

METHOD FOR OPTIMIZING A SERVICE FOR DISSEMINATING NETWORK QUALITY OF SERVICE INFORMATION

The subject of the present invention is a method and a device for
5 optimizing the management and dissemination of overlay link quality information. An overlay network is a logical network overlaid on a computer network, the nodes of the overlay network being interconnected by logical links. A logical link corresponds to the aggregation of the links and equipment which are located on the path between two nodes in the physical infrastructure of the underlying
10 network. The invention relates, for example, to the dissemination of overlay link quality information. More specifically, the invention performs an adaptation of the frequency at which the links are "monitored" according to their use. The term "monitoring" in the present invention refers to the evaluation of a link, the tracking of the performance levels offered by a logical link in an overlay network. The
15 invention notably performs an adaptation of the frequency at which the link quality information is disseminated according to its use.

The invention lies notably in the field of infrastructure networks made up of a plurality of nodes interconnected via one or a plurality of wide area networks, WAN, such as the Internet. A node in the overlay topology can be independent or
20 belong to a local area network LAN.

The invention relates to the field of the network architectures in which performance indicators are collected to take routing decisions for the communication flows.

A server meeting node in the present application is a node which will
25 determine the best route between a given pair of nodes and which will return this best route information (routing table) to the two nodes of the pair, nodes which are called client meeting nodes.

In an overlay-type architecture, the nodes of the network are interconnected through at least one wide area network (WAN). A wide area
30 network, such as the Internet, provides access to the communication medium and a packet transport service ensuring connectivity between the nodes of the overlay network.

A general technical problem associated with the networks of this type is the physical and logical routing of the data flows between the nodes so as to observe the broadcasting constraints associated with each of them. In practice, it is possible to implement a routing mechanism in an overlay network to route the data flows between two nodes by using an overlay path (aggregation of logical links and of overlay nodes) other than the logical link directly interconnecting these two nodes. This technique refers to the traffic engineering, which makes it possible, for example, to ensure that the level of quality required by certain data flows is reached. By extension, it also makes it possible to ensure a better level of resilience. The resilience can be defined as the capacity of a network to withstand various types of threats by detecting them and by applying corrective solutions within a minimal delay period to ensure operation even in degraded mode. The threats may be simple failures, a degradation of the available resources, critical situations damaging to the physical infrastructure (natural catastrophes, piracy, etc.) or even illegal traffic which could provoke a malfunction. Ensuring the resilience of the networks in themselves, but also and above all for services that they provide or support, represents a major issue.

Conducting traffic engineering in an overlay network requires the ability to compare the performance levels offered by the logical links. The interconnection network is assumed to be uncontrolled (it may be a single operator network or a set of interconnected operator networks), and to not supply information on its state and on the performance levels that it offers. Each node of the overlay network must therefore observe the state of the network and, more particularly, of the logical links which link it to the other nodes of the network.

One of the technical problems posed is therefore of “monitoring” the logical links, and therefore the state of the interconnection network, as is schematically represented in Figure 1 which represents an infrastructure network linking a number N of network enclaves.

In such a context, a node of the overlay is, for example, the router at the edge of each enclave, which performs the interconnection with the wide area network. A local area network (LAN) can also be called “enclave” in as much as the data flows which leave therefrom can be secured. The data flows are

transferred from one enclave to another through the uncontrolled interconnection network. This is represented by the route I between an enclave B and an enclave N in Figure 1. In order to track the state of this link, the delay between the two enclaves B, N is controlled, for example, by measuring the round-trip time (RTT) between these two enclaves. The routing plan in the interconnection network is not known to the enclaves and, in particular, it does not guarantee a minimum delay between the two enclaves, while it may be important to do so for flows such as voice communications, urgent messages. Conversely, it is possible to find a logical path with two hops in the overlay network which passes through a relay enclave and which reduces the delay on the route, as is represented by the path II of Figure 1, passing through the enclave D.

One of the technical problems posed is making it possible to find the best route with two hops, for example, in the overlay formed by the enclaves for a measured link quality criterion. In this example, the overlay is made up of the nodes of the enclaves and of the connections between all these nodes, and the criterion is the round-trip time (RTT).

