



(51) International Patent Classification:
H04L 12/56 (2006.01) *H04L 29/00* (2006.01)
H04L 1/00 (2006.01)

(21) International Application Number:
PCT/AU2009/001546

(22) International Filing Date:
26 November 2009 (26.11.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2008906199 1 December 2008 (01.12.2008) AU

(71) Applicant (for all designated States except US): **COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION** [AU/AU]; Limestone Avenue, Campbell, ACT 2612 (AU).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **ROSBERG, Zvi** [IL/AU]; PO Box 76, Epping, NSW 1710 (AU).

(74) Agent: **SHELSTON IP**; 60 Margaret Street, Sydney, NSW 2000 (AU).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— of inventorship (Rule 4.17(iv))

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR NETWORK TRAFFIC CONTROL

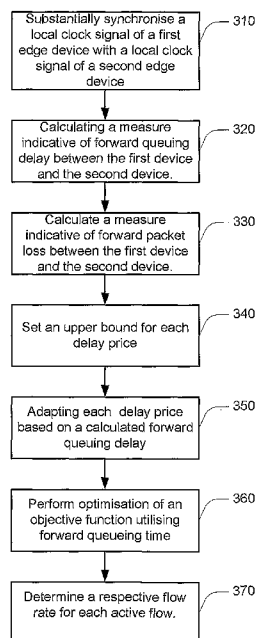


FIG. 3

(57) Abstract: A method and apparatus for controlling active network traffic flows between a first edge device and a second edge device, wherein said first device and said second device are coupled via a communication network, the method including the steps of: performing optimisation of an objective function utilising forward queuing time associated with each active flow; and determining a respective flow rate for each said active flow.



Published:

— *with international search report (Art. 21(3))*

- 1 -

METHOD AND APPARATUS FOR NETWORK TRAFFIC CONTROL**FIELD OF THE INVENTION**

[0001] The present invention relates to communication networks and in particular to communication network control.

- 5 [0002] The invention has been developed primarily for use as a method and apparatus for communication network control based on forward queuing time and forward packet loss probing and will be described hereinafter with reference to this application. However, it will be appreciated that the invention is not limited to this particular field of use.

BACKGROUND OF THE INVENTION

- 10 [0003] Any discussion of the prior art throughout the specification should in no way be considered as an admission that such prior art is widely known or forms part of the common general knowledge in the field.

- [0004] Rate control of flows in communication networks has been considered in the context of transfer control protocol (TCP). Previous analytical approaches have advocated *extended*
15 *proportional fair* rates whereby rate control is modelled as an optimization problem- subject to link capacity constraints using a fluid model. Two special cases of *extended proportional fairness* are *proportional fairness* and *max-min fairness*.

- [0005] Present control schemes for best-effort application flows typically aim at achieving fair rates subject to the link capacity constraints, with an aim for avoiding congestion. Moreover,
20 these control schemes are typically either based on round-trip time (RTT) feedback information or on some other feedback information produced by and delivered to the flow sources from one or more core routers along the forward path. RTT-based rate controls can result in unjustified low rates due to congestion on the backward path, even though the forward path is not congested at all. Alternative rate control based on standard feedback information from the core routers can
25 also result in substantially less than optimal network capacity utilization, due in part to a limited amount of information that can be extracted from standard feedback data produced in the prior art implementations. The controls that use non-standard feedback information from core routers further requires non-standard enhancement to be made to core routers along with non-standard signalling protocols. Such dependency on non-standard enhancements of core routers can make
30 these controls schemes less commercially attractive.

- [0006] United States Patent Application 11/608,834, by way of example, discloses a method and apparatus for communication network flow control. The approach taken in this disclosure involves a communication networks comprising user devices, edge routers, core routers, access and core links. An example is given for a method of computing and allocating fair transmission
35 rates to user data flows from a plurality of quality of service levels. These fair rates satisfy the minimum transmission rates, the end-to-end delays and the data loss rates required by each flow,

- 2 -

and can avoid network congestion. The method comprises an edge router process and a flow control policing unit for each edge router and a core router process for each edge and core router. All processes are executed in a distributed and asynchronous manner, are stable and converge to the desired fair rates.

- 5 [0007] Although best-effort flows are typically not associated with stringent delay and packet loss requirements, an RTT larger than 250 milliseconds or a packet loss rate larger than 1% to 3% are generally not acceptable in current network deployments. Current controls schemes have no means to guarantee such delay and packet loss requirements for best effort flows except by over-provisioning.
- 10 [0008] FIG. 1 shows a typical network infrastructure 100 associated with a wide area network (WAN). The wide area network 110 includes a plurality of core routers 120 for facilitating the transport and switching (routing) of packets across the network. An edge router 130 typically belongs to an end user (enterprise) organization and is located at the edge of a subscriber's network. The edge router provides an interface between the local or metropolitan network (for
- 15 example 140) and wide area network 110, such that devices 150 connected to a local or metropolitan network can transmit data onto - and received data from - the wide area network. [0009] The preferred embodiments can be implemented by edge network devices and does not rely on the reconfiguration of the core network infrastructure. This edge device is typically owned and controlled by enterprises or dedicated application developers that require a
- 20 substantially guaranteed end-to-end QoS across their systems.
- [0010] FIGS. 2A and 2B shows two example embodiments of an apparatus used in the implementation of a method for traffic control in a communications network.
- [0011] FIG. 2A shows a modified edge router 230 in which the processing element 210 is included in the edge router.
- 25 [0012] FIG. 2B shows a network device 235 coupled to an edge router 130. This network device 235 includes a processing element 210 and communication line adaptors 220 and 225, such that network packets can be received from - and transmitted to - the wide area network 110 to pass through the device and can be processed according to a method disclosed herein.
- [0013] It will be appreciated that processing of network traffic can occur in a standalone blade or rack-mounted device, or within network edge router.
- 30 [0014] By way of example, a method and apparatus for rate, delay and packet loss control of best effort application flows can be based on forward queuing time and forward packet loss probing and estimation to provide an edge-to-edge quality of service (QoS) solution for communication networks. This control method and apparatus is implemented at the network
- 35 edge in either an edge router or in a separate network device that is externally attaches to an edge

- 3 -

router. The control method utilizes only the information gathered in response to probe packets requesting provision for rate, delay and loss control of best effort flows.

[0015] Present control schemes for best-effort application flows are typically based on round-trip time (RTT) feedback information - or on some other feedback information - produced by
 5 and delivered to the flow sources from each core router along the forward path. As noted previously, RTT-based rate controls can result in unjustified low rates due to congestion on the backward path even though the forward path is not congested at all.

[0016] Alternative rate control methods that are based on feedback information typically depend on either the limited standard information provided by state of the art core routers or
 10 non-standard enhancements made to the core routers.

[0017] It will be appreciated that previous analytical alternatives provide two distributed control algorithms and have advocated *extended proportional fairness* rates whereby rate control is modelled as an optimization problem, subject to link capacity constraints, using a fluid model. Two special cases of *extended proportional fairness* are *proportional fairness* and *max-min*

15 *fairness*. (see A. Charny, "An algorithm for rate allocation in a packet-switching network with feedback" M.A. thesis, MIT, Cambridge, MA, 1994. J. Mo and J. Walrand, "Fair end-to-end window-based congestion control" IEEE/ACM TON, vol. 8, no. 5, pp. 556--567, Oct. 2000).

[0018] One previous rate-based algorithm which achieves proportional fair rates is based on a gradient search for an optimal solution of a primal problem (called the primal algorithm).

