

METHOD OF CONTROLLING A TELECOMMUNICATIONS NETWORK

This invention relates to a method of controlling a telecommunications network which includes at least one local exchange connected by routes to one or more parent
5 trunk exchanges each being one of a plurality of trunk exchanges interconnected by trunk routes. It is particularly concerned with the control of focused overloads in such networks.

Commonly a local exchange is associated with a home
10 exchange, through which incoming calls to the local exchange are routed, and a security exchange, through which outward calls from the local exchange are routed, in normal operation. The security exchange is so called because it can also be used to route incoming calls to the local
15 exchange if the home exchange fails. The security and home exchanges are collectively referred to as the parent exchanges of the local exchange. Exchanges other than the parent exchanges are referred to as far-end trunk exchanges of the local exchanges.

20 Near real-time network traffic management (NTM) is an essential component of network management if optimal traffic performance in terms of call throughput is to be ensured. To give an indication of the volume of traffic which may be involved, BT's trunk network in the United
25 Kingdom currently handles approximately six million call attempts per hour during the busy periods which is equivalent to 1,700 call attempts per second. Given such a volume of traffic it is essential that any network difficulties are detected and controlled as quickly as
30 possible. For example, difficulties are often encountered by network traffic managers due to abnormal traffic patterns which can be caused by events such as 'phone-ins, tele-votes and public holidays (for example Christmas Day and New Year's Eve/Day). In all these cases traffic in the
35 network varies widely from the normal level, sometimes

This Sheet Has Been Replaced



quite spectacularly, and the network must be controlled to maintain the best overall network performance.

With the introduction of digital switches such as System X it is possible to monitor closely the performance of each exchange and the routes between them and to the subscribers. BT's Network Traffic Management System (NTMS) currently receives statistics on upwards of 37,000 routes from 490 exchanges in the UK every five minutes, this measurement period being chosen to be a long enough period to be able to obtain a statistically reliable measurement of the network performance whilst being short enough to allow effective real-time control of the network.

(It might be noted that other communication networks carry different traffic patterns, and can be equipped with different network hardware and that therefore other measurement periods are sometimes found appropriate. For instance a lot of "Telcos", or telecommunications companies, use a measurement period of 15 minutes.)

The information received by the NTMS is processed to provide CCITT recommended parameters. The parameter values are compared with thresholds to determine if any difficulties exist on the monitored network elements.

Usually the first indication of a network problem is when an 'exception' is displayed on a wall-board, or on a graphical interface at an individual manager's workstation, at a Traffic Management Centre. Exceptions are those parameter values, calculated from network element measurements, which deviate sufficiently from a predetermined threshold for that value. The exceptions are ranked in a priority order with the top 20 displayed. However, due to the manner in which the thresholds are set by the network traffic managers, some exceptions do not necessarily indicate a difficulty as thresholds are percentage-based and set a value which ensures all potential difficulties are captured. This results in exceptions being displayed that are occasionally spurious

or insignificant. The exceptions therefore need to be examined in more detail to determine if a real difficulty exists and whether it warrants any action. To help in this activity several information sources are currently used by
5 the network traffic managers.

The NTMS provides near real-time surveillance and monitoring of the network's status and performance. It provides the network traffic managers with information to enable them take prompt action to control the flow of
10 traffic to ensure the maximum utilization of the network in all situations. The NTMS allows network traffic managers to look at the raw statistics as well as derived generic parameters and to compare traffic patterns over the last few measurement periods to isolate any trends.

15 An On-Line Traffic Information System (OTIS) takes the measurement of statistics from the NTMS system and processes them to provide summarised historical data for daily and weekly traffic patterns. This system allows the network traffic managers to examine historical traffic
20 patterns to detect any radical shifts in traffic.

A data management system provides the network traffic managers with an up-to-date copy of the routing tables at all trunk exchanges. This information is used to check the routes to which calls can be routed, which controls are in
25 force and the routing algorithms being used.

There is also a broadcast speaker facility which connects the world-wide network management centre to all the regional centres.

Once a potential difficulty has been detected,
30 acknowledged and analyzed, it is characterised and a decision made whether to control it using the available range of expansive and restrictive controls to either allow alternative traffic paths through the network or to restrict or block call attempts to particular areas,
35 respectively. The situation must then be monitored to ensure the controls are having the desired effect and that

they are removed as soon as a problem has been dealt with effectively.