Generally, traffic engineering solutions rely on the knowledge of quality of service, or QoS, metrics, such as the delay, jitter, loss rate, bandwidth in order to satisfy the quality of service constraints for a maximum of flows, or at least of the most critical flows. Other criteria, independent of the QoS indicators, can be taken into account, such as, for example, the security level, the geographic/administrative position, or the cost of the links. It is therefore necessary to have a mechanism that makes it possible notably:

- to control or monitor the connections between the nodes, that is to say all the logical links of the overlay network,
- to disseminate the information "monitored" in order for all the nodes to be able to reconstruct a useful view of the topology of the network and of the quality of the links, and thus make the best use of the available resources.

One of the problems which is posed is that the more the number of nodes increases, the more the number of connections to be monitored increases, and the more the quantity of monitoring information to be disseminated increases. The proportion of bandwidth consumed for obtaining and disseminating such

information therefore increases considerably, to the detriment of the bandwidth reserved for the data.

The document by D. Sontag, Y. Zhang, A. Phanishayee, D. Andersen, D. Karger, entitled "Scaling All-Pairs Overlay Routing" in ACM CoNEXT, Dec 5 2009, describes a method for reducing the load due to the dissemination of the link states, by logically organizing the nodes in grid form and by governing the exchanges between these nodes as follows: 1) only the nodes on a same row/column exchange the link state information, 2) each node cuts away the information collected in the step 1) and computes, for each other node on the 10 same row/column and for each destination in the overlay, the best path to be used. The result of these computations is transmitted to each node situated on the same row/column in the form of path recommendations, or more specifically, of relay nodes to be used to transfer the data flows to each other destination node. Figures 2A, 2B, 2C, 2D schematically represent the sequencing of the steps.

15 In Figure 2A, the node A and the node B will each control their links $L_{A,i}$ $L_{B,j}$ with the other nodes N_i, N_j of the network. Each node A, B measures the delay that exists between itself and each of the nodes of the network. Overall, each of the nodes N_i, N_j performs these measurements and ultimately has a connectivity map grouping together the results of these measurements to each of the other 20 nodes N_i, N_j . K_A and K_B respectively represent the connectivity maps of the nodes A and B.

Figure 2B shows the step in which each of the nodes A and B, then acting as client meeting node, transmits its connectivity map to server meeting nodes. A node of the network is, in this example, considered to be a server 25 meeting node when it is positioned on the same row or the same column to the node considered A, B. Each node A, B thus communicates its connectivity map to a set R_{sA} , respectively R_{sB} , of nodes which are its server meeting nodes.

Figure 2C is a representation of the step of computation of the recommendations. From the connectivity maps received from its client meeting 30 nodes R_c , a server meeting node R_s determines, for all the pairs of nodes made up of its client meeting nodes, the intermediate nodes to be used to transmit data flows in better conditions. Figure 2C illustrates the operation carried out by a

server meeting node that A and B have in common and which, from the connectivity maps received from these two nodes A and B, can determine which relay nodes provide better performance levels than on the direct path (in this example, a shorter delay) between the nodes A and B. During this step, the two
5 connectivity maps of the pair of nodes considered (K_A and K_B for the nodes A and B) are combined, and a single table K_{AB} is generated that comprises the results of this combination (in this example, this is a sum of delays, in milliseconds). For example, in Figure 2C, the node $N_{4,4}$ (fourth row, fourth column) is the best relay node since it provides the shortest delay (110 ms). Using this relay node rather
10 than the direct path between A and B (200 ms) therefore makes it possible to reduce the delay needed to route communications between these two nodes.

Figure 2D schematically represents the step of transmission of the recommendations from a server meeting node to the nodes A, B. The server meeting node common to the nodes A and B transmits to these two nodes a
15 message that includes, among other things, the recommendation indicating to them that the best relay node to be used to communicate is the node C.

This mechanism makes it possible notably to reduce the network overhead due to the dissemination of the controlled and managed information while guaranteeing that the best choices with two hops will be able to be made. On
20 the other hand, this approach does not deal with the preliminary link monitoring step which can be optimized in order to reduce the quantity of bandwidth needed to monitor the state of the network.