20 Another previous rate-based algorithm which achieves proportional fair rates is based on a gradient search for an optimal solution of a dual problem (called the dual algorithm). Yet another previous rate-based algorithm which achieves proportional fair rates combines the primal and the dual algorithms. (see D. X. Wei, C. Jin and S. H. Low, "FAST TCP: Motivation, architecture, algorithms, performance," IEEE/ACM TON, vol. 14, no. 6, pp. 1246--1259, Dec.

25 2006, and Z. Rosberg, "Control Plane for End-to-End QoS Guarantee: A Theory and Its Application," Proc. IWQoS'08, pp. 269--278, June 2008).

[0019] Other previous end-to-end window-based congestion controls (in contrast to the rate-based algorithms above) use packet round-trip-time (RTT) information to achieve either extended fair rates, or proportional fair rates or max-min fair rates.

30 [0020] Window-based controls using RTT information typically assume at any time t that, the rate of each n th flow ($x_n(t)$), its associated window size ($w_n(t)$), and its associated RTT ($RTT_n(t)$), are related by the following equation:

$$x_n(t) = \frac{w_n(t)}{RTT_n(t)}.$$

- 4 -

[0021] This relation can be described as a deterministic version of “Little’s Theorem” for ergodic queuing systems, and can be referred to as the *delay-window assumption*. It will be appreciated that this relation can be further relaxed.

5 [0022] A control method combining the explicit congestion notification (ECN) marking scheme with adaptive virtual queues has previously been developed, and stability of the primal and dual control algorithms under arbitrary time lags have been also studied. General sufficient conditions have been established for global stability in terms of the increase/decrease parameters of a congestion control algorithm and the price functions used at network links.

10 [0023] Although window-based controls use packet delay information, such as RTT, these delays are incorporated into the control method only through the *delay-window assumption*, which is an average law. More explicit incorporations of link delays have been previously proposed. For example, by representing each link delay as a function of its total load, and where the actual link delay trajectories are not left out of the framework. Alternatively, actual link delay trajectories have been included in the framework by using differential equations to specify
15 dynamic for which global asymptotic stable primal, dual and combined primal-dual controls were proposed. An earlier proposed combined rate and end-to-end delay control that aims at QoS guarantee flows which is based on RTT feedback information and requires new components to be included in core routers has also been proposed.

[0024] Typically, the control schemes for best-effort flows are trying to achieve fair rates
20 subject only to the link capacity constraints. It has been identified that these control schemes are typically based on round-trip time (RTT) feedback information rather than on forward time (FWT) feedback information. RTT-based rate controls are biased and may not achieve the desired effect since they also depend on the backward time (BWT). For example, the rate of flow from source to destination can be unjustifiably reduced due a congested backward path although
25 the forward path is not congested and has no packet loss. Although best-effort flows have no stringent delay and packet loss requirements, it is not acceptable to provide them a service with RTT larger than 250 milliseconds and packet loss rate larger than 1% to 3% in the core network.

OBJECT OF THE INVENTION

30 [0025] It is an object of the present invention to overcome or ameliorate at least one of the disadvantages of the prior art, or to provide a useful alternative.

[0026] It is an object of the invention in its preferred form to provide a method for network traffic control.

SUMMARY OF THE INVENTION

35 [0027] In accordance with a first aspect of the present invention, there is provided a method for network traffic control between a first edge device and a second edge device, wherein said first device and said second device are coupled via a communication network, said method

- 5 -

comprising the steps of: (a) performing optimisation of an objective function utilising the forward queuing time associated with each active flow; and (b) determining a respective flow rate for each said active flow.

[0028] Preferably, the optimisation applies a respective fixed delay price to each respective said
 5 active flow. Preferably, the optimisation maximizes $\sum_n U_n(x_n)/\alpha_n$ subject to link capacity constraints, wherein $U_n(x_n)$ is a rate utility function, and α_n is a positive delay price.

[0029] Preferably, the determining a respective flow rate for each said active flow includes the step of solving a discrete time formula of the form:

$$x_n(t + dt) = x_n(t) + dt \times (U'_n(x_n(t)) - \alpha_n(t) Q_n(t)),$$

10 [0030] wherein $x_n(t)$ is a flow rates at time t , $U_n(x_n)$ is a rate utility function, α_n is a positive delay price, $Q_n(t)$ is a forward queuing time, and dt is a discrete time tuning parameter defining the update frequency. The forward delay estimation can be used to control the rate of best effort application network traffic flow. Preferably, the method also includes the step of setting an upper bound $\bar{\alpha}_n$ for each said delay price. Preferably, the upper bound applied
 15 to each said delay price is inversely proportional to respective forward queuing delays. The method can also comprise the step of adapting each said delay price based on a calculated forward queuing delay and/or calculated forward packet loss. Preferably, each delay price is adapted based on a discrete time formula of the form:

$$\alpha_n(t + dt) = \begin{cases} \alpha_n(t) + \frac{dt}{\alpha_n(t)}, & \text{if excess delay or excess loss event occurs} \\ & \text{and } \alpha_n(t) < \bar{\alpha}_n \\ 0 & \text{otherwise} \end{cases}$$

20 [0031] wherein $\alpha_n(t)$ is a positive delay price at time t , $\bar{\alpha}_n$ is an upper bound for each delay price, and dt is discrete time tuning parameter for defining an update frequency.

[0032] Preferably, the method further comprises the step of substantially synchronising a local clock signal of said first device with a local clock signal of said second edge device. Preferably, the synchronising a local clock signal includes a step of sending to said second edge device a
 25 zero queuing packet comprising a measure indicative of said first device local time clock, such that receipt of said zero queuing packet at said second edge device is delayed substantially by a one forward propagation delay, and wherein said local clock synchronisation incorporates a one forward propagation delay offset.

[0033] Preferably, the method also comprises the step of calculating a measure indicative of
 30 forward queuing delay between said first device and said second device. Preferably, the method

- 6 -

also comprises the step of continuously calculating an arithmetic or damped average of at least two instances of said measure indicative of forward queuing delay. Preferably, the method also comprises the step of calculating a measure indicative of forward packet loss between said first device and said second device. Preferably, the method also comprises the step of calculating an arithmetic or damped average of at least two instances of said measure indicative of forward packet loss. Preferably, the method also provides rate, delay and packet loss control of best effort application flows based on forward queuing time probing and estimation and on forward packet loss probing and estimation, wherein said method provides an edge-to-edge quality of service solution across said communication network.

10 [0034] In accordance with a further aspect of the present invention, there is provided a method for controlling active network traffic flows between a first and second edge device coupled to a communication network, said method comprising a distributed control algorithms selected from the set including an algorithm based on a gradient search for an optimal solution of a primal problem; and an algorithm based on a gradient search for an optimal solution of a dual problem.

15 Preferably, the method uses forward time (FWT) feedback information and/or and forward packet loss such that said method is suitable for best-effort flow control.

[0035] In accordance with a further aspect of the present invention, there is provided a first edge device for controlling active network traffic flows between said first edge device and a second edge device, wherein said first device is coupled to said second device via a communication network, said first edge device comprising: at least one network interface for allowing said first device to communicate with said second edge device via said communication data network; a processor element coupled to said network interface for performing a method of for controlling active network traffic flows between said first edge device and said second edge device, wherein said method comprises the steps of: (a) performing optimisation of an objective function utilising forward queuing time associated with each active flow; determining a respective flow rate for each said active flow.

20

25

[0036] Preferably, the device is an edge router of said communication network or the device is coupled to an edge router of said communication network. The communication network can include a wired network and/or a wireless network.

30 **BRIEF DESCRIPTION OF THE DRAWINGS**

[0037] A preferred embodiment of the invention will now be described, by way of example only, with reference to the accompanying drawings in which:

FIG. 1 shows a typical network infrastructure for a wide area network (WAN);

FIG. 2A shows an example modified edge router in which a processing element is included for implementing a method according to the present invention;

35

- 7 -

FIG. 2B shows an example network device coupled to an edge router, where the network device includes a processing element and communication line adaptors for implementing a method according to the present invention; and

FIG. 3 shows an example flowchart for implementing a method according to the present invention.

