

COPYRIGHT NOTICE

[0001] Contained herein is material that is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the United States Patent and Trademark Office patent file or records, but otherwise reserves all rights to the copyright whatsoever. The following notice applies to the software and data as described below and in the drawings hereto: Copyright © 2002, Intel Corporation, All Rights Reserved.

FIELD

[0002] An embodiment of the invention relates to data networks in general, and more specifically to guaranteed service in a data network.

BACKGROUND

[0003] In a data network, quality of service (QoS) refers to the service provided for network traffic. A particular type of quality of service is “guaranteed service” in which the end-to-end delay for data transmission is guaranteed to be no more than a certain delay bound or limit. Guaranteed service is particularly important in services involving mission critical communications, including voice over Internet technology (voice over IP or VoIP), in which an Internet connection is used to transport voice data, and virtual private networking (VPN), in which the Internet or another public communication medium is used for transport of private network data.

[0004] In order to guarantee an end-to-end delay bound for data traffic on a data path, each node along the data path must guarantee a delay bound and the sum of the per node delay bounds is limited to the end-to-end delay bound. One data traffic capacity

system that may be instituted in a node of a network is a token bucket. However, in order to utilize a traffic capacity system such as a token bucket model in a network and ensure that a delay bound is met, minimum parameters need to be established for the traffic capacity system. The required minimum parameters that will enable the system to remain
5 within the delay upper bound are dependent upon the data traffic pattern that is encountered. In conventional network traffic systems, there is no efficient and effective mechanism for determining the minimum parameters that are required in order to meet guaranteed service requirements for a particular data traffic pattern.

BRIEF DESCRIPTION OF THE DRAWINGS

10 [0005] The invention may be best understood by referring to the following description and accompanying drawings that are used to illustrate embodiments of the invention. In the drawings:

[0006] Figure 1 is an illustration of token bucket operation under an embodiment of the invention;

15 [0007] Figure 2 is an illustration of traffic specification technology under an embodiment of the invention;

[0008] Figure 3 is an illustration of an exemplary network architecture in which an embodiment of the invention may be utilized; and

[0009] Figure 4 is a flow chart illustrating the determination of an optimal token
20 bucket size under an embodiment of the invention.

DETAILED DESCRIPTION

[0010] A method and apparatus are described for guaranteed service in a data network.

[0011] In the following description, for the purposes of explanation, numerous
25 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 circuits, structures, techniques, and devices have not been shown in detail.

[0012] Embodiments of the invention include various processes, which will be described below. The processes may be performed by hardware components or may be embodied in machine-executable instructions, which may be used to cause a general-purpose or special-purpose processor or logic circuits programmed with the instructions to perform the processes. Alternatively, the processes may be performed by a combination of hardware and software.

Terminology

[0013] Before describing an exemplary environment in which various embodiments of the present invention may be implemented, some terms that will be used throughout this application will briefly be defined:

[0014] As used herein, "quality of service" or "QoS" means the quality of service provided for network data traffic.

[0015] As used herein, "guaranteed service" means a quality of service that guarantees that the end-to-end delay for a data flow does not exceed an upper bound.

[0016] According to an embodiment of the invention, a data traffic pattern is transmitted in a network, with the network traffic being comprised of data packets. According to an embodiment of the invention, the network includes a traffic control system to control the data traffic flow. According to an embodiment of the invention, the network utilizes a token bucket controller model in a traffic control system. According to an embodiment of the invention, an upper bound is prescribed for delay of the data traffic and, utilizing information regarding a traffic pattern, a parameter for the traffic control system is determined such that no delay of a data packet exceeds the upper bound. Under

a particular embodiment, the determined parameter is the size of a token bucket.

According to an embodiment of the invention, the determined parameter is the minimum parameter that ensures operations complying with the delay upper bound.

[0017] Under an embodiment of the invention, a traffic pattern collector collects
5 raw traffic patterns from a network data flow and passes information regarding the traffic patterns to a measurement-based traffic specification (MBTS) engine. According to the embodiment, the engine converts the raw traffic pattern data into flow specifications and provides the flow specifications to a QoS service provider. The flow specifications are used in the specification of data traffic control.

10 [0018] Under an embodiment of the invention, the processing of traffic patterns is accomplished on a real-time basis. The utilization of real-time processing for data traffic control makes the efficiency of the processing algorithm extremely important in maintaining efficient network operation.

