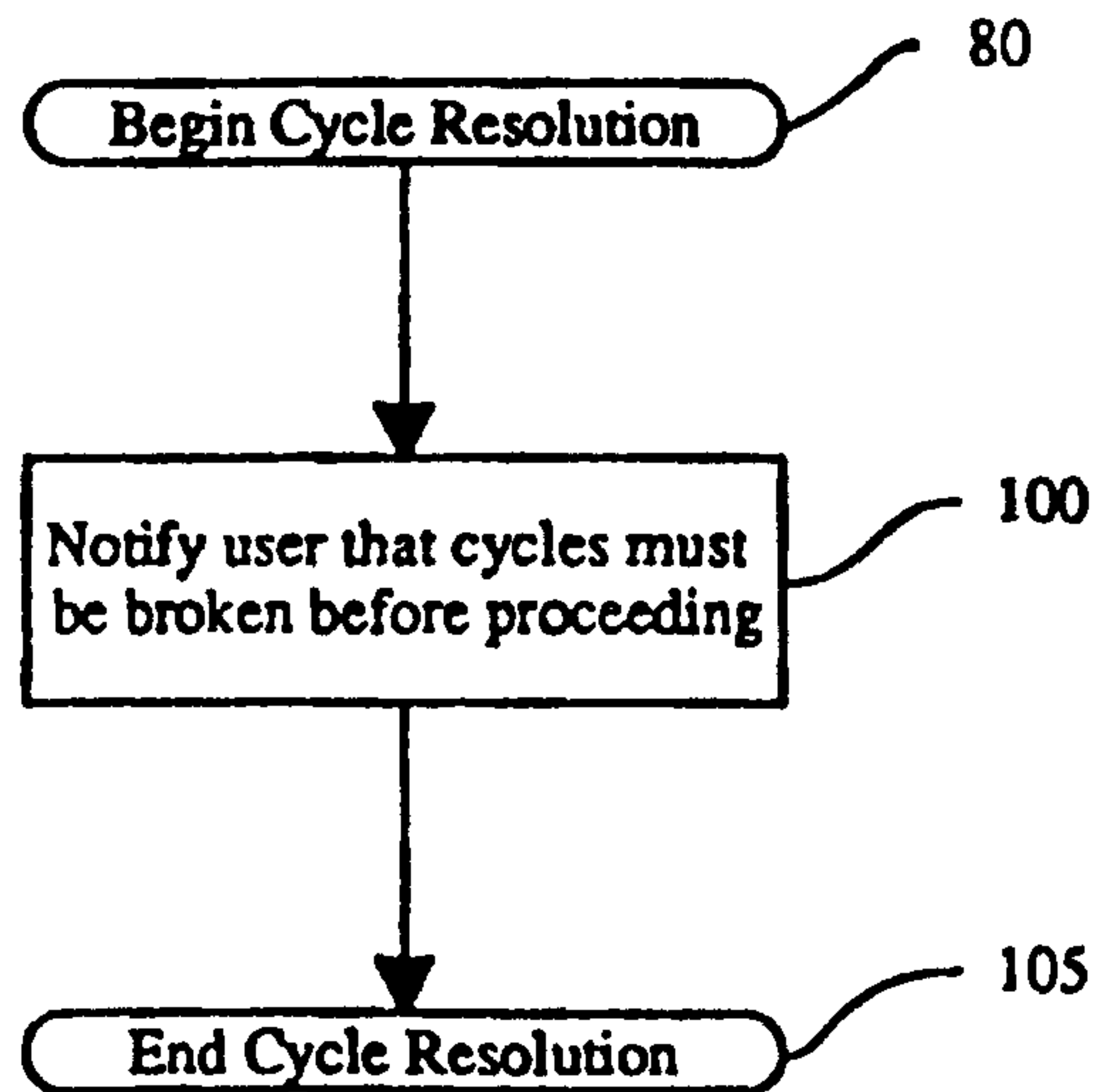


Fig. 6(d)

WO 94/15303

10 / 10

PCT/US93/12314

*Fig. 6(e)*