PREFERRED EMBODIMENT OF THE INVENTION

[0038] Unlike earlier schemes, the method of the preferred embodiment uses forward time FWT (rather than RTT) and does not require additional feedback information from the routers. The preferred embodiment can be implemented in an apparatus placed only the edge network nodes. Furthermore, this control method can be based on FWT and forward packet loss and is suitable for best-effort flow control.

[0039] It will be appreciated that present invention can be utilised by network equipment vendors, enterprise networked systems, and applications developers. As such, a deployment scenario could include the enterprise market, applications market (such as video/audio content delivery, real-time game systems) and vendors.

[0040] Referring to Fig. 2, the preferred embodiments, by way of example only, can include a processing element 210 which further comprising one or more from the set including Network Processors (NP), Application-Specific Integrated Circuit (ASIC), and Field Programmable Gate Array (FPGA). This processing element is further coupled to communication line adaptors 220 and 225.

[0041] The preferred embodiment can be deployed amongst a wired and/or wireless broadband architecture, and can be used to substantially ensure an appropriate responsiveness of critical control traffic, voice and video traffic over wide area communication networks.

[0042] It will be appreciated that an embodiment of the present invention discloses a method to estimate forward delay (or forward time FWT) that is used to control the rates of best effort application network traffic flows. This method substantially removes bias associated with backward path congestion. This method does not require feedback from core routers, facilitating a solution that is suitable for deployment in existing communication network.

[0043] It will be appreciated that a RTT larger than 250 milliseconds and packet loss rate larger than 1% to 3% are generally not acceptable in current network deployments even for best effort flows. In an embodiment, flow rate control can substantially guarantee the meeting of such delay and packet loss requirement, without requiring any changes in the core routers or requiring over-provisioning. This can facilitate a network that is more economical and adaptable to traffic changes.

[0044] In an embodiment, as part of an edge-to-edge quality of service (QoS) solution for communication networks, a method and apparatus for rate, delay and packet loss control of best

- 8 -

effort flows based on forward queuing time and forward packet loss estimation is disclosed.

This control method and apparatus is implemented at the edge network device (such as an edge router or a standalone device), and is facilitated by forward path probing.

[0045] FIG. 3 shows, by way of example only, a flow chart of the steps for controlling active
 5 network traffic flows between a first edge device and a second edge device coupled to a communication network, wherein the first edge device performs the method comprising the steps of:

- (a) performing optimisation of an objective function 360 utilising forward queuing time associated with each active flow; and
- 10 (b) determining a respective flow rate 370 for each said active flow.

[0046] In another embodiment, there is provided a series of steps for controlling active network traffic flows between a first edge device and a second edge device coupled to a communication network, wherein the first edge device performs the steps of:

- 15 (a) substantially synchronise a local clock signal 310 of a first edge device with a local clock signal of a second edge device;
- (b) calculating a measure indicative of forward queuing delay 320 between said a first edge device and said a second edge device;
- (c) calculating a measure indicative of forward packet loss 330 between said a first edge device and said a second edge device;
- 20 (d) setting an upper bound for each said delay price 340;
- (e) adapting each said delay price 350 based on a calculated forward queuing delay and calculated ;
- (f) performing optimisation of an objective function 360 utilising forward queuing time associated with each active flow; and
- 25 (g) determining a respective flow rate 370 for each said active flow.

[0047] These example methods will be described further in relation to two embodiments associated with:

- rate control with fixed delay prices; and
- combined rate, delay and loss control.

30 **System Model**

[0048] A network (e.g. 100 of Fig. 1) can be described as comprising a set of N best effort flows and L links with link capacities (bandwidths) of $c = (c_1, \dots, c_L)^T$. There are possibly higher priority flows in the network with reserved bandwidth. The higher priority flows use at most the bandwidth reserved to them.

- 9 -

[0049] A system model for a network can be developed as follows: Let

$x(t) = (x_1(t), \dots, x_N(t))^T$ be the flow rates at time t ; $R(n)$ be the set of links comprising the route of flow n ; and $T(l)$ be the set of flows traversing through link l . For every time t , let $y_l(t)$ be the total rates of flows traversing link l at time t . The total rate of flows traversing link l at time t can be expressed by:

$$y_l(t) = \sum_{n \in T(l)} x_n(t)$$

[0050] Each flow is associated with a rate utility function $U_n(x_n)$, and a route delay penalty function $L_n(t)$. It can be considered that: $U_n = U_n(x_n)$ is a differentiable, strictly increasing and strictly concave function; and $L_n(t) = x_n Q_n(t)$, where $Q_n(t)$ is the forward (from source to destination) queuing time of flow n at time t . The forward queuing time $Q_n(t)$ can be expressed mathematically by the following equation.

$$Q_n(t) = \sum_{l \in R(n)} q_l(t) \quad (1)$$

[0051] In this equation, $q_l(t)$ is the time to clear the packet backlog of flows residing in the buffer of link l at time t . It will be appreciated that queuing delays and buffer occupancies are related by $b_l(t) = q_l(t) c_l$. Thus, delay and buffer occupancy controls can be considered equivalent.

[0052] Rate utilities and delay penalties can be combined into a single objective function. By defining positive delay prices $\alpha = (\alpha_1, \dots, \alpha_N)^T$ and specifying a rate control with adapted delay prices α , an objective function can be expressed mathematically as follows.

$$\mathbf{J}_\alpha(\mathbf{x}) = \sum_{n=1}^N [U_n(x_n) - \alpha_n Q_n x_n] \quad (2)$$

Rate Control with Fixed Delay Prices

[0053] Rate control can be specified with fixed delay prices by applying a continuous-time gradient projection algorithm for optimizing $\mathbf{J}_\alpha(\mathbf{x})$:

$$\dot{x}_n(t) = U'_n(x_n(t)) - \alpha_n(t) Q_n(t) \quad (3)$$

25 where

- 10 -

$$\alpha_n(t) = \alpha_n; \text{ and}$$

$$U'_n(x_n(t)) := \left. \frac{dU_n(x)}{dx} \right|_{x=x_n(t)}.$$

[0054] A discrete time version of equation (3) can be expressed mathematically as

$$x_n(t + dt) = x_n(t) + dt \times (U'_n(x_n(t)) - \alpha_n(t) Q_n(t)), \quad (4)$$

[0055] where dt is a discrete time tuning parameter defining the update frequency.

5 [0056] Implementing a rate control method at an edge router only, $Q(t)$ is estimated at the flow source edge by applying “*forward queuing time probing and estimation*” as described below.

[0057] An advantage of using $Q_n(t)$ over other control methods based on route penalties delivered by the core routers is that it does not require core router modification (or enhancement). It will be further appreciated that the present control method (or algorithm)

10 maximizes the following utility function subject to the link capacity constraints.

$$\sum_n \frac{U_n(x_n)}{\alpha_n}$$

Combined Rate, Delay and Loss Control

[0058] Best effort flows do not have stringent delay and loss requirement; however, they often prefer a maximum round trip delay (for example 250 milliseconds) and a packet loss less than
15 few percents but not at the expense of a significantly low data rate. To this end, the present method can set an upper bound $\bar{\alpha}_n$ on the delay price of each flow n .

[0059] An interpretation of $\bar{\alpha}_n$ can be expressed as one over the minimum number of bits transmitted per one forward queuing delay. For example, in a TCP flow environment, $1/\bar{\alpha}_n$ translates into the congestion window size (in bits).