It is important that the controls applied are sufficient to suppress the problem yet not so severe that
5 they cause their own problems by congesting other parts of the network or by blocking potentially successful calls.

One class of exception associated with telecommunications networks is the focused overload. The All Circuits Engaged (ACE) CCITT parameter may be monitored
10 in order to determine when a focused overload occurs but this has been found to be not very satisfactory at present and will be less so in the near future, when call failures will only be shown (on a wallboard) with respect to the last of several routes tried. Currently, all the routes
15 tried are shown and focused overloads give rise to a graphic "star" effect on a wallboard.

A preferred method of detecting a focused overload is as described in a British patent application filed by the present applicant on 12th May 1992, being allocated number
20 9210173.2 (ie. GB9210173.2), and in the equivalent United States patent application number 115689 filed on 3rd September 1993, in the name CRABTREE, and the subject matter disclosed therein is incorporated herein by reference.

25 The present invention is, however, directed to a method of controlling a network once a focused overload condition has been detected by whatever method.

According to the present invention there is provided a method of controlling a network, which includes at least
30 one local exchange connected by one or more routes to one or more parent trunk exchanges, each parent trunk exchange being connected by trunk routes to a plurality of further trunk exchanges, in which call restriction is applied to one or more of the exchanges of the network on detection of
35 a focused overload with respect to said local exchange, the method being characterised in that call restriction of



numbers of the local exchange is applied to at least one exchange of the network, on detection of a focused overload with respect to the local exchange, but is not applied to any parent trunk exchange of that local exchange.

5 The call restriction may be applied for instance to a far-end trunk exchange of the local exchange whenever the Answer Seizure Ratio (ASR) from the far-end trunk exchange to a parent trunk exchange of the local exchange has stayed below a threshold for the duration of a measurement period.
10 This is one of the criteria for applying call restriction disclosed in the present applicant's above mentioned GB9210173.2. However, clearly other criteria could be used, in combination or separately. The important aspect is that, by applying call restriction beyond the relevant
15 parent trunk exchanges, it is possible to control the level of call restriction applied in a much more versatile manner.

 "Call restriction" is a term used herein to indicate a network control step which limits the number of calls
20 made by blocking attempts at a point in the network. It can be implemented in different ways, including call gapping, in which fixed length intervals are enforced during which no call can be made. For instance, five second call-free intervals may be imposed. Another form
25 of call restriction is proportional call blocking, in which only a percentage of bids are accepted. For instance, every tenth call might go forward. A third approach is to allow a maximum number of calls to go through in any one-time period, such as one call every five seconds.

30 The use of call restriction according to known techniques, at parent trunk exchanges, means that all calls incoming to those parent trunk exchanges will be subject to the same call restriction levels. But the known, fixed rate, call restriction, for instance 5 second call gapping,
35 applied "across the board" in this manner has been recognised as sometimes inappropriate in making the present



invention. It means that calls arising on a route which has contributed nothing to the focused overload, that is, where the traffic remains at normal levels, are restricted in the same way that calls on a route which has caused the focused overload are call gapped. Hence, in a particularly advantageous embodiment of the invention, call restriction is applied so as to restrict calls on different routes to the local exchange by different amounts, or at different rates. This could even mean that one or more of the routes is not restricted at all.

It is clearly preferable that the call restriction levels applied to routes contributing to different extents to a focused overload should be tailored for each route to reflect the extent to which that route is so contributing. This might be done by making the call restriction level for a route inversely proportional to that route's contribution.

In more detail, a call restriction level might suitably be set as follows. A value B_1 is selected as the total bids per measurement period which it is desired should be made to an overloaded number. A value B_0 is estimated as the total number of bids actually made, without call restriction, as a result of a focused overload. On each route to be call restricted a number of bids to be made on that route after call restriction has been applied, b_g , is allocated, this being determined by the proportion of B_0 carried by that route. The call restriction level on that route is then set appropriately. For instance, in the case of call gapping, the call gap rate in seconds for a single route, g , is then given by :

$$g = \frac{\text{measurement period in seconds}}{b_g}$$

A suitable measurement period might for instance be 5 minutes.

