

REPLACED SHEETS SG

☒ Description : paged 1 to paged 15
☒ Claim(s) : paged 16 to paged 18
☒ Drawing(s) : paged 1/6 to paged 6/6
☒ Abstract : paged 19 to paged -

SS: 9602697-6
 RTMP'S COPY

- 1 -

A METHOD OF CONTROLLING OVERLOADS
IN A TELECOMMUNICATIONS NETWORK

5 This invention relates to a method of controlling overloads in a telecommunications network.

An overload in a telecommunications network may arise for various reasons. For example, a television programme requesting telephone calls from its audience may cause an
 10 overload. An overload may also occur when a large number of people attempt to make calls to an entertainment or an information service or to a business which has something special on offer. It is clearly desirable to control such overloads.

15 According to this invention, there is provided a method of controlling overloads in a telecommunications network comprising a network of interconnected nodes arranged to provide connections between terminal resources, said method comprising the steps of: upon detecting an initial
 20 failed call setting a counter to an initial value; incrementing the counter upon detecting each further failed call whose call identity is the same as that of the initial failed call or whose call identity belongs to a common set of call identities which includes the call identity of said
 25 initial failed call; decrementing said counter at a predetermined rate; causing said counter to provide an output which is in a first state when the number of calls in the counter rises above a first threshold and which is in a second state when the number of calls in the counter falls
 30 below a second threshold; setting and updating a restriction parameter in accordance with the output of the counter; and restricting calls having said common call identity or whose call identities belong to said common set of call identities, the level of restriction applied in said step of
 35 restricting calls being determined by the value of said restriction parameter.

This Sheet Has Been Replaced



Preferably, in said step of setting and updating the restriction parameter, after each updating of the restriction parameter, during a first time period the restriction parameter is not updated, during a second time period the
5 restriction parameter is varied so as to increase the severity of the level of restriction if the output of the counter changes from the second state to the first state or so as to decrease the severity of the level of restriction if the output of the counter changes from the first state to the
10 second state, and if the second time period expires without a change in state of the output of the counter, the restriction parameter is varied so as to increase the level of restriction if the output of the counter is in the first state or so as to decrease the level of restriction if the
15 output of the counter is in the second state.

Conveniently, said initial failed call and said further failed calls are detected at a first one of said nodes, and said step of restricting calls is performed at one or more nodes which are upstream from said first node with
20 respect to the direction of call set up.

This invention will now be described in more detail, by way of example, with reference to the accompanying drawings in which:

Figure 1 is a block diagram showing some of the
25 switches which form a public telecommunications network;

Figure 2 shows a modification to the public telecommunications network of Figure 1 which includes an additional switch for providing additional services;

Figure 3 is a block diagram showing some of the
30 switches which form an intelligent telecommunications network;

Figure 4 is a block diagram showing processes which are added to a telecommunications network to provide a method of controlling overloads embodying this invention;

35 Figure 5 shows an alternative location for the processes shown in Figure 4;

This Sheet Has Been Replaced



Figure 6 shows a set of counters used for detecting and monitoring overloads;

Figure 7 shows the thresholds used in one of the counters of Figure 6;

5 Figure 8 is a flow chart of an algorithm used for detecting and monitoring an overload; and

Figure 9 is a flow chart of an algorithm used for calculating the parameter used in call restriction.

Referring now to Figure 1, there are shown some of the
10 switches used in a public telecommunications network. The switches shown in Figure 1 comprise two trunk exchange switches 10, 12 and two local exchange switches 14, 16. The trunk exchange switches 10, 12 are just two switches of a fully interconnected network of trunk exchange switches
15 located over a large geographical territory, such as the UK. The local exchange switches 14, 16 are part of a much larger number of local exchange switches which provide access to terminal resources, most of which are telephones or facsimile machines or integrated services digital network (ISDN)
20 terminals. Each local exchange switch may be connected to one, two or three trunk exchange switches. The switches of the telecommunications network are connected by routes 18 which are embodied by suitable traffic carriers such as coaxial copper cables, optical fibre cables and microwave
25 links. The traffic takes the form of voice and other data and also the signalling messages which are used for setting up calls. As well known, signalling messages include forward call set up messages which travel in the direction of call set up and backward call set up messages which travel in the
30 reverse direction.