20 [0060] It will be appreciated that forward queuing delays decrease with the delay prices, α . Hence, forward queuing delays can be made arbitrary small by a proper adaptation of α . Since packet loss is typically due to link buffer overflows, buffer sizes can also be made arbitrary small by the virtue of queuing delays and buffer occupancies being related by $b_l(t) = q_l(t) c_l$ (as discussed earlier).

25 [0061] The rate control expressed in equation (3) can be amended by adapting α_n . This amendment is based on two estimators, the forward queuing delay $Q_n(t)$, and the forward packet loss $L_n(t)$.

- 11 -

[0062] For every flow n , $\overline{Q}_n$ is defined to represent the associated upper bound on the end-to-end forward queuing time and $\overline{L}_n$ is defined to represent the associated upper bound on the end-to-end forward packet loss. $\alpha(t)$ is defined to represent a delay price vector used at time t .

[0063] An *excess delay event* can be expressed at an updated delay estimation time t by
 5 $\{Q_n(t) > \overline{Q}_n\}$. An *excess loss event* can be expressed at an updated loss estimation time t by
 $\{L_n(t) > \overline{L}_n\}$.

[0064] In an embodiment, the delay price vector used at time t , $\alpha(t)$, can be adapted as follows:

$$\dot{\alpha}_n(t) = \begin{cases} \frac{1}{\alpha_n(t)}, & \text{if excess delay or excess loss event occurs} \\ & \text{and } \alpha_n(t) < \overline{\alpha}_n \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

10 [0065] Equation (6) below shows a discrete time form of equation (5).

$$\alpha_n(t+dt) = \begin{cases} \alpha_n(t) + \frac{dt}{\alpha_n(t)}, & \text{if excess delay or excess loss event occurs} \\ & \text{and } \alpha_n(t) < \overline{\alpha}_n \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

[0066] In equation (6), dt is discrete time tuning parameter that defines an update frequency. It will be appreciated that equations (3) and (6) express a mathematical relationship applicable to a method of discrete rate, delay and loss control.

15 **Virtual Clock Synchronization and Forward Queuing Delay-Loss Probing**

[0067] An apparatus and method is taught for estimating the forward queuing time $Q(t)$ and packet loss $L(t)$ at time t .

[0068] Forward queuing delay estimation apparatus can be defined for networks having a zero queuing (ZQ) mechanism that forwards packets from a source to a destination without any
 20 queuing time. For example, a zero queuing (ZQ) mechanism can include the low latency queuing (LLQ) of Cisco IOS. Packets employing a ZQ mechanism are marked accordingly and receive top priority scheduling in each router along their traffic path. To guarantee zero queuing time, associated routers are typically configured to police (drop) all data packets marked as ZQ that exceed a predefined bandwidth allocated to ZQ packets. This queuing/scheduling scheme
 25 results in the total forward delay from source to destination of the un-policed ZQ packets equaling the forward propagation delay.

Virtual Clock Synchronization Between Edge Nodes

[0069] Substantial synchronization between edge nodes, defined by a source node s and a destination node d , can be achieved using a virtual clock at d . A virtual clock, v_s , of a source s can be maintained at each destination d of a source-destination pair (s, d) associated with an
 5 *active flow* comprising an exchange of data packets.

[0070] A source s can regularly send ZQ packets to the destination d , to provide data indicative of the source local clock, t_s . Upon receiving the source local clock update t_s , the destination d calculates (or updates) a time offset from s . The destination node d can calculate (or update) this time offset as $OFS(d, s) = t_d - t_s$, where t_d is representative of the
 10 clock at the destination node upon reception.

[0071] The virtual clock of s can be specified at a time instant t_d of node d , using the following equation:

$$v_s = t_d - OFS(d, s)$$

[0072] It will be appreciated that a ZQ packet, providing data indicative of the source local
 15 clock t_s , arrives at a destination node d after a one-way propagation delay. Thus, the virtual clock of s at d is lagging by a one-way propagation delay. Regularly updating $OFS(d, s)$ can assist in minimizing an offset due to individual clock drifts.

Forward Queuing Time Probing and Estimation

[0073] Forward queuing time probing and estimation can be established between edge nodes,
 20 defined by a source node s and a destination node d . Each source s of a source-destination pair, (s, d) , with an active flow can maintain a forward queuing delay estimator, Q .

[0074] The source s can sample values of Q by regularly sending probe packets to the destination d . Probe packets sent to the destination can be used for multiple purposes, one of which is delay estimation. The probe packets are sent as best effort packets, i.e. they are not
 25 marked as ZQ packets, and therefore may experience queuing time. Each probe can be labeled with a sequential number, $n = 1, 2, \dots$. Before transmission to a destination node d , each probe packet is time stamped by the source s with its local clock t_s . Upon receiving a probe packet at destination d , the destination time stamps the probe with its virtual clock v_s of the relevant source node s . The destination then can respond by sending the corresponding probe
 30 packet back to source s .

- 13 -

[0075] When a source node receives a packet corresponding to a previously sent probe packet n , the source records q_n as the n th sample of Q . This can be expressed mathematically by the following equation, where v_s and t_s are time stamps carried in the fields of the n th probe packet:

$$q_n = v_s - t_s$$

[0076] It will be appreciated that q_n is the forward queuing time of probe n . The value of Q can be initialized to a very large number and updated as a probe packet arrives. For example as the source node receives a returned probe packet n , Q can be updated by one of plurality of estimation procedures based on the sample $\{q_i; 1 \leq i \leq n\}$. Two example embodiments of estimation procedures are disclosed.

[0077] A first example embodiment for estimating forward queuing time, involving calculation of the arithmetic average of the last k samples, can be expressed mathematically in the following equation.

$$Q = \frac{1}{k} \sum_{i=n-k+1}^n q_i$$

[0078] By keeping Q and q_{n-k+1} in a local data structure, upon every arrival of q_n , the average forward queuing time can be implemented on-the-fly using the following equation.

$$Q \leftarrow Q + \frac{q_n - q_{n-k+1}}{k}$$

[0079] A second example embodiment for estimating forward queuing time, involving calculation of a damped average with factor $0 < \beta < 1$, can be expressed mathematically in the following equation.

$$Q \leftarrow (1 - \beta) Q + \beta q_n$$

Forward packet loss probing and estimation

[0080] Each source node s of a source-destination pair, (s, d) , with an active traffic flow, can maintain a forward packet loss estimator, L_s .

[0081] A source node s can sample values of L by regularly sending probe packets to the destination d . Probe packets sent to the destination can be used for multiple purposes, one of which is packet loss estimation. Each probe packet can be labeled with a sequential number, $n = 1, 2, \dots$ by containing a representation of the sequence number in a data field of the packet.

- 14 -

[0082] The destination d also maintains a current forward packet loss estimator on behalf of the source s , $L1_s$. The last sequence number received from a respective source node s is denoted by n_{last} . Initially, $L1_s$ and L_s can be set to one. When the destination node d receives probe packet representing with sequence number n , it records the next sample (for example the i th sample), of the number of losses so far denoted by l_i . This can be represented mathematically by the following equation.

$$l_i = n - n_{\text{last}} - 1 \quad (7)$$

[0083] The destination node d can then also update n_{last} as expressed below.

$$n_{\text{last}} \leftarrow n$$

[0084] $L1_s$ can be estimated by one of a plurality of estimation procedures based on the samples $\{l_j; 1 \leq j \leq i\}$. Two embodiments for estimating forward packet loss are disclosed by way of example only.