For an overlay network of N nodes, the basic approach described in the prior art entails, in each node, monitoring, for each period, the N-1 links and
25 sending this information to the N-1 nodes of the network. The overhead is N^2 . The abovementioned paper proposes a method of the order of $N\sqrt{N}$.

One of the objectives of the present invention is notably to optimize both the monitoring phase and the phase of dissemination of the information in
30 order to reduce the network overhead and therefore increase the bandwidth available for the users, to achieve a network overhead due to the (collection and the) dissemination that is lower than the abovementioned approach. The invention

makes it possible however to optimize the dissemination mechanism, a much greater consumer of bandwidth, independently of the fact that the monitoring mechanism is optimized or not optimized (no adjustment of the frequencies).

5 The objective of the invention is to couple the monitoring method to the dissemination method to reduce the load on the network. The method makes it possible to put in place the monitoring of a subset of the links at high frequency as a function of their use and their quality of service, the rest of the links being monitored at a lower frequency. The invention operates with a notification mechanism between the server meeting nodes and their client meeting nodes.
10 The server meeting nodes aggregate the monitoring information and decide on the best routes that the client meeting nodes must use. They notify their client meeting nodes of the links to be monitored as a function of the use that they make thereof.

The invention relates to a method for optimizing the dissemination of
15 link quality information in an overlay network comprising a plurality of nodes N_i , N_j , of which a number of nodes are server meeting nodes R_s , at least one client meeting node A and at least one client meeting node B, in which each client meeting node controls and performs measurements on the links that it has with the nodes of the network by using a number of frequency values F_i , the list G_i , of the
20 links to be monitored at a given frequency value being established by the server meeting nodes.

According to a variant embodiment, a client meeting node acts, for example, also as a server meeting node.

25 The object of the invention also relates to a method for optimizing the dissemination of link quality information in an overlay network comprising a plurality of nodes N_i , N_j , of which a number of nodes are server meeting nodes R_s , at least one client meeting node A and at least one client meeting node B, said method comprising:

- 30
- a step in which a server meeting node R_s determines at least one first subset G_1 of nodes for which the probability of being used as intermediate node for the communications between A and B is high and at least one

second subset G_2 of nodes that have a low probability of being used as intermediate node at a given instant, and transmits to the client nodes A and B, a list of the nodes to which the links $L_{A,i}$, $L_{B,j}$ are to be controlled at a first frequency F_1 , the other links being monitored at a second frequency F_2 ,

- 5 • a step of link state dissemination executed at said first frequency F_1 in which only the information concerning the links $L_{A,i}$, $L_{B,j}$ according to the first subset G_1 is transmitted from a node A, B to a server meeting node R_s having requested this information,
- a step of link state dissemination executed at said second frequency F_2 in
10 which the set of quality information concerning the other links $L'_{A,i}$, $L'_{B,j}$ is returned from the nodes A, B to the server meeting nodes R_s ,
- a step in which each server meeting node determines the best route or the best path to route the communications from the node A to the node B,
- a step in which a server meeting node sends a route recommendation to
15 the nodes A, B indicating the routes which offer better properties than the direct connection for joining other nodes.

The best path determined for the client meeting nodes is, for example, a route with two hops comprising a single intermediate node.

To determine a subset G_1 , G_2 , the n nodes which offer the shortest
20 delays between A and B can be considered.

Another possibility: to determine a subset G_1 , a server meeting node uses, for example, as criterion, the travel time delay RTT between a node A, B and a node of the network.

To determine the subset of nodes G_1 , a server meeting node can also
25 use as criterion the loss rate and/or the bandwidth value.

To determine the subsets of nodes G_1 , a server meeting node can also use as criterion the stability of a link $L_{A,i}$, $L_{B,j}$ or the cost of the link, or the level of security of the link, or, more generally, a composite metric of a set of quality or stability/cost/security criteria.

30 According to a variant embodiment, a server meeting node determines a set of intermediate nodes by using as criterion the value of the round trip time

RTT and a server meeting node transmits this information to the client meeting nodes.

5 A network is used, for example, that has a structure in which the server meeting nodes are situated one and the same row and one and the same column as a node A, B or a network of DHT, central server type.

According to a variant embodiment, the values of the link monitoring frequencies are varied as a function of the nature of the network, or of the traffic engineering objective, or of events occurring in the network.