Figure 2 shows the addition of a switch 20 for providing additional services to the telecommunications network of Figure 1. The switch 20 is connected to the trunk exchange switches. The additional services may include
35 information and entertainment services and also the facility for callers to make free calls or calls charged at the local rate to business numbers. For example, in BT's public

This Sheet Has Been Replaced



telecommunications network in the UK, telephone codes which commence with "0891" and "0898" relate to entertainment and information services. Codes to business numbers which are free to the caller or charged at the local call rate
5 commence, respectively, with "0800" and "0345".

Figure 3 shows some of the switches which form an intelligent network. The switches include a service switching point 30, a trunk exchange switch 32, two local exchange switches 34, 36, and a service control point 38.
10 The service switching point 30 and the trunk exchange switch 32 are part of a network of interconnected service switching points and trunk exchange switches, and the local exchange switches 34, 36 are part of a much larger number of local exchange switches. Each service switching point is connected
15 to the service control point 38. In addition to providing access to the service control point 38, each service switching point also provides the function of a trunk exchange switch. A service switching point can also provide the function of a local exchange switch. The service
20 switching points together with the service control point 38 provide the network with intelligent services. One example of an intelligent service is number translation, which takes place in the service control point.

The trunk switches 10, 12 and the local exchanges
25 switches 14, 16 of Figure 1, the additional services switch 20 of Figure 2, the switches 32, 34 and 36 and also the service switching point 30 and the service control point 38 of Figure 3 are all examples of network nodes. In this specification the term "node" should be construed as any
30 point in a network which is used in setting up a call.

The three networks shown in Figures 1, 2 and 3 represent three examples of telephone networks in which the present invention may be implemented. However, implementation of the present invention is not limited to
35 these three types of network and a mobile telecommunications network represents a further example of networks in which the invention can be implemented.

This Sheet Has Been Replaced



Figure 4 shows an example of the additional processes which are provided at nodes 40, 41 and 42 of a telecommunications network in order to control overloads in accordance with this invention. In the example shown in 5 Figure 4, the node 40 is an exchange switch which has direct access to terminal resources while the nodes 41 and 42 are upstream from node 40 with reference the direction of call set up. Thus, the nodes 40, 41 and 42 of Figure 4 may correspond to the exchange switches 16, 12 and 10 shown in 10 Figure 1, or the exchange switches 16, 12 and the additional services switch 20 shown in Figure 2. In the case of the intelligent network of Figure 3, the node 40 could be the local exchange switch 36, the node 41 could be the trunk exchange switch 32 and the node 42 could be the service 15 switching point 30. The processes for overload control are implemented by modifying the software which controls the nodes. These modifications will now be described in general terms for the nodes 40, 41 and 42.

The software of the node 40 includes a resource access 20 control module 44 which controls access to terminal resources. The module 44 is of conventional design but modified to provide an output signal 45 in the event of call failure. The output signal 45 gives the call identity of a failed call. The normal criterion for registering a call as 25 a failed call is that the terminal instrument is engaged or unobtainable. However, if desired, other criteria may be used. For example, an excessive delay in setting up a call could also be classified as a failure.

The signal 45 indicating call failure is supplied to 30 an overload detection and monitoring module 46. The module 46 provides an output signal 47. The module 46 may be arranged simply to detect overloads to individual full length called numbers. As will be explained in more detail below, the module 46 may also be arranged to detect overloads 35 falling within a common set of call identities. A set of call identities may comprise a number of full length called party numbers, or a number of full length calling party

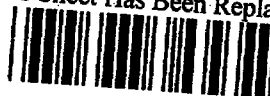


numbers, or all the numbers of a particular service such as the "0800" service mentioned above, or even all the called party numbers accessed by an exchange switch. The data in signal 47 comprises an indication of overload and the call
5 identity or the set of call identities causing the overload. The indication of overload can have only two states, namely, overload and no overload. The signal 47 forms part of the backward call set up message. Where the overload is caused by calls to a single full length called party number, the
10 backward call set up message already contains the call identity. Consequently, the only modification to the conventional backward call set up message is the addition of an extra bit of data which has a value of binary "1" for an overload and a value of binary "0" where there is no
15 overload. Where the overload relates to a set of call identities, the backward set up message must be modified to specify the set of call identities.