[0019] Under an embodiment of the invention, a token bucket controller is used
15 for purposes of traffic control. Token bucket data controllers, which may be utilized in traffic shapers, switches, routers, and other networking devices, allow for limited delay for data packets. According to an embodiment, delayed data packets are temporarily cached in a queue associated with the token bucket controller. For a given traffic pattern and a prescribed delay upper bound, token bucket parameters are established to ensure that no
20 delay of a data packet exceeds the upper bound. In particular, a token bucket size parameter is determined that is the minimum size that can guarantee that the delay upper bound is not violated.

[0020] Figure 1 illustrates a token bucket model that may be utilized in
connection with an embodiment of the invention. In this illustration, a transmitting device
25 105 is transmitting data to a receiving device 115 over a communication path 110. The

transmitting device 105 and receiving device 115 may be any device or system in a computer network, including any network server, switch, router, client system, or other device or system. The data is transmitted in the form of data packets, with the data packets being of varying sizes and arriving at various times. The data packets shown in
5 Figure 1 are a first data packet 120 of size p_1 and arriving at time t_1 , a second data packet 125 of size p_2 and arriving at time t_2 , a third data packet 130 of size p_3 and arriving at time t_3 , and a fourth data packet 135 of size p_4 and arriving at time t_4 .

[0021] As shown in Figure 1, the variance in the size and arrival time of the data packets results in certain time periods with little activity and certain time periods with
10 "bursts" of data. The data packets are processed through a flow controller utilizing a scheduling counter, which may be illustrated symbolically as a token bucket 140. In this illustration, the token bucket 140 receives tokens periodically from a token generator 145. The token bucket controller regulates the flow of data. The token bucket controller allows certain bursts of data, such as represented in Figure 1 by the second data packet 125, the
15 third data packet 130, and the fourth data packet 135 arriving in close sequence, but the system limits how large the bursts of data traffic can be. The level of data flow is regulated by the tokens. For each particular number of bytes released to the network, one token is used and removed from token bucket 140. Under a particular embodiment, one token is required for each byte of data that is released. A token generation rate r
20 represents the number of tokens that are generated per unit of time. The token bucket 140 can hold a certain maximum number of tokens b , the token bucket capacity.

[0022] According to the illustration shown in Figure 1, a data packet that has arrived will be released when there are sufficient tokens in the token bucket for the number of bytes contained in the data packet. If there are not sufficient tokens available at
25 a particular time to release a data packet, the data packet is held, or stored in a queue 150,

until token generator 145 generates sufficient tokens for token bucket 140 to allow transmission. Under a particular embodiment, the token bucket 140 includes parameters (r, b) , where r is the token bucket rate and b is the token bucket size. Under an embodiment of the invention, the queue 150 that is utilized for a token bucket is

5 sufficiently large to hold any backlogged data packets. The traffic pattern that is observed will have a certain maximum size of a data packet, which may be denoted as m .

[0023] Under an embodiment of the invention, token generator 145 and token bucket 140 comprise a counter that increments at a constant rate r with a maximum count of b . The counter is decremented for each data packet that is released, with, for example,

10 the count being reduced by one for each byte that is released. If the counter is not at a sufficient count for the release of a data packet, the data packet is held in queue 150 until the counter reaches a sufficient count. However, holding a data packet in queue 150 creates a delay for the data. If there is an upper bound prescribed for data delay, there is a question regarding the minimum size required for the token bucket in order to ensure

15 guaranteed service in which the delay upper bound is not exceeded for any data packet.

[0024] Figure 2 is a block diagram illustrating the operation of a measurement-based traffic specification system 200 under an embodiment of the invention. The system 200 may exist for any observation point in a network. In this illustration, data in the form of data packets is transferred from the preceding point in the

20 network, designated as Point A 205, to the next point in the network, designated as Point B 210. The data traffic will flow through traffic pattern collector 215 and QoS provider 220 on to the second point 210. The traffic pattern collector 215 collects information regarding the traffic pattern and provides the information to a measurement-based traffic specification (MBTS) engine 225. The information may include the number of data

packets received, the size of each data packet, and the time at which each data packet arrived at the traffic pattern collector 215.

[0025] The MBTS engine 225 converts the raw information regarding the data pattern to data flow specifications. Under an embodiment of the invention, the flow specifications determined by the MBTS engine 225 are in the form of optimal token bucket parameters for the data traffic pattern. The results determined by the MBTS engine 225 are then made available to the QoS service provider 220 through a reporting mechanism 230. Note that in this description, in order to facilitate explanation, the traffic pattern collector 215, the QoS service provider 220, the MBTS engine 225, and the reporting mechanism 230 are generally discussed as if each is a single functional unit. However, the functions performed by these various elements may be consolidated and/or distributed differently than as described. For example, any function can be implemented on any number of logical or physical devices or on a single logical or physical device.