[0085] A first example embodiment for estimating forward packet loss, involving calculation of the arithmetic average of the last k samples, can be expressed mathematically in the following equation.

$$L1 = \frac{\sum_{j=i-k+1}^i l_j}{k + \sum_{j=i-k+1}^i l_j}$$

[0086] By retaining the values of $\sum_{j=i-k+1}^i l_j$ and l_{i-k+1} in a local data structure, the average can be calculated on-the-fly upon the receipt of each probe packet and evaluation of the associated l_i specified in (2). This on-the-fly calculation of a forward packet loss estimation can be implemented by setting

$$L1 = \frac{S - l_{i-k+1} + l_i}{k + S - l_{i-k+1} + l_i}$$

[0087] A second example embodiment for estimating forward packet loss, involving calculation of the damped average with factor $0 < \beta < 1$, can be expressed mathematically in the following equation.

- 15 -

$$L1 \leftarrow (1 - \beta) L1 + \beta \frac{l_i}{1 + l_i}$$

[0088] After every update of $L1$ at destination node d , the destination writes $L1$ in the probe packet and sends the probe back to source s . Upon probe reception at node s , the source updates L by setting $L1$ into L , i.e.,

5
$$L \leftarrow L1$$

Interpretation

[0089] The invention may be embodied using devices conforming to other network standards and for other applications, including, for example other WLAN standards and other wireless standards. Applications that can be accommodated include IEEE 802.11 wireless LANs and
10 links, wired and wireless Ethernet.

[0090] In the context of this document, the term “wireless” and its derivatives may be used to describe circuits, devices, systems, methods, techniques, communications channels, etc., that may communicate data through the use of modulated electromagnetic radiation through a non-solid medium. The term does not imply that the associated devices do not contain any wires,
15 although in some embodiments they might not. In the context of this document, the term “wired” and its derivatives may be used to describe circuits, devices, systems, methods, techniques, communications channels, etc., that may communicate data through the use of modulated electromagnetic radiation through a solid medium. The term does not imply that the associated devices are coupled by electrically conductive wires.

20 [0091] Unless specifically stated otherwise, as apparent from the following discussions, it is appreciated that throughout the specification discussions utilizing terms such as “processing”, “computing”, “calculating”, “determining” or the like, refer to the action and/or processes of a computer or computing system, or similar electronic computing device, that manipulate and/or transform data represented as physical, such as electronic, quantities into other data similarly
25 represented as physical quantities.

[0092] In a similar manner, the term “processor” may refer to any device or portion of a device that processes electronic data, e.g., from registers and/or memory to transform that electronic data into other electronic data that, e.g., may be stored in registers and/or memory. A “computer” or a “computing device” or a “computing machine” or a “computing platform” may
30 include one or more processors.

[0093] The methodologies described herein are, in one embodiment, performable by one or more processors that accept computer-readable (also called machine-readable) code containing a set of instructions that when executed by one or more of the processors carry out at least one of the methods described herein. Any processor capable of executing a set of instructions

- 16 -

(sequential or otherwise) that specify actions to be taken are included. Thus, one example is a typical processing system that includes one or more processors. The processing system further may include a memory subsystem including main RAM and/or a static RAM, and/or ROM.

[0094] Furthermore, a computer-readable carrier medium may form, or be included in a computer program product.

[0095] In alternative embodiments, the one or more processors operate as a standalone device or may be connected, e.g., networked to other processor(s), in a networked deployment, the one or more processors may operate in the capacity of a server or a client machine in server-client network environment, or as a peer machine in a peer-to-peer or distributed network environment.

10 The one or more processors may form a web appliance, a network router, switch or bridge, or any machine capable of executing a set of instructions (sequential or otherwise) that specify actions to be taken by that machine.

[0096] Note that while some diagram(s) only show(s) a single processor and a single memory that carries the computer-readable code, those in the art will understand that many of the components described above are included, but not explicitly shown or described in order not to obscure the inventive aspect. For example, while only a single machine is illustrated, the term "machine" shall also be taken to include any collection of machines that individually or jointly execute a set (or multiple sets) of instructions to perform any one or more of the methodologies discussed herein.

20 [0097] Thus, one embodiment of each of the methods described herein is in the form of a computer-readable carrier medium carrying a set of instructions, e.g., a computer program that are for execution on one or more processors. Thus, as will be appreciated by those skilled in the art, embodiments of the present invention may be embodied as a method, an apparatus such as a special purpose apparatus, an apparatus such as a data processing system, or a computer-readable carrier medium. The computer-readable carrier medium carries computer readable code including a set of instructions that when executed on one or more processors cause a processor or processors to implement a method. Accordingly, aspects of the present invention may take the form of a method, an entirely hardware embodiment, an entirely software embodiment or an embodiment combining software and hardware aspects. Furthermore, the present invention may take the form of carrier medium (e.g., a computer program product on a computer-readable storage medium) carrying computer-readable program code embodied in the medium.

30 [0098] The software may further be transmitted or received over a network via a network interface device. While the carrier medium is shown in an example embodiment to be a single medium, the term "carrier medium" should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, and/or associated caches and servers) that store the one or more sets of instructions. The term "carrier medium" shall also be taken to include

- 17 -

any medium that is capable of storing, encoding or carrying a set of instructions for execution by one or more of the processors and that cause the one or more processors to perform any one or more of the methodologies of the present invention. A carrier medium may take many forms, including but not limited to, non-volatile media, volatile media, and transmission media.

5 [0099] It will be understood that the steps of methods discussed are performed in one embodiment by an appropriate processor (or processors) of a processing (i.e., computer) system executing instructions (computer-readable code) stored in storage. It will also be understood that the invention is not limited to any particular implementation or programming technique and that the invention may be implemented using any appropriate techniques for implementing the
10 functionality described herein. The invention is not limited to any particular programming language or operating system.

[00100] Reference throughout this specification to “one embodiment” or “an embodiment” means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, appearances
15 of the phrases “in one embodiment” or “in an embodiment” in various places throughout this specification are not necessarily all referring to the same embodiment, but may. Furthermore, the particular features, structures or characteristics may be combined in any suitable manner, as would be apparent to one of ordinary skill in the art from this disclosure, in one or more embodiments.

20 [00101] Similarly it should be appreciated that in the above description of example embodiments of the invention, various features of the invention are sometimes grouped together in a single embodiment, figure, or description thereof for the purpose of streamlining the disclosure and aiding in the understanding of one or more of the various inventive aspects. This method of disclosure, however, is not to be interpreted as reflecting an intention that the claimed invention
25 requires more features than are expressly recited in each claim. Rather, as the following claims reflect, inventive aspects lie in less than all features of a single foregoing disclosed embodiment. Thus, the claims following the Detailed Description are hereby expressly incorporated into this Detailed Description, with each claim standing on its own as a separate embodiment of this invention.

30 [00102] Furthermore, while some embodiments described herein include some but not other features included in other embodiments, combinations of features of different embodiments are meant to be within the scope of the invention, and form different embodiments, as would be understood by those in the art. For example, in the following claims, any of the claimed embodiments can be used in any combination.

35 [00103] Furthermore, some of the embodiments are described herein as a method or combination of elements of a method that can be implemented by a processor of a computer system or by

other means of carrying out the function. Thus, a processor with the necessary instructions for carrying out such a method or element of a method forms a means for carrying out the method or element of a method. Furthermore, an element described herein of an apparatus embodiment is an example of a means for carrying out the function performed by the element for the purpose of carrying out the invention.

[00104] In the description provided herein, numerous specific details are set forth. However, it is understood that embodiments of the invention may be practiced without these specific details. In other instances, well-known methods, structures and techniques have not been shown in detail in order not to obscure an understanding of this description.

[00105] As used herein, unless otherwise specified the use of the ordinal adjectives “first”, “second”, “third”, etc., to describe a common object, merely indicate that different instances of like objects are being referred to, and are not intended to imply that the objects so described must be in a given sequence, either temporally, spatially, in ranking, or in any other manner.

