

On Rate Adjustment Mechanism for Reliable Multicast Transmission in ForCES

Liang Gong¹, Yining Wang², and Chuanhuang Li¹

¹College of Information & Electronic Engineering Zhejiang Gongshang University,
Hangzhou, Zhejiang Province, China
756695230@qq.com, chuanhuang_li@mail.zjgsu.edu.cn

²College of Computer Science and Technology, Zhejiang University,
Hangzhou, Zhejiang Province, China
3110102753@zju.edu.cn

Abstract. In the specific environment for reliable multicast transmission of ForCES protocol messages, this paper firstly analyzed the congestion control difficulties and problems in the process of the reliable multicast transmission, and then designed and proposed a reliable multicast rate adjustment mechanism which meets the architectural requirements of ForCES routers. Test results also show that the mechanism basically meets the relevant performance requirements.

Keywords: ForCES; Reliable Multicast; Congestion Control; Rate Adjustment.

1 Introduction

ForCES (Forward and Control Element Separation) Working Group [1] is one of IETF working groups in the field of routing, which has committed to the research and development of open programmable router architecture and protocol since its inception. ForCES router architecture is composed by the control unit (CE) and the forwarding unit (FE) in the structure, and CE and FE can transmit information through the ForCES protocol.

Within ForCES routers, one CE needs multiple FEs to transmit protocol messages. Compared with multiple TCP unicasts, IP multicast can save bandwidth between ForCES transport mapping layers. Because the protocol message is only copied at the branching node of the multicast tree, it avoids CE producing redundant protocol message packets, and reduces the load of the CE port and improves the transmission efficiency. However, as the multicast transmission is based on UDP protocols, CE transmits protocol messages at a fixed rate, if the network environment is deteriorated, the sending rate of CE cannot be adjusted to the protocol messages flow, which may lead to a further deterioration of the network environment, or even a collapse. There are already a number of multicast congestion control protocols, such as Pragmatic General Multicast Congestion Control (PGMCC) [2], TCP-friendly Multicast Congestion Control (TFMCC) [3], Receiver-driven Layered Multicast (RLM) [4], and other agreements and so on.

Currently, there are mainly two types of reliable multicast transport protocols which are based on acknowledgment (ACK) or based on negative acknowledgment (NACK). Reliable multicast protocol messages within ForCES adopt the mechanism of error detection and recovery, which is based on the negative acknowledgment (NACK), and it can better ensure the reliable transmission of multicast protocol messages.

In order to ensure the friendly fairness to TCP, Ma *et al.* [5] proposed Fair Active Congestion Control (FACC), which starts with a single rate by the receiver, and takes a transmission rate as congestion control parameters. In order to increase the robustness and adaptivity of the algorithm, the protocol took congestion parameter convergence strategy to redesign the filtering algorithm of the congestion control parameter. To eliminate the phenomenon of “slowest priority”, Zhai, Wu and Gu [6] proposed two mechanisms that are Composite Multicast Congestion Control (CMC) and Layered Multicast (LM) Congestion Control by the Particle Swarm Optimization (PSO) (LM-PSO). Both of these mechanisms do not use a single-rate congestion control algorithm, but use an algorithm where each layer of the Layered Multicast adjusts the transmission rate dynamically.

In this paper, in view of the research of the Multicast Congestion Control Mechanisms, we have taken the detailed analysis and research about the issues of the reliable multicast rate adjustment mechanisms within ForCES, and adopted the Wireshark capturing software to test the traffic conditions of the CE port during the process of protocol messages reliable multicast. What is more, by drawing a graph of throughput situations of data packets, we analyzed the problem of reliable multicast about fairness and heterogeneity within ForCES.

2 Main Problem of the Reliable Multicast Rate Adjustment Mechanism within ForCES Routers

Similar to the TCP congestion control, currently the ForCES routing channel protocol message is based on multicast on IP layer. The reliable multicast protocol message within ForCES adopts the point-to-point rate adjustment mechanism. Therefore, protocol messages reliable multicast rate adjustment mechanism within ForCES can learn from the TCP congestion control mechanism absolutely, research from the regulator parameters, feedback mechanisms, and other aspects of rate regulation algorithm. However, considering the environment of reliable multicast is much more complex than that of unicast, TCP congestion control mechanisms cannot be applied simply on reliable multicast, and the specific scenarios of reliable multicast within ForCES must be combined to carry out a detailed analysis and design for each point.