[0026] Figure 3 is a block diagram of a exemplary network architecture that may include an embodiment of the invention. Figure 3 is intended as a simplified example, but the embodiments of the inventions may be implemented in many different environments. In Figure 3, a network 300 is illustrated, which may include one or more servers. In this example, the network 300 is shown to include a first server 305, a second server 310, and a third server 315. The network may include one or more clients, shown in Figure 3 as a first client 320, a second client 325, and a third client 330. In this illustration, the network devices are connected by switches 335 and 340 and by routers 345, 350, and 355. In one example, the first client 320 may request certain data. In this example, the data requested may be supplied by multiple sources, including the first server 305 and the second server 310. With data traffic control under an embodiment of the invention, data traffic in the network may be measured at any point in the network. According to an embodiment of

the invention, switch 335 may sense the data traffic associated with the first server 305 and the second server 310. In particular, the amount of traffic at any time may be determined, including whether there are bursts of high traffic levels. Using the traffic data information, the data loads on the first server 305 and the second server 310 may be
5 adjusted or balanced as necessary to, for example, insure that data traffic does not exceed an upper bound prescribed for data delay.

[0027] According to an embodiment of the invention, the third server 315 may be subject to a profile when supplying data to the network, the profile including limitations on the rate of data transmission and the “burstiness” of the data traffic from the server. An
10 observation point may be implemented at switch 340 to determine the traffic pattern characteristics. Under an embodiment of the invention, a data controller may control the traffic to insure that the required profile is maintained while insuring that guaranteed service is provided such that no data is delayed more than a certain upper bound.

[0028] Under an embodiment of the invention, a token bucket is utilized in a data
15 controller for a traffic flow of data packets, the token bucket having a rate that is at least equal to the average data traffic flow. The traffic flow is comprised of n data packets. Each data packet in the traffic flow has a particular size and arrives at a particular time. To determine a size for a token bucket that will ensure that the delay of any data packet in the traffic is no more than a prescribed upper bound, certain calculations are performed for
20 each data packet from 1 to n . Under an embodiment of the invention, the calculations are performed in real time as data packets are observed. The result of such calculations is a size parameter for the token bucket, the size parameter being the minimum token bucket size that will meet the delay upper bound requirement.

[0029] Under an embodiment of the invention, the size parameter for a token
25 bucket for a given delay upper bound is calculated pursuant to an algorithm that may be

denoted as the “Any-Delay TB Algorithm”, where TB indicates a token bucket. Under an embodiment of the invention, the size parameter for the token bucket may be determined as follows:

Token bucket rate = r , where the rate is at least equal to the average data rate for the data flow

D = delay upper bound

b , d , and m are integer variables, where:

b = a placeholder for a token bucket size

d = an intermediate value

m = an intermediate value that will become the maximum packet size for a given traffic pattern

[0030] A data traffic pattern comprised of data packets may be described as:

$$(p_1, t_1), (p_2, t_2), \dots, (p_n, t_n)$$

which indicates a data flow with n data packets, where the i -th data packet has a

size of p_i bytes and arrives at a time t_i . Utilizing this terminology, the

relation of the token bucket rate to the average data rate may be expressed as follows:

$$r \geq \frac{\sum_{i=1}^n p_i}{(t_n - t_1)}$$

[0031] Under an embodiment of the invention, for a given traffic pattern

comprised of data packets, a token bucket size may be determined as follows, as expressed in pseudo-code:

$$b = d = p_1 - r \cdot D$$

$$m = p_1$$

For $k = 2$ to n

$$\begin{aligned}
 & \{ \\
 & \quad d = \max[d + p_k - r \cdot (t_k - t_{k-1}), p_{k-1} + p_k - r \cdot (t_k + D - t_{k-1}), p_k - r \cdot D] \\
 & \quad b = \max(b, d) \\
 & \quad m = \max(p_k, m) \\
 5 \quad & \} \\
 & b = \max(b, m)
 \end{aligned}$$

[0032] Figure 4 illustrates the selection of the size parameter for a token bucket under an embodiment of the invention. According to an embodiment of the invention, a traffic pattern comprised of data packets is observed, process block 405. The first data packet is evaluated, process block 410. A first variable b and a second variable d are initially set at the size of the first data packet minus the rate of the token bucket times the delay upper bound D , process block 415. A third variable m , which ultimately will be equal to the maximum packet size in the traffic stream, is initially set at the size of the first data packet, process block 420.