In the module 41, the signal 47 is supplied to a software module 48 for setting and updating the restriction
20 parameter. The restriction parameter specifies the level of call restriction which is to be applied in the node 41 to calls destined for node 40. There are various methods of applying call restriction. In the present example, call restriction is applied by proportional blocking. Thus, the
25 restriction parameter specifies the proportion of calls which are blocked and consequently the level of restriction increases with the value of the restriction parameter. Alternatively, the restriction parameter could specify the proportion of calls which are allowed with the result that
30 the level of restriction would increase as the value of the restriction parameter falls. Another method of applying call restriction is call gapping in which each call is followed by a gap interval during which all calls are blocked.

The restriction parameter is supplied by the module 48
35 to a module 49 which applies call restriction. The module 49 also identifies failed calls and supplies the call identities of the failed calls to an overload detection and monitoring

This Sheet Has Been Replaced



module 53, which is identical to the module 46. In the module 41, a call is identified as a failed call if it is blocked because of call restriction or if it fails due to a cause occurring at node 41 or a cause such as no circuits available occurring between nodes 41 and 40. If the call fails at node 40, it is not identified as a failed call in module 49 because such a call will be identified as a failed call in the module 44.

The module 53 supplies an output signal 50 indicating overload or no overload to a software module 51 located in node 42. The module 51 sets and updates the restriction parameter for controlling the level of restriction to be applied by node 42 to calls destined for node 41. The restriction parameter is supplied to a module 52 which implements call restriction. The module 51 is identical to the module 48 and the module 52 is identical to the module 49 except that the module 52 does not detect failed calls.

As will be explained in more detail below, the module 48 progressively increases the restriction parameter when the output signal from the module 46 indicates the presence of an overload, thereby increasing the proportion of calls which are blocked, and progressively decreases the restriction parameter when the output signal from the module 46 indicates there is no overload, thereby decreasing the proportion of calls which are blocked. Consequently, the node 40 oscillates in and out of an overload state until the cause of the overload subsides. As a result, the rate at which the node 41 sends call set up messages to the node 40 should be close to the rate at which calls can be completed successfully. By preventing calls which have a low chance of success from reaching node 40, there is removed the risk that such calls will interfere with other calls which have a high chance of success.

The module 51 operates in a similar manner to the module 48. Consequently, if there is an overload at node 41, the process of increasing and decreasing the restriction parameter in module 51 will ensure that the node 41

This Sheet Has Been Replaced



oscillates into and out of an overload state. However, in most telecommunications networks, the node 41 will be only one of several nodes which are sending call set up messages to the node 40 for calls having the call identity or falling
5 within the set of call identities which are causing the overload. Where the overall proportion of failed and blocked calls is comparatively modest, the node 41 will not be in an overload condition. Where the proportion of failed and blocked calls is severe, the node 41 will also be in an
10 overload condition and so call restriction will also be applied at node 42. Thus, with increasing severity of the overload condition, call restriction is applied progressively further away from node 40.

Because the software modules for controlling overload
15 in node 41 and the other nodes which send call set up messages to the node 40 are the same, the level of call restriction applied in the various nodes in the event of overload will be similar. However, because the signal from the module 46 indicating overload is transmitted in the
20 backward call set up message and not continuously, there will be some variation in the level of restriction. Similarly, the level of restriction applied in the nodes which send set up messages to the node 41 will be similar but not identical.

Referring now to Figure 5, there is shown a
25 modification to the arrangement of Figure 4 in which the modules 48 and 51 for setting and updating the restriction parameters are located at the nodes 40 and 41 rather than at the nodes 41 and 42. Consequently, the level of restriction in all of the nodes which send call set up messages to the
30 node 40 will be identical. Likewise, the level of restriction in all of the nodes which send call set up messages to node 41 will be identical. Thus, with proportional blocking this arrangement provides total fairness for the callers. However, this arrangement has the disadvantage that the
35 backward call set up messages have to specify the value of the restriction parameter and this requires more bits of data

This Sheet Has Been Replaced



than that required simply to specify the presence or absence of an overload.