10 The invention also relates to a device for optimizing the dissemination of link quality information in an overlay network comprising a plurality of nodes N_i , N_j , of which a number of nodes are server meeting nodes R_s , at least one client meeting node A and at least one client meeting node B, characterized in that it comprises at least the following elements:

For a server meeting node:

- 15
- a communication module,
 - a database comprising the parameters associated with the strategy implemented by the method,
 - an optimization control module,
 - a notification mechanism control module,

20

 - a module suitable for controlling the broadcasting of the recommendations stored in a recommendations database,
 - a module for controlling the state of the links and a link states database,
 - a module for defining and controlling the structure of the overlay network,
 - a database containing a table of the nodes of the overlay network which is

25

 - known to the controller.

For a client meeting node:

- a communication module,
- a module for controlling the optimization of the monitoring of the links and of the dissemination of the information at the node level,

30

- a database containing parameters of the information dissemination strategy,
- a routing table,

- a module for controlling notifications which will recover the notification commands/messages sent by a server meeting node,
- a database comprising notifications sent by a server meeting node,
- 5 • a module, for controlling recommendations and a database comprising the tables of recommendations in the choice of nodes to be used,
- a module for measuring and controlling the state of the links stored in the link states database,
- a monitoring control module,
- 10 • a module for controlling the structure of the overlay network suitable for defining the arrangement of the nodes stored in a database.

Other features and advantages of the device according to the invention will become more apparent on reading the following description of an exemplary embodiment given as a nonlimiting illustration with appended figures which represent:

- Figure 1, a representation of an overlay network according to the invention,
- Figures 2A, 2B, 2C, 2D, an illustration of the steps of a method according to the prior art,
- 20 • Figures 3A, 3B, 3C, the steps for determining the subsets of nodes to be monitored at different frequencies,
- Figures 4A, 4B, 4C, the steps for determining the best path, and
- Figure 5A schematically represents an exemplary architecture of the modules of a server meeting node, and Figure 5B, an exemplary architecture of a client meeting node.

The following example illustrates the operation of the method according to the invention. It is given in the case where the nodes of the overlay network are organized in grid form in which the nodes of one and the same row/column are server meeting nodes Rs and client meeting nodes Rc to one another.

The method according to the invention relies on a dynamic adaptation of the strategy for “monitoring” the connections and the strategy for “disseminating” the monitored information enabling each node:

- to maintain information on the state of the links and their use,
- to exchange this information with certain other nodes in order to construct a view of the topology of the overlay network that includes information associated with the links,
- 5 • to determine, according to the constraints of the different flows, the paths best suited for transmitting them,
- to thus implement the routing policies chosen while reducing the network overhead.

In the example given to illustrate the method implemented by the invention, a pair of source/destination nodes, A and B, are considered and the quality of the routes with two hops will be examined. The server meeting nodes are, in the present example, nodes situated on the same row and the same column as a given node. Another logical structure that can be used is, for example, a distributed hash table (better known by the acronym DHT). A single
15 central server, in this case called server meeting node, in which all the nodes, then called client meeting nodes, would send their links state information, could also be used.

For the pair of nodes A, B, a server meeting node will, initially, determine the performance levels offered by all the routes with two hops between
20 these two nodes, by combining their connectivity maps.

The method then consists in generating, for each pair of nodes A, B, a first subset G_1 of the nodes of the network and a second subset G_2 of nodes, from the result of the combination of their connectivity maps.

The first subset G_1 of nodes comprises the nodes which have a high
25 chance of being chosen as best intermediate node to convey the data flows between A and B. The second subset G_2 of nodes is made up of the nodes which have a small chance of serving as intermediate node. It is also possible to construct more than two subsets.

To generate these two subsets, the process may use a number of
30 methods. The first method consists in determining a distribution of the nodes in these subsets. For example, the N nodes are ranked in ascending order according to the chosen metric (for example in the order of the delays that they offer as

intermediate node) and it is decided that n of these nodes belong to the subset G_1 and that $N-n$ of these nodes belong to the subset G_2 . Provision can also be made for this distribution to be adapted over time. The n nodes are, for example, those which offer the shortest delays between A and B. Another way of proceeding would be to set a threshold, the nodes for which the result of the combination is under the threshold form part of a first subset and the other nodes form part of the second subset. This threshold can be adapted over time.