[00106] In the claims below and the description herein, any one of the terms comprises, comprising, comprised of or which comprises is an open term that means including at least the elements/features that follow, but not excluding others. Thus, the term comprising, when used in the claims, should not be interpreted as being limitative to the means or elements or steps listed thereafter. For example, the scope of the expression a device comprising A and B should not be limited to devices consisting only of elements A and B. Any one of the terms including or which includes or that includes as used herein is also an open term that also means including at least the elements/features that follow the term, but not excluding others. Thus, including is synonymous with and means comprising.

[00107] Similarly, it is to be noticed that the term coupled, when used in the claims, should not be interpreted as being limitative to direct connections only. The terms “coupled” and “connected”, along with their derivatives, may be used. It should be understood that these terms are not intended as synonyms for each other. Thus, the scope of the expression a device A coupled to a device B should not be limited to devices or systems wherein an output of device A is directly connected to an input of device B. It means that there exists a path between an output of A and an input of B which may be a path including other devices or means. “Coupled” may mean that two or more elements are either in direct physical or electrical contact, or that two or more elements are not in direct contact with each other but yet still co-operate or interact with each other.

[00108] Thus, while there has been described what are believed to be the preferred embodiments of the invention, those skilled in the art will recognize that other and further modifications may be made thereto without departing from the spirit of the invention, and it is intended to claim all such changes and modifications as fall within the scope of the invention. For example, any

- 19 -

formulas given above are merely representative of procedures that may be used. Functionality may be added or deleted from the block diagrams and operations may be interchanged among functional blocks. Steps may be added or deleted to methods described within the scope of the present invention.

- 5 [00109] Although the invention has been described with reference to specific examples, it will be appreciated by those skilled in the art that the invention may be embodied in many other forms.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A method for controlling active network traffic flows between a first edge device and a second edge device, wherein said first device and said second device are coupled via a communication network, said method comprising the steps of:
 - 5 (a) performing optimisation of an objective function utilising forward queuing time associated with each active flow; and
 - (b) determining a respective flow rate for each said active flow.
2. A method according to claim 1, wherein said optimisation applies a respective fixed delay price to each respective said active flow.
- 10 3. A method according to any one of the preceding claims, wherein said optimisation maximizes $\sum_n U_n(x_n)/\alpha_n$ subject to link capacity constraints, wherein $U_n(x_n)$ is a rate utility function, and α_n is a positive delay price.
4. A method according to any one of the preceding claims, wherein said determining a respective flow rate for each said active flow includes the step of solving a discrete time

15 formula of the form:

$$x_n(t + dt) = x_n(t) + dt \times (U'_n(x_n(t)) - \alpha_n(t) Q_n(t)),$$
 wherein $x_n(t)$ is a flow rates at time t , $U_n(x_n)$ is a rate utility function, α_n is a positive delay price, $Q_n(t)$ is a forward queuing time, and dt is a discrete time tuning parameter defining the update frequency.
- 20 5. A method according to claim 4, wherein forward delay estimation is used to control the rate of best effort application network traffic flow.
6. A method according to claims 4 or claim 5, wherein said method comprises the step of setting an upper bound $\bar{\alpha}_n$ for each said delay price.
7. A method according to claim 6, wherein said upper bound applied to each said delay price

25 is inversely proportional to respective forward queuing delays.
8. A method according to any one of claims 4 to 7, wherein method further comprises the step of adapting each said delay price based on a calculated forward queuing delay and/or calculated forward packet loss.
9. A method according to claim 8, wherein each said delay price is adapted based on a discrete

30 time formula of the form:

$$\alpha_n(t + dt) = \begin{cases} \alpha_n(t) + \frac{dt}{\alpha_n(t)}, & \text{if excess delay or excess loss event occurs} \\ & \text{and } \alpha_n(t) < \bar{\alpha}_n \\ 0 & \text{otherwise} \end{cases}$$

- 21 -

wherein $\alpha_n(t)$ is a positive delay price at time t , $\overline{\alpha}_n$ is an upper bound for each delay price, and dt is discrete time tuning parameter for defining an update frequency.

10. A method according to any one of the preceding claims wherein said method further comprises the step of substantially synchronising a local clock signal of said first device
5 with a local clock signal of said second edge device.
11. A method according to claim 10, wherein said synchronising a local clock signal includes a step of sending to said second edge device a zero queuing packet comprising a measure indicative of said first device local time clock, such that receipt of said zero queuing packet at said second edge device is delayed substantially by a one forward propagation delay, and
10 wherein said local clock synchronisation incorporates a one forward propagation delay offset.
12. A method according to any one of the preceding claims wherein said method comprises the step of calculating a measure indicative of forward queuing delay between said first device and said second device.
- 15 13. A method according to claim 12, wherein said method comprises the step of continuously calculating an arithmetic or damped average of at least two instances of said measure indicative of forward queuing delay.
14. A method according to any one of the preceding claims wherein said method comprises the step of calculating a measure indicative of forward packet loss between said first device and
20 said second device.
15. A method according to claim 14, wherein said method comprises the step of calculating an arithmetic or damped average of at least two instances of said measure indicative of forward packet loss.
16. A method according to any one of the preceding claims, wherein said method provides rate,
25 delay and packet loss control of best effort application flows based on forward queuing time probing and estimation and on forward packet loss probing and estimation, wherein said method provides an edge-to-edge quality of service solution across said communication network.
17. A method for controlling active network traffic flows between a first and second edge
30 device coupled to a communication network, said method comprising a distributed control algorithms selected from the set including an algorithm based on a gradient search for an optimal solution of a primal problem; and an algorithm based on a gradient search for an optimal solution of a dual problem.
18. A method according to claim 17, wherein said method uses forward time (FWT) feedback
35 information and/or and forward packet loss such that said method is suitable for best-effort flow control.

- 22 -

19. A method for controlling active network traffic flows between a first and second edge device coupled to a communication network, substantially as herein described with reference to any one of the embodiments of the invention illustrated in the accompanying drawings and/or examples.
- 5 20. A device for controlling active network traffic flows, said device comprising a processing element coupled to at least one network interface, said processor element being adapted for performing a method according to any one of the preceding claims.
21. A first edge device for controlling active network traffic flows between said first edge device and a second edge device, wherein said first device is coupled to said second device
10 via a communication network, said first edge device comprising:
at least one network interface for allowing said first device to communicate with said second edge device via said communication data network;
a processor element coupled to said network interface for performing a method of for controlling active network traffic flows between said first edge device and said
15 second edge device, wherein said method comprises the steps of:
(a) performing optimisation of an objective function utilising forward queuing time associated with each active flow;
(b) determining a respective flow rate for each said active flow.
22. A device according to claim 20 or claim 21, wherein said device is an edge router of said
20 communication network.
23. A device according to claim 20 or claim 21, wherein said device is coupled to an edge router of said communication network.
24. A device according to any one of claims 20 to 23 wherein said communication network includes a wired network and/or a wireless network.
- 25 25. A device according to any one of claims 20 to 24, wherein said processing element comprises one or more from the set including Network Processors (NP), Application-Specific Integrated Circuit (ASIC), Field Programmable Gate Array (FPGA).
26. A device for controlling active network traffic flows, substantially as herein described with reference to any one of the embodiments of the invention illustrated in the accompanying
30 drawings and/or examples.
27. A computer-readable carrier medium carrying a set of instructions that when executed by one or more processor elements cause the one or more processor elements to carry out a method of according to any one of claims 1 to 19.