This Sheet Has Been Replaced



If proportional call blocking is substituted for call gapping, the call gap route value, g , can simply be used as a measure of the proportion of calls which will be blocked, for instance nineteen out of every twenty potential bids.

5 As discussed in the above mentioned GB9210173.2, and US 115689/93, the criteria used for applying call restriction, in a method according to the present invention, might include a rise in the Bids per Circuit per Hour (BCH) from a parent exchange to the local exchange.
10 The BCH parameter on a route between a pair of exchanges gives a normalised indication of the number of call attempts down that route and will generally stay well below the maximum BCH that can be handled by the exchanges under normal conditions. It will however tend to move to and
15 above that value during a focused overload.

The BCH will also rise due to an increase in country-wide traffic to a large number of subscribers served by a given local exchange, for example when a disaster befalls an area, but this will not necessarily result in a focused
20 overload if a sufficiently high number of calls continue to be connected. The inclusion of a test of the ASR, together with a test of the relevant BCH, serves to distinguish between these two scenarios. Preferably, therefore, call restriction is applied according to the criteria disclosed
25 in GB9210173.2 and US 115689/93. That is, it is applied when the BCH from a parent exchange to the relevant local exchange rises above a first threshold, and the ASR on a primary traffic route from that parent exchange to the
30 local exchange stays below a second threshold for the duration of a measurement period. (A primary traffic route is the route first tried for a call in a network, all alternative routes then being secondary.) This takes account of the fact that, during a focused overload, the average ASR on the trunk primary traffic routes into the
35 home exchange drops very sharply and can remain low for a number of measurement periods before slowly returning to



its normal value whereas a general increase in traffic to a local exchange does not have this effect.

An embodiment of the present invention will now be described, by way of example only, with reference to the
5 accompanying drawings in which:

Figure 1 shows a schematic representation of part of a telecommunications network to which call restriction might be applied;

Figure 2 shows in more detail aspects of the network of
10 Figure 1;

Figure 3 shows a schematic representation of a control system for implementing a method of the present invention; and

Figure 4 shows a schematic representation of a control
15 system for implementing a method of the present invention, using direct control of a network by a computer.

Referring to Figure 1, a telecommunications network to which call restriction might be applied comprises a local exchange 1, together with other exchanges connected
20 directly or indirectly. The other exchanges fall into categories comprising directly connected local exchanges 2, the home exchange 3 of the local exchange 1, the security exchange 4 of the local exchange 1, indirectly connected local exchanges 5 and far end trunk exchanges 6.

25 Routes between the exchanges that are monitored for the purposes of controlling a focused overload fairly, routes 7, are depicted in Figure 1 by arrows, the base of the arrow indicating in each case where call restriction control would be applied in a method according to an
30 embodiment of the present invention. Hence, call restriction would be applied at the directly connected local exchange end of the route 7 between a directly connected local exchange 2 and the local exchange 1 which is subject to a focused overload. In the case of other
35 exchanges, call restriction might be applied at the far end trunk exchanges 6 or at the indirectly connected local

exchanges 5, but not at a parent exchange 3, 4. There are two unmonitored routes 8, between the parent exchanges 3, 4 and the local exchange 1 subject to a focused overload.

In practice, once a focused overload has been detected, to control that overload in a fair manner, the following steps are carried out: -

(i) an estimate of the total increase in bids due to the overload, B_o , is made according to the following;

$$B_o = \sum_{D_n} (b_c - b_a) + \sum_{I_n} (b_c - b_a) + \sum_{F_n} (b_c - b_a)$$

where

b_c is the number of bids on a route in the measurement time period immediately after the overload is detected, and b_a is the number of bids on that route in the measurement time period immediately before the overload is detected.

In the notation used here, " D_n " refers to a sum over routes from all the directly connected local exchanges 2, " I_n " refers to a sum over routes from all the indirectly connected local exchanges 5 and " F_n " refers to a sum over routes from all the far end trunk exchanges 6.

(ii) to ensure that there is more than enough traffic offered to the overloaded number to avoid for example loss of revenue, an estimate is made of the maximum number of bids to the overloaded number that could be handled in a measurement period, based on average call duration and the number of customer lines that can be answered simultaneously, this number then being multiplied upwards, for instance typically in the UK public network by a factor in the range from 4 to 10, to arrive at a value B_i for total bids to the customer.