Figure 3A schematically represents a step in which each server meeting node R_s determines the subsets of nodes. This step follows the step of computation of the best paths with two hops for their respective client meeting nodes A and B. Just as in Figure 2C, this phase consists in combining the tables K_A and K_B in order to obtain, in a resulting table K_{AB} , the characteristic of the paths which link the nodes A and B by passing through each other node of the network. A server meeting node R_s then ranks the intermediate nodes according to the corresponding values in the table K_{AB} . These values represent, for example, the round-trip time (RTT) on the path concerned, its stability, its capacity (bandwidth), or any other metric that might be made up of these different indicators. This ranking thus makes it possible to determine which nodes form part of the subset G_1 and which other nodes form part of another subset G_2 . A server meeting node R_s needs only a subset of link state measurements to choose the best intermediate node used to transmit information between the nodes A and B. According to the example, the nodes that have resulting values 110 ms, 140 ms, 150 ms and 160 ms are the nodes which would make it possible to reduce the delay if they were used as relay between the nodes A and B. They constitute, for example, the subset G_1 since they are situated under the threshold value of the delay on the direct path between A and B (at 200 ms). The links between A and these nodes (respectively, B and these nodes) will then be monitored at a higher frequency. The other nodes form part of the subset G_2 , and the corresponding links will be monitored less frequently.

According to the development of the measurements collected, and therefore of the values contained in the resulting table K_{AB} , the threshold defining the fact that a node belongs to one or other of the subsets, or the rule of

distribution of the nodes in these subsets can be adapted during operation. Figure 3B represents the recommendation step during which a server meeting node R_{sA} , R_{sB} will transmit to a client meeting node A, B, the information on the best relay node to be used to reach each other destination node out of its client meeting nodes.

Figure 3C schematically represents the notification step in which a server meeting node R_s will transmit to its client meeting nodes R_{cA} , R_{cB} a list comprising the nodes which have to be monitored by the client meeting nodes at a higher frequency, frequency F_1 . A server meeting node R_s thus notifies its client meeting nodes of its interest in frequently receiving the states of the links to the nodes which have the greatest chance of being selected as intermediate nodes. Consequently, the client meeting nodes monitor and send the link states concerned (generally, those where the metric is the best and most stable) at a higher frequency. The regular notification mechanism also makes it possible to adjust the monitoring phase of each node by modulating the measurement frequency according to the use which is made of its result. This mechanism also makes it possible to adjust the dissemination of the monitoring information by sending only the information that is updated and necessary to each server meeting node R_s .

For a node A, if the set of links which link it to nodes belonging to G_1 , to be monitored at a high frequency, is greater than the set of links which link it to nodes belonging to G_2 , to be monitored at a lower frequency, it will then be more effective to send the identifiers of this second set. Figure 4A schematically represents a modified monitoring step for the two nodes A and B according to the steps of the method according to the invention.

Each node A, B will “monitor” the links that it has with the other nodes of the network, but at two different frequencies, F_1 and F_2 depending on whether a node N_i , N_j belongs to the subset G_1 or to the subset G_2 .

The node A will monitor, at a frequency F_1 , all the links $L_{A,i}$, for which it will have received a notification from one of its server meeting nodes R_{sA} (generally, links that it has with the nodes belonging to the subset G_1), and will monitor, at a frequency F_2 lower than F_1 , all the links $L'_{A,i}$ for which it has not

received any notification. Thus, the overhead due to the monitoring can be modulated according to the state of the network, and reduced in the case where certain links offer poor performance levels and are therefore little used in the computation of the recommendations. Similarly, the node B will monitor, at a frequency F_1 , all the links $L_{B,j}$ (corresponding to the subset G_1) for which it will have received a notification from one of its server meeting nodes R_{sB} , and at a frequency F_2 lower than F_1 , all the links $L'_{B,j}$ for which it has not received any notification.

The value of the frequencies F_1 and F_2 with which a node will monitor the links of the network depends, for example, on the nature of the network and/or on the traffic engineering objective. When the network is set to be unstable, that is to say when the results of the measurements fluctuate widely, then the monitoring frequency can be greater in order to take into account the variation of the quality of the links between the nodes. In the case of a stable network, the results of the measurements are relatively stable, and the monitoring frequency can be lower since the quality of the links varies very little.