15 [0033] According to an embodiment of the invention, a series of calculations (or loop) is performed for data packets 2 through n . Beginning with the second data packet, process block 425, each such data packet is evaluated, process block 430. For each of such data packets, the variable d is set to the maximum of three numbers, process block 435. The first number is equal to the previous value of variable d plus the size of the current data packet minus the rate of the token bucket times the difference between the time of the current data packet and the time of the preceding data packet. The second number is equal to the size of the preceding data packet, plus the size of the current data packet, minus the rate of the token bucket times a number equal to the time of the current data packet plus the delay upper bound D minus the time of the last data packet. The third

number is equal to the size of the current packet minus the rate of the token bucket times the delay upper bound D .

[0033] The variable b is set to the maximum of either the previous value of the variable b or the current value of the variable d , process block 440. The variable m is set
5 to the maximum of either the size of the current data packet or the previous value of the variable m , process block 445. As m is set to the maximum of either the current data packet size or the previous maximum size, m will equal the size of the largest data packet after all data packets have been evaluated.

[0035] If there are additional data packets to evaluate, process block 450, the
10 next data packet is chosen, process block 455. After the series of calculations are performed for each of the data packets, process block 450, the token bucket size parameter b is the maximum of either the previous value of the variable b or the current value of the variable m , process block 460. Under an embodiment of the invention, the value of b determined expresses the minimum token bucket size that will satisfy the delay upper
15 bound condition.

[0036] Under an embodiment of the invention, the loop in the algorithm for calculation of the token bucket size parameter operates only $n - 1$ times. For this reason, the complexity of the algorithm in time is $O(n)$, which indicates that the complexity of computation is linearly related to the number of data packets in the traffic pattern. Under
20 an embodiment of the invention, the space required for computation of the token bucket size parameter is comprised of only the three integer variables d , b , and m . Therefore, in space the complexity of the algorithm is $O(1)$, which indicates that the memory required for computation is constant. The linear complexity of computation and the constant memory requirements indicate that the computations utilized to establish the token bucket

size are extremely efficient and light weighted and thus have a minimal impact on system operation.

[0038] The invention has been described in terms of several embodiments. However, those skilled in the art will recognize that the invention is not limited to the 5 embodiments described, but rather that modifications and changes may be made without departing from the broader spirit and scope of the invention. The specification and drawings are thus to be regarded as illustrative rather than limiting.

CLAIMS

What is claimed is:

1. A data controller comprising:
a traffic pattern collector to observe a data traffic pattern, the data traffic being
5 comprised of a plurality of data packets;
a traffic service engine to receive data from the traffic pattern collector concerning
the data traffic pattern and to determine a data flow specification for the
data traffic pattern, the received data including the size and arrival time of
each of the data packets; and
10 a quality of service provider to receive the data flow specification, the data flow
specification representing the minimum data flow specification that will
allow each of the data packets to be transferred by the data controller within
a delay upper bound.
2. The data controller of claim 1, further comprising a reporting mechanism to report
15 the data flow specification to the quality of service provider.
3. The data controller of claim 1, wherein the data controller includes a token bucket,
the data flow specification comprising a minimum token bucket size that complies
with the delay upper bound.
4. The data controller of claim 3, wherein the token bucket size is based at least in
20 part on the sizes and arrival times of the data packets.
5. The data controller of claim 4, wherein data controller considers each data packet
only once in the determination of the token bucket size.

6. A method comprising:
- observing a data traffic pattern, the data traffic pattern being comprised of a stream of data packets;
- determining information regarding the data traffic pattern, the information
- 5 including the size of each of the data packets and the time of arrival of each of the data packets; and
- determining a traffic control parameter value based upon the information regarding the data traffic pattern, the determined parameter value being the smallest parameter value that will enable a network device to transport the data
- 10 traffic without any data packet being delayed more than a delay upper bound.
7. The method of claim 6, further comprising:
- setting a traffic control parameter of a data traffic controller associated with the network device to the determined parameter value.
- 15 8. The method of claim 6, wherein the traffic control parameter comprises a token bucket size.
9. The method of claim 8, wherein calculations to determine the token bucket size are performed as data packets are observed.
10. The method of claim 9, wherein the complexity in time of the calculations to
- 20 determine the token bucket size is $O(n)$ and wherein the complexity in space of the calculations to determine the token bucket size is $O(1)$.

11. A machine-readable medium having stored thereon data representing sequences of instructions that, when executed by a processor, cause the processor to perform operations comprising:
- 5 observing a data traffic pattern, the data traffic pattern being comprised of data packets;
- determining information regarding the data traffic pattern, the information including the size of each of the data packets and the time of arrival of each of the data packets; and
- determining a traffic control parameter value based upon the information regarding
- 10 the data traffic pattern, the determined parameter value being the smallest parameter value that will enable a network device to transport the data traffic without any data packet being delayed more than a delay upper bound.
12. The medium of claim 11, further comprising:
- 15 setting a traffic control parameter of a data traffic controller associated with the network device to the determined parameter value.
13. The medium of claim 11, wherein the traffic control parameter comprises a token bucket size.
14. The medium of claim 13, wherein calculations to determine the token bucket size
- 20 are performed on a real time basis.
15. The medium of claim 14, wherein the complexity of the calculations to determine the token bucket size is linear in time and is fixed in space.