(iii) on each outgoing route from a far end trunk exchange 6, an indirectly connected local exchange 5 or directly connected local exchange 2, the number of bids after call restriction has been applied, b_r is given by

$$b_g = \frac{(b_i - b_g)}{B_0} B_i$$

5 (iv) finally, if call gapping is the restrictive technique to be used, the call gap rate, g , in seconds, on each outgoing route is given by

$$g = \frac{300}{b_g}$$

10 (The measurement period is set at five minutes here, to give the constant 300 which is the number of seconds in a five minute period.)

Hence, call gap rates depending on the contribution a route makes to a focused overload are applied to achieve a
15 fairer distribution of successful calls in a case of a focused overload.

As mentioned above, if the restrictive technique is to be proportional call blocking, then " g " can instead be used as a measure of the proportion of call attempts to be
20 blocked.

In practice, there may be routes which have negligible or no detectable traffic contributing to a focused overload, which would lead to an infinite or ridiculously high call restriction level, and it may be preferable to
25 set a maximum restriction level which might be used in these cases.

In general, however, it can be seen that the call restriction level on each outgoing route 7 is inversely proportional to the excess traffic generated by that route.
30 Therefore, the higher the bid rate, the lower the call restriction, that is, the lower the call gap rate or the proportion of calls blocked, ensuring that traffic allowed through is proportional to the traffic offered. Because the call restriction levels are selected on the basis of
35 offering a total number of bids to the customer which is, say, 4 to 10 times higher than the maximum which can be

dealt with, the above also ensures that the total traffic offered to the customer via the local exchange 1 is always in excess of the traffic they can handle.

Referring to Figure 2, a telecommunications network may generally comprise a number of digital main switch units (DMSUs) -trunk exchanges - of which only five are shown for clarity and are referenced 102, 104, 106, 108 and 110. These are equivalent to the parent and far end trunk exchanges, 3, 4, 6, of Figure 1. Subscribers' customer premises equipment, of which only two sets are shown, referenced 112 and 114, are connected to these DMSUS, 102 to 110, via respective digital local exchanges (DLEs) 116 and 118.

One DLE 116 is connected to a DMSU 102 through which incoming calls to the subscribers attached to the DLE 116, including subscriber 112, are routed. The DMSU 102 is therefore a home trunk exchange for the DLE 116, equivalent to the home exchange 3 of Figure 1. The DLE 116 is also connected to a second DMSU 104 through which outgoing calls from the DLE 116 are routed. This second DMSU 104 is a security exchange since incoming calls to the exchange 116 can be routed through it should the home trunk exchange DMSU 102 fail, and it is therefore equivalent to the security exchange 4 of Figure 1.

Similarly, there is a home exchange DMSU 108 for the other digital local exchange 118 shown in Figure 2.

The home and security exchanges 102, 104 are thus the parent exchanges associated with the first DLE 116. Those exchanges other than the parent exchanges, the DMSUs referenced 106, 108, 110 in Figure 2, are the far-end trunk exchanges of the first DLE 116 and therefore equivalent to the far-end trunk exchanges 6 of Figure 1.

If a large number of calls are attempted to a subscriber's customer premises equipment 112 within too short a period, the digital local exchange 116 may not be able to handle the calls and in extreme cases the DLE might

also go into overload. This will block unrelated traffic in the area around the destination exchange. This is a focused overload which needs to be controlled if congestion is to be avoided.

5 A method of controlling a telecommunications network such as that illustrated in Figure 2 will now be described with additional reference to Figure 3 which shows a network control system implementing the method of the present invention. This method is also disclosed in the above
10 mentioned GB9210173.2, and US 115689/93, and is merely repeated herein for completeness.

Referring to Figure 3 the network of Figure 2 is denoted by box 202. Every five minutes, which is the selected measurement period of the network of Figure 2 in
15 this case, a set of statistics is generated by the digital exchanges of the network 202 which is processed by an NTMS system 204 to provide generic measurement values including those of the CCITT recommendation. These parameters are input to a run-time system 206 which applies rules to the
20 received parameters from the NTMS 204 by means of an appropriately coded expert system. The run-time system 206 provides recommendations to aid a network traffic manager 208 to detect and control focused overload problems in the network 202 according to the method of the present
25 invention.