The link states dissemination phase comprises, for example, the following two phases:

1) at a frequency $F_1 = 1/T_1$ a client meeting node sends a server meeting node only the states of the links $L_{A,i}$, $L_{B,j}$ which have been requested of it by this server meeting node during the notification phase,

2) at a frequency $F_2 = 1/T_2$, each client meeting node sends all its link states to all its server meeting nodes.

Thus, the monitoring information dissemination overhead is reduced because only the information of interest is transmitted at given instants.

Figure 4B schematically represents the step of exchanging information for monitoring.

The node A transmits, initially, at a first frequency F_1 , the measurements characteristic of the links to the server meeting node which has asked to receive the measurement. This leads to the generation of a first table K_{AF1} relating to the node A and containing the measurements required by a server

meeting node and performed at the frequency F_1 by the node A towards the nodes of the network which belong to the subset G_1 . The table K_{AF1} therefore contains the results of the measurements on the links that are used most, that is to say which link this node A to the nodes which have the highest probability of serving as intermediate node in the transmission of the data flows. In the example of Figure 4B, the table K_{AF1} for the node A contains the delay values 100 ms, 70 ms, 70 ms, 70 ms, 40 ms for the nodes N14, N22, N41, N43, N44 respectively.

Similarly, the node B transmits to its server meeting node R_{sB} which has asked to obtain information on the state of the links, the results of the measurements that it has performed on the links to the different nodes belonging to the subset G_1 . The table K_{BF1} corresponding to the node B contains the delay values 50 ms, 80 ms, 90 ms, 70 ms, 70 ms for the nodes N14, N22, N41, N43, N44 respectively.

In a second stage, all the quality measurements carried out for all the links, regardless of whether the node belongs to the subset G_1 or to the subset G_2 , are transmitted to all the server meeting nodes R_{sA} , R_{sB} at the frequency F_2 or $1/T_2$, with $T_2 > T_1$. This leads to two new tables K_{AF2} , K_{BF2} , for the node A and for the node B respectively, which contain the measurements of the delays on the links associated with all the nodes of the network.

From these last two tables K_{AF2} , K_{BF2} , each server meeting node will determine the best path with two hops for the nodes A, B and rank the intermediate nodes according to a given criterion. This criterion corresponds, for example, to the shortest delay RTT, to the stability of the network, to its capacity, etc.

A server meeting node does not wait for the dissemination at the frequency F_2 to recompute the paths and send recommendations. The computation is done at the end of each dissemination step; the values which are monitored at the frequency F_2 remain simply identical for iterations at the frequency F_1 if an iteration at the frequency F_2 has not occurred between the two.

In the example given in Figure 4C, in the table K_{AB} resulting from the combination of the tables K_{AF2} , K_{BF2} , it can be seen that the recommended intermediate node N_i is the node $N_{4,4}$ because it offers the smallest delay value

(100 ms). The method will also select a subset G_1 of nodes of the network which are of interest in this choice, that is to say which are likely to become intermediate nodes. Each server meeting node R_s notifies the client nodes of this subset.

A node can be both a server meeting node and a client meeting node.

- 5 Depending on the logical structure that is being considered, a client node can even be its own server, that is to say that it has all the information to compute the best path with two hops with one or more other nodes.

The proportion of nodes selected to be monitored at a higher frequency can be adjusted over time. If the network is very stable, the measurements vary
10 little over time. The proportion may be low, for example, 10% of the best nodes according to the criterion concerned. On the other hand, if the network is unstable, that is to say, if the measurements vary greatly over time, this same proportion of nodes will be great, for example 80%, in order for the server meeting nodes to be able to find the best route with two hops. In the case where a server meeting node
15 notifies a client meeting node to monitor at high frequency all the nodes of the network, the overhead generated will then be equal to that of the prior art modulo the overhead introduced by the notifications (very low, however, since only node identifiers are sent).