Another issue of the reliable multicast rate adjustment mechanism within ForCES that must be solved is fairness, which appears much more important in multicast environment. For reliable multicast, the fairness of congestion control mechanism has two meanings. On the one hand, the fairness is that, the protocol message stream that transmits in the form of TCP shares the network bandwidth with a reliable multicast stream. On the other hand, the fairness is that each FE of multicast receivers is fair. For example, the sending rate of CE cannot be decided by the FE which has the worst network condition, and this is obviously unfair for other FE ports. Therefore, the rate adjustment

mechanism of reliable multicast protocol messages within ForCES must satisfy the heterogeneity between each FE. In order to achieve intra protocol fairness, and to ensure the sufficient and effective use of the network bandwidth of ForCES transport mapping layer, all of these problems should be considered and designed in the research of this paper.

3 Analysis of Concrete Problems of Reliable Multicast Rate Adjustment Mechanism within ForCES Routers

3.1 Preferences of Reliable Multicast Rate Regulation

The CE sender of the reliable multicast within ForCES relays on the hardware platform, the transmission rate of which is impacted by the processor hardware itself. If adopting the rate adjustment mechanism based on the sending frequency, the network processor of CE should be made certain modifications on the kernel, and every rate adjustment process should be switched between the kernel space and the user space constantly, which will increase the expense of time, and cannot meet the real-time requirement of the reliable multicast rate adjustment mechanism. Therefore, the reliable multicast within ForCES adopts a rate adjustment based on a window parameter. In order to adjust the rate, and to achieve the rate adjustment mechanism by the software, and to avoid modifying hardware, CE can adjust the size of the sending window according to the feedback messages sent by FE, so that the whole rate adjustment mechanism can have greater flexibility and feasibility.

Referring to TCP congestion control mechanisms, during the process of reliable multicast within ForCES, the size of a sending window maintained by CE is set as W_0 , and then simulating the AIMD mechanism to adjust the size of the sending window. For reliable multicast — a one-to-many model in transmission, CE cannot grasp the network congestion of FE timely and accurately. When the size of sending window increases in an additive way, it will become too large and cause a network congestion. As this result, we cannot introduce TCP congestion control mechanisms to the multicast environment simply.

We use the method that setting a sending threshold denoted as W' at CE, which is used to adjust the size of the sending window. The sending threshold is the maximum size of the sending window at CE during the process of reliable multicast, that is to say, when the size of the sending window has reached the threshold, namely W' , if we continue to increase the size of the sending window, the efficiency of multicast will not be improved. So a sending threshold can be set according to the size of the measured values for each multicast test or empirical values. Thus, the adjustment parameter for CE should be set between the sending window and the threshold, the sending window represents the size of each multicast initialization process, namely the size of the minimum window. If the size of the sending window is less than this value, FE may receive multicast messages abnormally, which will waste the multicast network bandwidth, so that the normal performance of ForCES routers cannot be guaranteed.

In conclusion, the parameter selection of the reliable multicast rate adjustment within ForCES bases on the following principles: the initial sending window size of CE, denoted as W_0 . When W_0 is lower than the available network bandwidth, denoted

as $W1$, increases the size of sending window rapidly; when $W0$ is between $W1$ and W' , increases the size of sending window slowly to avoid the network congestion; when the network is congested, reduces the size of the sending window, but when $W0$ is between $W1$ and W' , decreases slowly to avoid the rate oscillation.

3.2 Fairness Problem of Reliable Multicast Transmission

Among reliable multicast protocols within ForCES, the control protocol message sent in the form of multicast competes with the control protocol message within TML in the bandwidth in the rigid way. When the network environment becomes deterioration, the control protocol message flow transmitted in the form of TCP will reduce the sending rate according to the AIMD mechanism, but the multicast protocol message flow will continue being transmitted in the original rate, or even a greater rate. Because of the lack of the corresponding rate adjustment mechanism, it will lead to the collapse of network bandwidth. Therefore, in order to achieve the fairness of TCP, a window mechanism similar to TCP's should be taken, but the AIMD window adjustment mechanism of TCP has a great jitter [7], which is not conducive to transmitting the multicast protocol message. So the AIMD mechanism cannot be applied simply in the multicast environment within ForCES. In order to get an ideal control curve of the sending rate, we should improve the AIMD mechanism, for example, we can add a series of adjustment rules based on the history records to make it smoother.

In the curve of the TCP AIMD mechanism, the increasing factor is set as b , the multiplicative factor is set as a , the size of the sending window is a time function whose feedback time interval is set as t , during the cycle of the sending window, the function is denoted as follow:

$$W(t) = \begin{cases} W_0 * a^t & 0 < a < 1 \\ b * (t - t_0) & b > 0 \end{cases} \quad (1)$$

Herein, W_0 is the initial value of the sender congestion window, t_0 is the initial time, t is an integer multiple of the feedback time which is also called RTT , namely, $t = n * RTT$ ($n = 0, 1, \dots$), so the curve is discrete. However, in order to facilitate the analysis, ignoring the effect of timeout and retransmission during the practical transmission, the discrete curve can be simplified as a continuous curve.