The NTMS 204 is an extensive and complex system which collates incoming data. It is, in the British PSTN, provided with huge visual screens and "dumb" terminals, for use by operators, managers or others. Data can be
30 processed by the NTMS or pulled off into workstations to do analysis and output recommendations or controls. In the following, data is pulled off into a "run-time system" 206 provided at a "Sparc" workstation.

The run-time system 206 monitors data coming from each
35 local exchange and determines when the BCH along the route from the home trunk exchange 102 of a local exchange,

assumed in this case to be local exchange 116 of Figure 2, -
exceeds 30 and also the BCH averaged over trunk primary
traffic routes to that home exchange 102 have increased by
40% since the last measuring period. When these conditions
5 are satisfied, the run-time system supplies a
recommendation to the network traffic manager 208 to apply
focused overload controls with respect to the local
exchange 116 affected.

The run-time system 206 also provides recommendations
10 to the network traffic manager 208 based on ASR conditions.
In particular, call restriction will be applied with
respect to numbers of the local exchange 116 once the ASR
from the parent trunk exchange to the local exchange 116
has remained at less than 20% for the duration of a five
15 minute measuring period and the BCH along the route from
the parent trunk exchange to the local exchange 116 is
greater than 10.

The run-time system 206 will also recommend call
restriction to numbers of the local exchange 116 at far-end
20 trunk exchanges 106 to 110 of the local exchange 116 when
the ASR along the route from the far-end trunk exchange to
a parent exchange 102, 104 of the local exchange 116 falls
below 45%.

- As a focused overload eases up, the run-time system
25 206 will recommend to the network traffic manager 208 when
the call restriction at a given trunk exchange should
cease. Advice is given to remove call restriction with
respect to the local exchange 116 once the ASR to the local
exchange 116 is greater than 50% and the percentage
30 occupancy (OCC) along the route from the parent trunk
exchange to the local exchange 116 is less than 80% for the
measuring period, and to remove the call gapping of the
far-end trunk exchanges 106 to 110 of the local exchange
116 once the ASR along the route from the far-end trunk
35 exchange to the parent exchange 102 of the local exchange
116 once again rises above 45%. (The percentage occupancy

is the total traffic in Erlangs, divided by the number of in-service circuits on a route).

These focused overload controls are applied to the trunk exchanges 102 to 110 unless the run-time system 206 recommends that focused overload controls no longer need to be applied notwithstanding that a local control may still be in force to ensure all controls are eventually removed. In the present embodiment this occurs once the BCH along the route from the home exchange 102 to the local exchange 116 falls to a value less than five.

It is envisaged that the run-time system 206 may directly control the network as shown in Figure 4 but at present it is expected that it will be necessary for a network traffic manager 208 to implement the method to allow overriding of the recommendations at his or her discretion.

The method described above can be simulated on a computer representation of BT's UK telecommunications network in which the run-time system 206 comprises a PROLOG-based expert system coded with the rules necessary to provide the recommendations described above in response to the appropriate CCITT parameters from the NTMS 204. This is also described in the above mentioned GB9210173.2 and US 115689/93. In embodiments of the present invention, however, an arrangement as described above with reference to Figures 2, 3 and 4 can be used to generate recommendations to apply differing call restriction levels to different routes of the network 202 so as to avoid, for instance, heavy call restriction on routes which make almost no contribution to a focused overload. That is, the arrangement described in general terms with reference to Figure 1 can be embodied by the arrangement described with reference to Figures 2, 3 or 4.

In an example of such a system, parameters supplied to the NTMS system 204 are processed by the run-time system 206 so as to provide recommendations for specific call

restriction levels to be applied at different points in the network 202. To do this, as described above, the change in bids on each route under consideration in the event of a focused overload is looked at in the context of the total change in bids, B_o , generated by the overload. Thus, as a starting point, the total change in bids, B_o , is again calculated, followed by the total bids to the customer, " B_i ", which it is desired should be offered after call restriction to ensure call restriction leaves very little or no "idle" time. Typically, the factor might be selected to be in the range from 4 to 10 in the British public switched telephone network. (This factor may vary according for instance to the type of network in which the call restriction is being applied, or according to prior agreement with a relevant customer.)