1/3

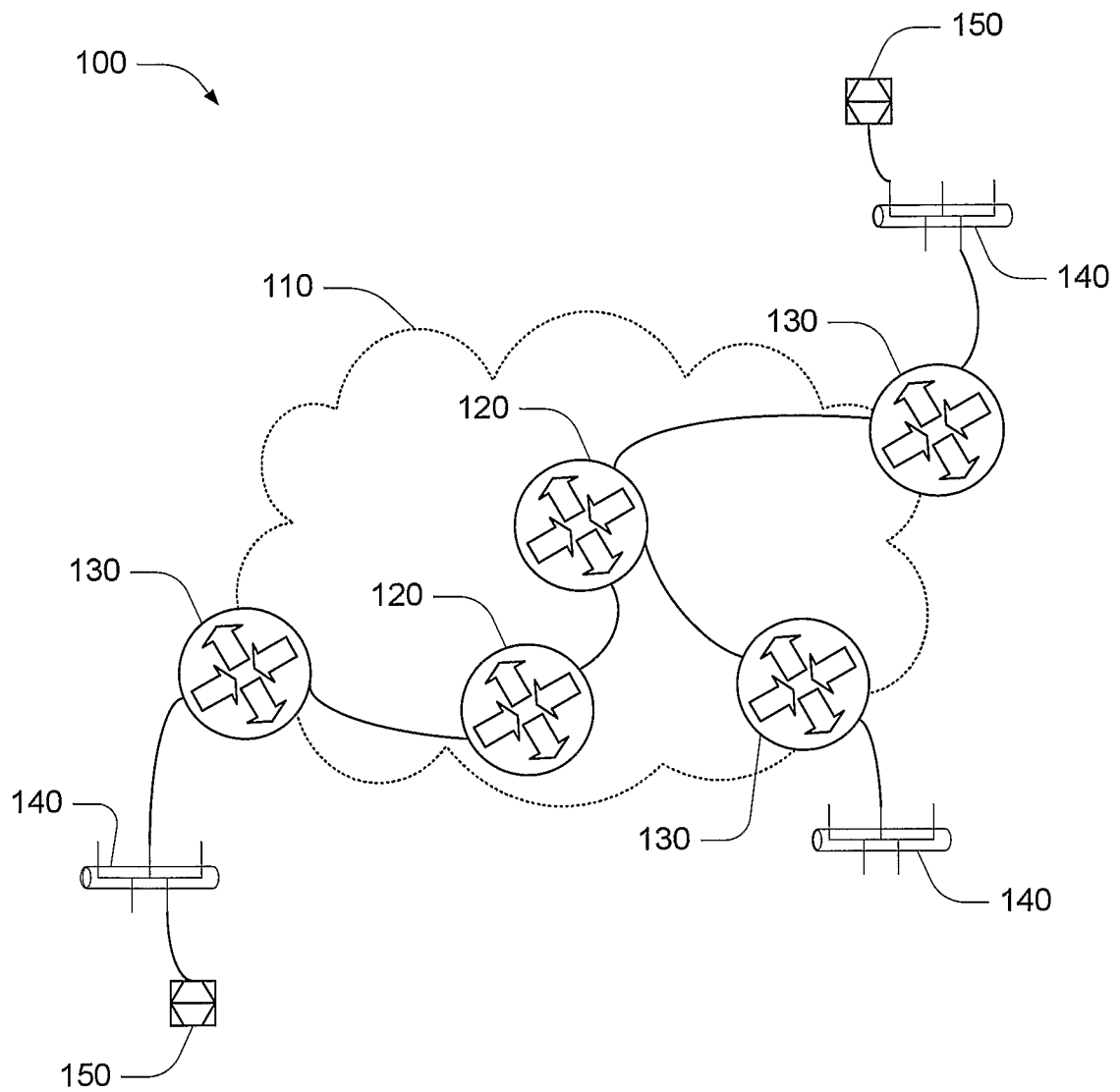


FIG. 1

2/3

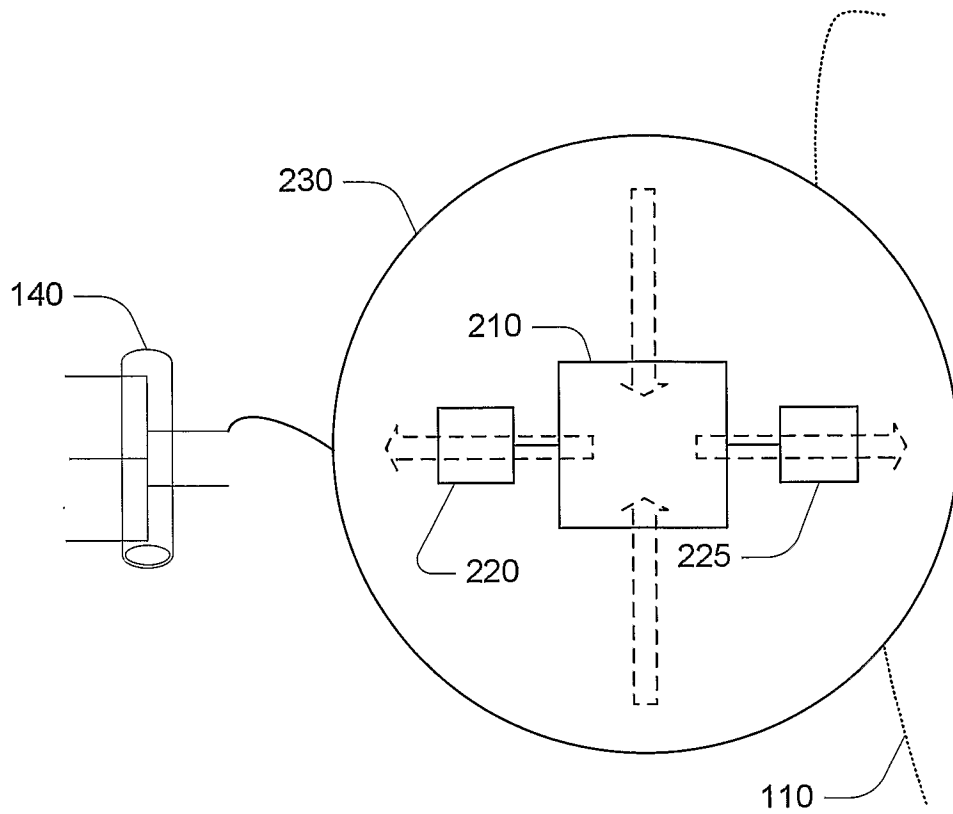


FIG. 2A

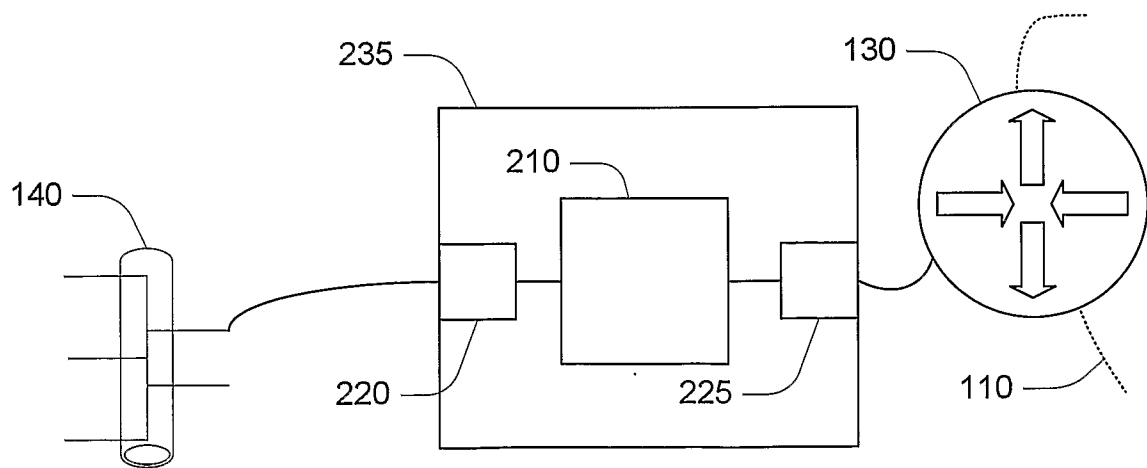
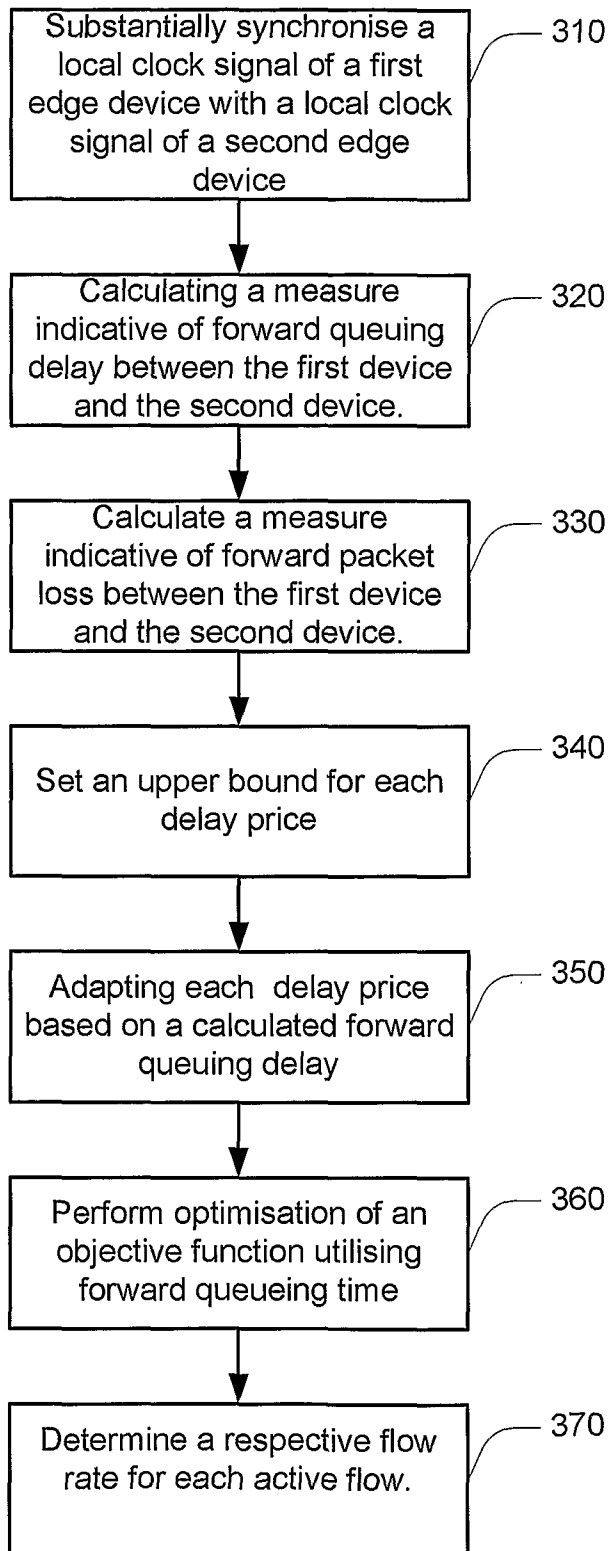


FIG. 2B

3/3**FIG. 3**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU2009/001546

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

H04L 12/56 (2006.01), H04L 1/00 (2006.01), H04L 29/00 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPOQUE (inspec)- (communication, network, flow, control, queuing, time, delay) and similar terms;
EPOQUE (epodoc and wpi)- 'optimisation' and terms same as above; Google Scholar, IEEE Explore, Google Patents, esp@cenet- 'TCP', 'rate' and similar terms as above.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PAPACHRISTODOULOU, A., "Global stability analysis of a TCP/AQM protocol for arbitrary networks with delay", Decision and Control, 2004, CDC, 43 rd IEEE Conference, Volume 1, 17 December 2004, Pages 1029-1034 (Title, abstract, figure 1, sections 'Introduction', 'PROBLEM FORMULATION')	1, 10, 12, 20-25, 27
A	WEI, D X. et al., "FAST TCP: Motivation, Architecture, Algorithms, Performance", IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 14, NO. 6, DECEMBER 2006, Pages 1246-1259 (Abstract)	
A	WO 1998/047166 A2 (FLASH NETWORKS LTD. [IL/IL]) 22 October 1998 (Abstract)	

☐ Further documents are listed in the continuation of Box C☒ See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"E" earlier application or patent but published on or after the international filing date

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"O" document referring to an oral disclosure, use, exhibition or other means

"&" document member of the same patent family

"P" document published prior to the international filing date but later than the priority date claimed

Date of the actual completion of the international search
08 February 2010

Date of mailing of the international search report

17 FEB 2010

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustralia.gov.au
Facsimile No. +61 2 6283 7999

Authorized officer

Pradnya Satarkar

AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No. : +61 2 6283 2751

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2009/001546

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: **19 and 26**
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
These claims do not comply with Rule 6.2(a) because they rely on references to the description and/or drawings.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

- Invention 1: Claims 1-16, 20 (*when appended to claim 1*), 21-25 and 27 (*when appended to claim 1*).
- Invention 2: Claims 17-18, 20 (*when appended to claim 17*) and 27 (*when appended to claim 17*).

Note: See 'Supplemental Box 1' for details.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: **1-16, 20-25 and 27**.

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

Supplemental Box 1