From Equation (1), the control curve of AIMD mechanism under the congestion condition obeys the exponential function curve with a as its base, when t is large enough, the size of the window is close to 0, but under the non-congestion conditions it obeys the linear function curve with b as its slope.

According to above analysis, the exponential curve of AIMD mechanism declines much more rapidly at first, when the value is close to the lower limit, it declines slower, the characteristic of which is more appropriate for the protocol message multicast transmission within ForCES. Considering that, during the protocol message multicast transmission process, the initial sending window of CE is denoted as W_0 , the sending window function of CE under the congestion condition is adjusted as:

$$W(t) = W_0 * a + W_0 \quad (2)$$

At the sending window adjusting algorithm of CE, we adopt the multiplicative decrease method under the congestion condition:

$$W(t) + RTT = a * (W_t - W0 + W_0) \quad 0 < a < 1. \quad (3)$$

In order to be friendly to TCP, we adopt the method which uses a maximum value to limit the additive increase under the non-congestion conditions:

$$t + RTT = \min \{W', W_t + b\} \quad b > 0. \quad (4)$$

Under this window adjustment algorithm, the curve of the transmission rate is linear increasing, and will stop at the sending threshold, denoted as W' . As a result, it cannot achieve the initial requirements. Therefore, the rising amplitude of the additive increasing function curve changes from the constant to the variable, the rising amplitude of which is related to the difference between its current value and sending threshold. So we can design as follow:

$$W_t + RTT = W_t + b(W' - W_t) \quad 0 < b < 1. \quad (5)$$

We can know from above, the function of Equations (5) and (3) is complementary, and the effect is the opposite, both of which have decided the sending window curve rising at a faster rate at the initial time, and then the rate of the rising becomes much slower due to the value of the sending window is close to the sending threshold. When b is set as an appropriate value, we can get an ideal performance of the curve, as shown in Fig. 1. By the testing, we can see that if the mechanism of the AIMD is improved, and considered strengthening the smoothing of the original curve, it will have smaller amplitude than that of the original curve. However, it is still the additive increase and multiplicative decrease in theory, and friendly to the TCP. So it can satisfy the fairness requirement perfectly.

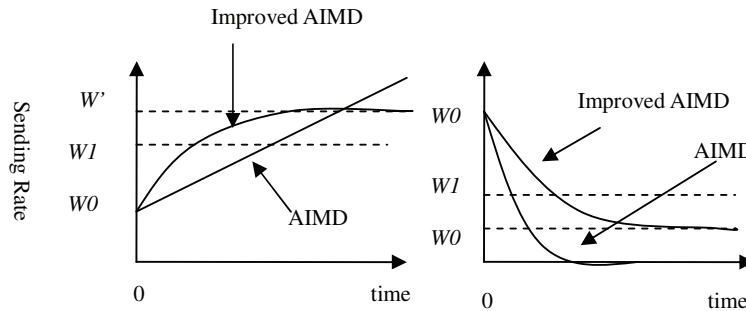


Fig. 1. Rate curve compared between AIMD and improved AIMD.

3.3 Principles of Feedback and Rate Adjustment Mechanism

The reliable multicast within ForCES takes the way of feedback (which combines with the network congestion trends according to the network bandwidth) to adjust the sending rate.

The principles which should be adopted during the reliable multicast within ForCES are as follows:

1. FE makes the congestion judgment according to the average packet-dropout interval packets during each multicast process, and sends feedback to CE;
2. CE makes the judgment about the trend of the network bandwidth by grasping the network congestion of FE and combining the current sending rate with the historical rate during the last multicast, and adjusts the sending rate according to the change of the network bandwidth;
3. CE updates the history records of multicast after every sending rate adjustment and keep the history records, in preparation for the next sending rate adjustment.

In order to reflect the congestion situation of the current network correctly, we adopt an exponential weighted moving average method for the historical packet-dropout rate to judge the condition of the network congestion.

In order to assist the judgment of the network bandwidth variation trend, the reliable multicast within ForCES needs keeping a group of history records, denoted as $H = \{S_c, S_{nc}\}$, wherein, S_c stands for the control status of the sending rate under the recent condition of the multicast congestion, and S_{nc} stands for the control status of the sending rate under the recent condition of the non-congestion. At the beginning of the reliable multicast, the trends of the network bandwidth of CE, denoted as B , are documented as follows:

- $B = 0$: The available network bandwidth is stable;
- $B = 1$: The available network bandwidth is increasing;
- $B = -1$: The available network bandwidth is decreasing.