16. A method comprising:

observing a traffic pattern comprising n data packets, the size and time of arrival of

such data packets being $(p_1, t_1), (p_2, t_2), \dots, (p_n, t_n)$; and

determining a size of a token bucket utilized in connection with a data traffic

5 controller, the determined size being the minimum size that will enable the data traffic controller to transfer each data packet within a delay upper bound.

17. The method of claim 16, wherein the rate of the token bucket is r and the delay upper bound is D , and wherein the determination of the token bucket size

10 comprises:

for the first data packet, wherein $k = 1$:

setting values for variable d and a variable b as follows:

$$b = d = p_1 - r \cdot D, \text{ and}$$

setting a value for a variable m as follows:

15
$$m = p_1; \text{ and}$$

for each the second through n -th data packets, wherein $k = 2$ through n ,

performing the following calculations:

$$d = \max[d + p_k - r \cdot (t_k - t_{k-1}), p_{k-1} + p_k - r \cdot (t_k + D - t_{k-1}), p_k - r \cdot D],$$

$$b = \max(b, d), \text{ and}$$

20
$$m = \max(p_k, m).$$

18. The method of claim 17, wherein the token bucket size is variable b calculated as:

$$b = \max(b, m).$$

19. A machine-readable medium having stored thereon data representing sequences of instructions that, when executed by a processor, cause the processor to perform operations comprising:

observing a traffic pattern comprising n data packets, the size and time of arrival of

5 such data packets being $(p_1, t_1), (p_2, t_2), \dots, (p_n, t_n)$; and

determining a size of a token bucket utilized in connection with a data traffic

controller, the determined size being the minimum size that will enable the

data traffic controller to transfer each data packet within a delay upper

bound.

- 10 20. The medium of claim 19, wherein the rate of the token bucket is r and the delay upper bound is D , and wherein the determination of the token bucket size comprises:

for the first data packet, wherein $k = 1$:

setting values for variable d and a variable b as follows:

15 $b = d = p_1 - r \cdot D$, and

setting a value for a variable m as follows:

$m = p_1$; and

for each of the second through n -th data packets, wherein $k = 2$ through n ,

performing the following calculations:

20 $d = \max[d + p_k - r \cdot (t_k - t_{k-1}), p_{k-1} + p_k - r \cdot (t_k + D - t_{k-1}), p_k - r \cdot D],$

$b = \max(b, d)$, and

$m = \max(p_k, m).$

21. The medium of claim 20, wherein the token bucket size is variable b calculated as:

$$b = \max(b, m).$$

22. A computer network comprising:

one or more servers;

5 a plurality of network devices, including a first network device; and

a token bucket data traffic controller sensing data traffic associated with the first network device, the traffic controller comprising:

a traffic pattern collector to observe data traffic associated with the first network device, the data being comprised of a plurality of data packets,

10

a traffic service engine to receive information from the traffic pattern collector concerning the data traffic and to determine a size for the token bucket,

a quality of service provider to receive the determined token bucket size,

15

the token bucket size representing the minimum token bucket size that will allow each of the data packets to be transferred by the traffic controller within a delay upper bound, and

a reporting mechanism to report the token bucket size to the quality of service provider.

- 20 23. The computer network of claim 22, wherein the token bucket size is based at least in part on the sizes and arrival times of the data packets.

24. The computer network of claim 23, wherein the data traffic controller considers each data packet only once in determining the token bucket size.

25. A data traffic controller comprising:
- a means for observing a stream of data packets in a network;
 - a means for determining information regarding the data packets, the information including the size and time of arrival of each data packet; and
 - 5 a means for determining a data flow specification, the data flow specification being the minimum specification that allows each data packet to be transferred within a limit on delay.
26. The data traffic controller of claim 25, wherein the determined data flow specification comprises a token bucket size.
- 10 27. The data traffic controller of claim 26, wherein the token bucket size is based at least in part on the sizes and arrival times of the data packets.
28. The data traffic controller of claim 27, wherein the means for determining the traffic control specification includes a means for considering each data packet only once in the determination.
- 15 29. The data traffic controller of claim 28, wherein the determination of the data flow specification is performed in real time.