(To be used when the space in any of Boxes I to IV is not sufficient)

Continuation of Box No: III

This International Application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept.

In assessing whether there is more than one invention claimed, I have given consideration to those features which can be considered to potentially distinguish the claimed combination of features from the prior art. Where different claims have different distinguishing features they define different inventions.

This International Searching Authority has found that there are **2** different inventions as follows:

- Claims 1-16, 20 (*when appended to claim 1*), 21-25, 27 (*when appended to claim 1*) are directed to a method and device for controlling active network traffic flows between a first edge device and a second edge device, wherein both of these are coupled via a communication network comprising steps of: performing optimisation of an objective function utilising forward queuing time associated with each active flow and determining a respective flow rate for each said active flow. It is considered that '*performing optimisation of an objective function utilising forward queuing time associated with each active flow and determining a respective flow rate for each said active flow*' comprises a **first** distinguishing feature.
- Claims 17-18, 20 (*when appended to claim 17*), 27 (*when appended to claim 17*) are directed to a method for controlling active network traffic flows between a first edge device and a second edge device, wherein both of these are coupled via a communication network, the method comprising steps of: using a distributed control algorithms selected from the set including an algorithm based on a gradient search for an optimal solution of a primal problem and an algorithm based on a gradient search for an optimal solution of a dual problem. It is considered that '*using a distributed control algorithms selected from the set including an algorithm based on a gradient search for an optimal solution of a primal problem and an algorithm based on a gradient search for an optimal solution of a dual problem*' comprises a **second** distinguishing feature.

PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art.

Each of the abovementioned groups of claims has a different distinguishing feature and they do not share any feature which could satisfy the requirement for being a special technical feature. Because there is no common special technical feature it follows that there is no technical relationship between the identified inventions. Therefore the claims do not satisfy the requirement of unity of invention *a priori*.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/AU2009/001546

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report				Patent Family Member	
WO	1998/047166	AU	69332/98	US	6273622
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.					
END OF ANNEX					