Figures 4 and 5 each show an arrangement in which call restriction is applied at two nodes along a call set up path. If it is desired to apply call restriction at only one node along the path, this may be achieved by omitting the overload control arrangement of node 41 as shown in Figure 4 or Figure 5. On the other hand, if it is desired to apply restriction at more than two places along the call set up path, this may be achieved by repeating the overload control arrangement of node 41 as shown in Figure 4 or Figure 5 as many times as desired.

Figure 4 and Figure 5 each show an arrangement in which the overload is detected initially at the node which accesses the terminal resources and this usually represents the most desirable point to do this. However, if desired, the initial detection may occur at an upstream position with regard to call set up. For example, the nodes 40, 41 and 42 of Figure 4 could correspond to the switches 12, 10 and 14 of Figure 1, or the switches 20, 10 and 14 of Figure 2 or the service control point 38, the service switching point 30 and the local exchange switch 34 of Figure 3.

By way of modification, call restriction may be applied at the node where the overload is detected. For example, in the intelligent network of Figure 3, overload detection and call restriction may both occur at the service switching point 38.

There will now be described the two algorithms which are used, respectively, for detecting and monitoring the overload and for setting and updating the restriction parameter. These will be described initially with respect to detecting and controlling an overload to a full length called party number.

Referring now to Figure 6, the algorithm for detecting and monitoring overloads uses a set of counters 59. As illustrated in Figure 7, each of these counters has an initial value, which is normally 0 and four thresholds,

This Sheet Has Been Replaced



namely, a release threshold 60, an overload abatement threshold 61, an overload onset threshold 62 and a maximum threshold 63. The flow chart of the algorithm for detecting and monitoring overloads is shown in Figure 8 and this algorithm will be described with reference to this figure. In a step 21, when an initial call failure to a full length called party number is detected, one of the counters 59 is associated with that number and set to its initial value. Then, in a step S22, a check is made to determine if a further failed call has been detected to the number. If a further failed call has been detected to the number, in a step S23, the counter is incremented by 1. Thus, the counter is incremented each time a failed call is detected to the number. The counter is also decremented at a constant rate. In order to achieve this, in a step S24, a check is made to determine if it is time to decrement the counter. If it is time to decrement the counter, it is decremented in a step S25. If the count rises above the overload onset threshold, the counter goes into its overload state. The counter is prevented from counting above its maximum threshold. When the value of the count falls below the overload abatement threshold, the counter goes into its no overload state. When the count falls below the release threshold, the counter is no longer associated with the called party number. In order to achieve this, in a step S26, a check is made to determine if the count has fallen below the release threshold. If it has fallen below the release threshold, in a step S27, the counter is released from the called party number.

By providing separate onset and abatement thresholds, the counter has hysteresis. These thresholds should be set close enough so that the overload is detected with sufficient precision but far enough apart to give adequate hysteresis. The overload onset threshold should be set high enough to prevent the counter from going into the overload state when calls are failing for innocuous reasons but it should be set low enough to provide sensitivity to genuine overloads. The maximum threshold should not be set so high that there is an

This Sheet Has Been Replaced



undue delay in reducing the level of call restriction after a sudden surge in calling rate.

Each of the software modules which sets and updates the restriction parameter contains a number of copies of the algorithm for doing this. A flow chart for this algorithm is shown in Figure 9. When one of these modules receives an indication of overload, it associates one of the copies of the software algorithm with the called party number which is suffering the overload. Then, and referring to Figure 9, in a step S1, the restriction parameter is set to an initial value. Consequently, call restriction, which in the present example is achieved by proportional blocking, commences at a level specified by this initial value.

Then, in a step S2, two timers are started for timing intervals t_1 , t_2 . Then, the algorithm enters and remains in a step S3 until the elapsed time t is equal to the preset value t_1 . Thus, during the time interval which ends when the elapsed time is t_1 , no change is made to the restriction parameter.

The algorithm then enters a step S4 in which the overload status is monitored and the elapsed time is compared with the second preset value t_2 . If a change in the overload status occurs when the elapsed time is between t_1 and t_2 , the restriction parameter is updated in a step S6. In the step S6, the restriction parameter is increased if the overload status has changed from no overload to overload and it is decreased if the overload status has changed from overload to no overload. After step S6, the algorithm passes to a step S8 which is described below.

If the overload status does not change in the interval when the elapsed time is between t_1 and t_2 , the restriction parameter is updated in a step S7. In step S7, the restriction parameter is increased if there is an overload and it is decreased if there is no overload. After step S7, the algorithm continues with step S8.