A client can also vary F_1 and F_2 according to the stability of the network
20 reported by the server meeting node. If nothing changes in the overlay, F_1 and F_2 can decrease to a minimum imposed by configuration. The decrease may be linear as a function of time. When a disturbance affecting a client is observed by the server meeting node, this can be indicated to it by notification. The client then restores F_1 and F_2 to their initial levels.

25 Without departing from the framework of the invention, the example given for the method according to the invention with two frequencies can be implemented with k frequencies.

In the example which follows, given in Figures 5A and 5B, all the nodes can act as server meeting nodes and client meeting nodes simultaneously (this is
30 the case in the example of organization in grid form). These two architectures are then merged.

Figure 5A schematically represents an exemplary architecture for a server meeting node according to the present invention. A server meeting node comprises at least the following modules:

- a communication module 500,
- 5 • a database 501 comprising the parameters associated with the strategy implemented by the method, for example, the choice of the frequencies to be used, the number of subsets of nodes or subclasses chosen,
- an optimization control module 502 according to the invention,
- a notification mechanism control module 503 which sends the notification
10 messages to the client meeting nodes, module 503 communicates with the module 604 via the module 500 and the module 600,
- a module 504 suitable for controlling the broadcasting of the recommendations stored in a database 505, that a server meeting node transmits to the client meeting nodes, module 504 communicates with the
15 module 606 via the module 500 and the module 600,
- a module 506 for controlling the state of the links, or LSA (Link State ...) which receives the LSAs from the client meeting nodes and transmits the state of the links to a database 507, the state of the links also being known to the optimization module, module 506 communicates with the module 608
20 via the module 500 and the module 600,
- a module 508 for defining and controlling the structure of the overlay network, for example, the module may consist of a distribution of the nodes in the network structure,
- a database 509 containing a table of the nodes of the overlay network
25 which is known to the controller.

Figure 5B schematically represents an architecture of a client meeting node comprising at least the following elements:

- a communication module 600, said communication module 600 of a client meeting communicates with all the nodes of the network, for the working of
30 module 601, and also with the communication module 500 of its meeting server. Symmetrically, the communication module 500 communicates with all the communication module 600 of its meeting clients,

- a module 601 for controlling the optimization of the monitoring of the links and of the dissemination of the information at the node level,
- a database 602 containing, like the database 501, the parameters of the information dissemination strategy,
- 5 • a routing table 603 comprising the routing recommendations transmitted from the server meeting nodes to the client meeting nodes,
- a module 604 for controlling notifications which will recover the notification commands/messages sent by a server meeting node,
- a database 605 comprising notifications sent by a server meeting node,
- 10 • a module 606 for controlling recommendations and a database 607 comprising the tables of recommendations in the choice of nodes to be used to obtain the best path for the transmission of information from a first node to a second node,
- a module 608 for measuring and controlling the state of the links stored in the link states database 609, the control module 608 receives from the optimization controller, information on the adopted dissemination strategy, the information being able to be the value and the number of the frequencies used to scan the node,
- 15 • the function of the monitoring control module 610 is notably, during the monitoring step, to measure the state of the links by taking into account monitoring strategy information transmitted by the optimization controller, this information may be the value of the frequencies which are used for the monitoring.
- 20 • a module 611 for controlling the structure of the overlay network whose function is notably to define the arrangement of the nodes, the nodes being stored in a database 612.
- 25

The method and the device according to the invention notably make it possible to optimize the maintenance of a view of the topology of the overlay network that includes performance indicators associated with the logical links in order to carry out traffic engineering, and thus:

- best exploit the resources available in these different networks,

- reduce the reservation costs by implementing an adaptive routing,
- ensure the routing of the critical flows in the event of failure or threat.