On each outgoing route, from each of the far end trunk exchanges or DMSUs 106, and the directly or indirectly connected local exchanges (none shown in Figure 2), the number of bids desired per measurement time period after restriction has been applied, " b_g ", is assessed which in turn gives a measure of the extent of call restriction which will be necessary. For instance, if call gapping is to be applied, the call gap rate, " g ", in seconds, on each of these outgoing routes, can be calculated as;

$$g = \frac{300}{b_g}$$

where a measurement time period is 5 minutes.

The effect of applying selective call gapping in the way described above as an embodiment of the present invention is that the call gap rate in seconds on a relevant route is inversely proportional to the excess traffic on that route due to the overload. It also ensures the total number of bids offered is high enough to avoid any loss of traffic actually dealt with. That is;

- a. the traffic allowed through is proportional to the traffic offered, so ensuring fairness, and

b. the total traffic offered to the customer is always in excess of the traffic they can handle. However, as a general principle, it is the application of call gapping beyond or excluding the routes between the parent exchanges 102, 104 and the local exchange 116 which is overloaded which allows call gapping to be applied in a selective manner.

If other call restriction methods such as proportional call blocking are used in place of call gapping, then "g" can instead be used as a measure of the proportion of calls to be blocked, rather than as a call gap rate in seconds.

It is not essential that call restriction be applied so as to allocate available bids in the proportions determined by the relative contributions of routes to an overload. It may be that call restriction is applied in a different manner. However, it is important that call restriction can be applied in a differential manner to routes beyond the parent exchanges of a local exchange subject to overload, or to routes directly connecting other local exchanges to said local exchange.

This Sheet Has Been Replaced





CLAIMS

1. A method of controlling a network, which includes at least one local exchange connected by one or more routes to one or more parent trunk exchanges, each parent trunk
5 exchange being connected by trunk routes to a plurality of further trunk exchanges, in which call restriction with respect to one or more numbers of a local exchange can be applied to one or more of the exchanges of the network, on
10 detection of a focused overload with respect to said local exchange, but is not applied to any parent trunk exchange of that local exchange.
2. A method as claimed in claim 1 in which said call restriction can be applied to a different extent at
15 different exchanges of the network, in response to said focused overload.
3. A method as claimed in claim 2 in which the extent to which call restriction is applied at each exchange is selected in accordance with the relative contributions made to the focused overload via the different exchanges of the
20 network.
4. A method as claimed in any preceding claim wherein said call restriction comprises call gapping.
5. A method as claimed in any preceding claim wherein said call restriction comprises proportional call blocking.
- 25 6. A method as claimed in any preceding claim wherein said call restriction comprises allowing calls up to a maximum number of calls to be made in predetermined time intervals.

This Sheet Has Been Replaced





7. A method as claimed in any one of claims 2 to 6 comprising the following steps on said detection of a focused overload;

- 5 i) an estimate is made of the total number of bids occurring in a measurement time period as a result of the focused overload, " B_0 ";
- 10 ii) for each of said different exchanges of the network the number of bids at that exchange occurring in the measurement time period as a result of the focused overload, " $b_c - b_s$ ", is estimated;
- 15 iii) the maximum number of bids desirably presented to the target customer premises equipment of the focused overload, B_1 , is apportioned to each of said different exchanges of the network, in
- direct relation to the ratio of " $b_c - b_s$ " to " B_0 " for each of said different exchanges; and
- 20 iv) the said call restriction is applied to each of the different exchanges of the network to an extent in accordance with said apportionment.

8. A method as claimed in claim 7, wherein a maximum call restriction level is applied whenever said apportionment results in less than a minimum number of bids being apportioned to an exchange in accordance with step iii).

25 9. Apparatus for controlling a network according to a method as set out in any one of the preceding claims, comprising a network traffic management system configured to detect occurrence of a focused overload to a target local exchange, wherein the network traffic management
30 system comprises trigger means such that detection of such occurrence triggers comparison of the number of bids on each route between a local or far end trunk exchange and the target local exchange measured in measurement time periods set immediately before and immediately after onset



of the focused overload, to calculate the total number of bids occurring as a consequence of the overload, and to apportion call restriction recommendations for said routes in accordance with the results of said comparisons.

- 5 10. Apparatus according to claim 9 wherein data concerning detection of the focused overload can be downloaded from the primary network traffic-management system and processed by means of a run-time system.



This Sheet Has Been Replaced