This Sheet Has Been Replaced



When the restriction parameter is increased, the new value b_n is calculated from the old value b_{n-1} by the following equation:

$$5 \quad b_n = (1 - \alpha) + \alpha \cdot b_{n-1} \quad \dots (1)$$

In equation (1), the constant α will usually be chosen close to a value of 1.

When the restriction parameter is decreased, the new value b_n is calculated from the old value b_{n-1} by the following equation:

$$b_n = b_{n-1} - \beta \quad \dots (2)$$

15 In step S8, the restriction parameter is compared with a threshold which is less than the initial value of the restriction parameter. If its value is below the value of this threshold, call restriction ceases and the algorithm is no longer associated with the full length called party number. If the restriction parameter is above this threshold, the two timers are restarted in a step S9 and the algorithm then returns of step S3.

The algorithms for detecting and monitoring overloads and for setting and updating the restriction parameter have been described above with reference to monitoring and controlling overloads to full length called party numbers. There will now be given an example which shows how these algorithms may be modified for monitoring and controlling calls when arranged by sets of call identities. In the following example, the call identities are the complete group of call identities for called party numbers used in BT's UK public telecommunications network.

In this example, the complete group of called party identities is arranged as three collections of sets of call identities. Each collection comprises the complete group of call identities arranged either as a single set of call identities or divided into a plurality of non-intersecting

This Sheet Has Been Replaced



sets of call identities. Specifically, the first collection comprises a single set of all the individual call identities. The second collection comprises an individual set for each of the service codes "0345", "0800", "0891" and "0898", and a
5 fifth set for all the remaining call identities. The third collection comprises a single set for the two full length call party numbers "0891 000000" and "0891 000001", and an individual set for each of the remaining full length called party numbers. In this example, the two number "0891 000000"
10 and "0891 000001" are two numbers used for television voting. The first number is the one used for a yes vote and the second number is the one for a no vote. These two numbers are put together in a single set as it is clearly essential for exactly the same level of call restriction to be applied
15 to each of these numbers.

In each copy of the module for detecting and monitoring overload, a number of counters is dedicated to each collection. Specifically, a single counter is dedicated to the first collection, four counters are dedicated to the
20 second collection, and three counters are dedicated to the third collection. Thus, the sets of each collection and the number of counters dedicated to each collection is as shown in the following table.

25 <u>Sets in Collection</u>	<u>Number of Counters</u> <u>for collection</u>
Single set consisting of all called party numbers	1
30 0345, 0800, 0891, 0898, remainder	4
Each full length number and (0891 000000 + 0891 000001)	3
35	

As may be observed from the table, the single set in the first collection is divided into five sets in the second

This Sheet Has Been Replaced



collection, and each set in the second collection is divided into a large number of sets in the third collection. For example, the set in the second collection for the "0345" service code is divided in the third collection into a large number of sets, each of which comprises a full length called party number which commences with "0345".

Not all of the sets of call identities are monitored. In the present example, the set in the second collection comprising the remaining full length call party number is not monitored.

With the call identities arranged in sets as set out in the table above, each module for detecting and monitoring overload operates as follows. When a call arrives and fails at the node in which the module is located, an indication of the call identity of the failed call is sent to the module. In each collection of sets of call identities, the call identity for the failed call will belong to a unique set of call identities. The action taken will then depend upon the state of that set. If the set is one that is not monitored, which is the case for the set containing the remaining called party numbers in the second collection, no action is taken. If a counter is already associated with the set containing the call identity for the failed call, then that counter is incremented. If no counter is presently associated with a set containing the call identity of the failed call, and there is a free counter, then the free counter is associated with that set and set to its initial value.

In each module for detecting and monitoring an overload, an indication of overload status is sent to the appropriate module or modules for setting and updating the restriction parameter in the following manner.

When a counter initially passes into the overload state, if no other counter is active or no active counter has made the initial transition into the overload state, an indication is sent to the module or modules for setting and updating the restriction parameter. This indication contains the overload status and data to identify the set of call

This Sheet Has Been Replaced



identities which are being monitored. In the or each module for setting and updating the restriction parameter, one of the copies of the algorithm is then associated with the set which is monitored and the algorithm then calculates the
5 restriction parameter in the manner which has been described above.