CLAIMS

- 1 – Method for optimizing the dissemination of link quality information in an overlay network comprising a plurality of nodes N_i , N_j , of which a number of nodes are server meeting nodes R_s , at least one client meeting node A and at least one client meeting node B, in which each client meeting node controls and performs measurements on the links that it has with the nodes of the network by using a number of frequency values F_i , the list G_i , of the links to be monitored at a given frequency value being established by the server meeting nodes.
- 2 – Method according to Claim 1, in which a client meeting node also acts as a server meeting node.
- 3 – Method for optimizing the dissemination of link quality information in an overlay network comprising a plurality of nodes N_i , N_j , of which a number of nodes are server meeting nodes R_s , at least one client meeting node A and at least one client meeting node B, said method comprising:
- a step in which a server meeting node R_s determines at least one first subset G_1 of nodes for which the probability of being used as intermediate node for the communications between A and B is high and at least one second subset G_2 of nodes that have a low probability of being used as intermediate node at a given instant, and transmits to the client nodes A and B, a list of the nodes to which the links $L_{A,i}$, $L_{B,j}$ are to be controlled at a first frequency F_1 , the other links being monitored at a second frequency F_2 ,
 - a step of link state dissemination executed at said first frequency F_1 in which only the information concerning the links $L_{A,i}$, $L_{B,j}$ according to the first subset G_1 is transmitted from a node A, B to a server meeting node R_s having requested this information,
 - a step of link state dissemination executed at said second frequency F_2 in which the set of quality information concerning the other links $L'_{A,i}$, $L'_{B,j}$ is returned from the nodes A, B to the server meeting nodes R_s ,

- a step in which each server meeting node determines the best route or the best path to route the communications from the node A to the node B,
- a step in which a server meeting node sends, after each dissemination, a route recommendation to the nodes A, B indicating the routes which offer better properties than the direct connection for joining other nodes.

4 – Optimization method according to Claim 3, characterized in that the best path determined for the client meeting nodes is a route with two hops comprising a single intermediate node.

5 – Method according to Claim 3, characterized in that, to determine a subset G_1 , G_2 , the n nodes which offer the shortest delays between A and B are considered.

6 – Method according to one of Claims 1 or 3, characterized in that, to determine the subset G_1 , a server meeting node uses as criterion the travel time delay RTT between a node A, B and a node of the network.

7 – Method according to one of Claims 1 or 3, characterized in that, to determine the subset of nodes G_1 , a server meeting node uses as criterion the loss rate and/or the bandwidth value.

8 – Method according to one of Claims 1 or 3, characterized in that, to determine the subsets of nodes G_1 , a server meeting node uses as criterion the stability of a link $L_{A,i}$, $L_{B,j}$ or the cost of the link, or the links security level, or, more generally, a composite metric of a set of quality or stability/cost/security criteria.

9 – Method according to one of the preceding claims, characterized in that a server meeting node determines a set of intermediate nodes by using as criterion the value of the round trip time RTT and in that a server meeting node transmits this information to the client meeting nodes.

10 – Method according to one of the preceding claims, characterized in that a network is used that has a structure in which the server meeting nodes are situated on one and the same row and one and the same column as a node A, B or a network of DHT, central server, type.

5

11 – Method according to one of the preceding claims, characterized in that the values of the link monitoring frequencies are varied as a function of the nature of the network, or of the traffic engineering objective, or of events occurring in the network.

10

12- Device for optimizing the dissemination of link quality information in an overlay network comprising a plurality of nodes N_i , N_j , of which a number of nodes are server meeting nodes R_s , at least one client meeting node A and at least one client meeting node B, characterized in that it comprises at least the following elements:

15

For a server meeting node

- a communication module (500),
- a database (501) comprising the parameters associated with the strategy implemented by the method,
- an optimization control module (502),
- a notification mechanism control module (503),
- a module (504) suitable for controlling the broadcasting of the recommendations stored in a recommendations database (505),
- a module (506) for controlling the state of the links and a link states database (507),
- a module (508) for defining and controlling the structure of the overlay network,
- a database (509) containing a table of the nodes of the overlay network which is known to the controller.

25

30 For a client meeting node:

- a communication module (600),

- a module (601) for controlling the optimization of the monitoring of the links and of the dissemination of the information at the node level,
- a database (602) containing parameters of the information dissemination strategy,
- 5 • a routing table (603),
- a module (604) for controlling notifications which will recover the notification commands/messages sent by a server meeting node,
- a database (605) comprising notifications sent by a server meeting node,
- 10 • a module (606) for controlling recommendations and a database (607) comprising the tables of recommendations in the choice of nodes to be used,
- a module (608) for measuring and controlling the state of the links stored in the link states database (609),
- a monitoring control module (610),
- 15 • a module (611) for controlling the structure of the overlay network suitable for defining the arrangement of the nodes stored in a database (612).