If a counter makes its initial transition to the overload state at a time when another counter is active and has previously made its initial transition to the overload
10 state, the following procedure is followed. If the two counters are monitoring sets in the same collection, the new counter to go into its overload state sends an indication of its overload status and the details of the set which is being monitored to the module or modules for setting and updating
15 the restriction parameter. The or each module for setting and updating the restriction parameter then associates one of its spare algorithms with the new set and the restriction parameter is calculated in the manner described above. However, if the two counters are monitoring sets in different
20 collections and one monitored set contains the other monitored set, then the overload status is fed back only for the larger monitored set. Thus, for example, if the first counter which goes into an overload state is associated with a full length number in the third collection beginning with
25 the service code "0345", then call restriction will be applied initially just on this full length number. However, if subsequently a counter for the second collection which is associated with the service "0345" goes into an overload state, then overload restriction will be applied to all calls
30 which commence with this service code.

This Sheet Has Been Replaced





CLAIMS

1. A method of controlling overloads in a
5 telecommunications network comprising a network of
interconnected nodes arranged to provide connection between
terminal resources, said method comprising the steps of:

upon detecting an initial failed call setting a
counter to an initial value;

10 incrementing the counter upon detecting each further
failed call whose call identity is the same as that of the
initial failed call or whose call identity belongs to a
common set of call identities which includes the call
identity of said initial failed call;

15 decrementing said counter at a predetermined rate;
causing said counter to provide an output which is in
a first state when the number of calls in the counter rises
above a first threshold and which is in a second state when
the number of calls in the counter falls below a second
20 threshold;

setting and updating a restriction parameter in
accordance with the output of the counter; and

restricting calls having said common call identity or
whose call identities belong to said common set of call
25 identities, the level of restriction applied in said step of
restricted calls being determined by the value of said
restriction parameter.

2. A method as claimed in claim 1, comprising the further
30 steps of:

upon detecting said initial failed call associating
said counter with the call identity of the initial failed
call or said common set of call identities which include the
call identity of the failed call; and

35

This Sheet Has Been Replaced





- 17 -

releasing said counter from its association with the call identity of the failed call or said common set of call identities which include the call identity of the failed call when the number of calls in the counter falls below a threshold value which is set less than said initial value.

3. A method as claimed in claim 1 or claim 2, in which, in said step of setting and updating the restriction parameter, after each updating of the restriction parameter, during a first time period the restriction parameter is not updated, during a second time period the restriction parameter is varied so as to increase the severity of the level of restriction if the output of the counter changes from the second state to the first state or so as to decrease the severity of the level of restriction if the output of the counter changes from the second state to the first state, and if the second time period expires without a change in state of the output of the counter, the restriction parameter is varied so as to increase the level of restriction if the output of the counter is in the first state or so as to decrease the level of restriction if the output of the counter is in the second state.

4. A method as claimed in claim 3, in which, in said step of setting and updating the restriction parameter, the restriction parameter is set to an initial value when the output of the counter initially passes into the first state, call restriction being deactivated when the restriction parameter passes through a threshold in the direction of decreasing call severity, said threshold indicating a lower level of call restriction than that of said initial value.

5. A method as claimed in any one of claims 1 to 4, comprising the further steps of:
establishing a group of call identities relating to calls handled by said telecommunications network;

This Sheet Has Been Replaced





- 18 -

dividing said group into at least one collection of non-intersecting sets of call identities, the or each collection comprising at least one set of call identities; and

5 . dedicating a respective set of counters to the or each collection of sets of call identities, each set of counters comprising at least one counter.

6. A method as claimed in any one of claims 1 to 5, in
10 which said initial failed call and said further failed calls are detected at a first one of said nodes, and said step of restricting calls is performed at one or more nodes which are upstream from said first node with respect to the direction of call set up.

15

7. A method as claimed in claim 6, in which backward call set up messages from said first node to said one or more upstream nodes are used to convey information relating to the overload resulting from said failed calls.

20

8. A method as claimed in claim 7, in which said step of setting and updating said restriction parameter is performed at the or each of said one or more upstream nodes.

25 9. A method as claimed in claim 7, in which said step of setting and updating said restriction parameter is performed at said first node.

This Sheet Has Been Replaced