The trend of network bandwidth is judged by the rules which are as follows:

1. At beginning, the size of the sending window at CE is denoted as W_0 , if receiving the congestion feedback from FE during the multicast transmission process, and the size of the current sending window, denoted as W , is less than the value of S_c in the history of H , the network bandwidth will be considered decreasing, so $B = -1$;
2. If not receiving the congestion feedback from FE during the multicast transmission process, and the size of the current sending window, denoted as W , is larger than the value of S_{nc} in the history of H , the network bandwidth will be considered increasing, so $B = 1$;
3. If the size of the current sending window of CE, denoted as W , is between S_c and S_{nc} , or does not meet the two necessary conditions of above two at the same time, the network will be considered stable, so $B = 0$.

As a result, the reliable multicast rate adjustment algorithm within ForCES, which is based on the sending window, is as follow:

1. If the network bandwidth variation trend is identified as $B = -1$, updating the size of the sending window according to Equation (5);

2. If the network bandwidth variation trend is identified as $B = 1$, updating the size of the sending window according to Equation (3);
3. If the network bandwidth variation trend identified as $B = 0$, updating the size of the sending window according to Equation(6);

$$W = k * S_c + (1-k) * S_{nc} \quad 0 < k < 1. \quad (6)$$

The history records of CE will be updated after the sending rate being adjusted. If $B = 0$, this problem will be divided into two parts, if the packet-dropout feedback of FE is not received, and the value of S_{nc} in the history of H is less than the size of the current sending window, the value of S_{nc} will be updated to equal the size of the current sending window; if the packet-dropout feedback of FE is received, and the value of S_c in the history of H is larger than the size of the current sending window, the value of S_c will be updated to equal the size of the current sending window. If $B = -1$, the value of S_c will be set as W_0 , and the value of S_{nc} will not be changed. If $B = 1$, the value of S_{nc} will be set as W' , and the value of S_c will not be changed. So the multicast message curve diagram of CE through the simulation analysis can be shown in Fig. 2:

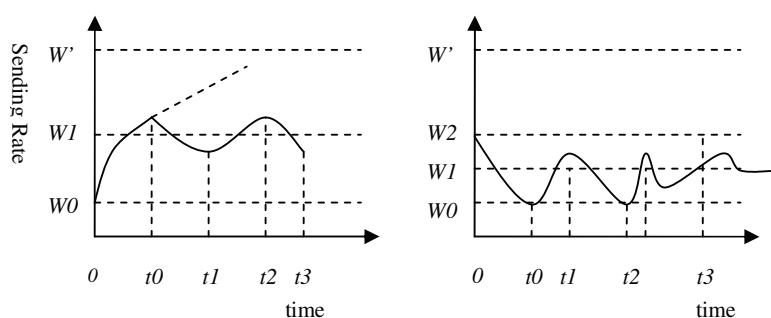


Fig. 2. CE sender multicast message curve.

3.4 Heterogeneity Problem of Reliable Multicast Transmission

Heterogeneity is also called the intra protocol fairness. In the current design of the reliable multicast congestion control, the sender selects its sending rate according to the bandwidth requirement of the worst receiver, namely the sending rate of the receiver is the minimum transmission rate, which is called the Worst-path fairness [7]. The transport model of ForCES protocol messages reliable multicast is an one-to-many communication model, where there are many FE ports at the multiple receiver, and each FE has a different utilization status to the network bandwidth in the different path, and a FE receiver in congestion should not affect the rate of the other FE receivers to receive the multicast packets, and allows each FE receiver to select the received bandwidth according to the congestion status of its own network. So this is the intra protocol fairness, which the reliable multicast rate adjustment mechanism within ForCES should be able to meet with.

The rate regulation feedback mechanism of the protocol message multicast within ForCES adopts the mechanism which selects a set of worse FE receivers to represent a multicast congestion conditions. During the reliable multicast process of intra protocol messages, initially, CE sends the Session Messages to FE, and FE does not calculate the packet-dropout at the same time, so FE does not send the feedback message. When the second session message arrives, every FE sends the ACK feedback packets which include the packet-dropout during the last multicast process and the RTT timestamps. Among them, because the transmission of session messages between CE and FE is the one-way, the session message is transmitted by TCP. But, each FE forwarding element may serve as an important role in the architecture of ForCES router, as a result, each FE must be ensured receiving protocol messages reliably. However, because the function of the ACK feedback message is to feedback the congestion status to CE, the feedback message is transmitted by UDP. As CE has the overall grasp of the receiving capacity and throughput of each FE to coordinate the rate adjustment and avoid the congestion, losing the ACK feedback message is allowed under the premise of satisfying certain QoS.

We select a group of ACK feedback messages from the session packets dynamically every time, which is required not only to represent the bandwidth utilization and throughput under poor reception conditions of FE, but not to cause an intra protocol to be unfair because of the excessive inhibition of the sending rate of CE. As shown in Fig. 3:

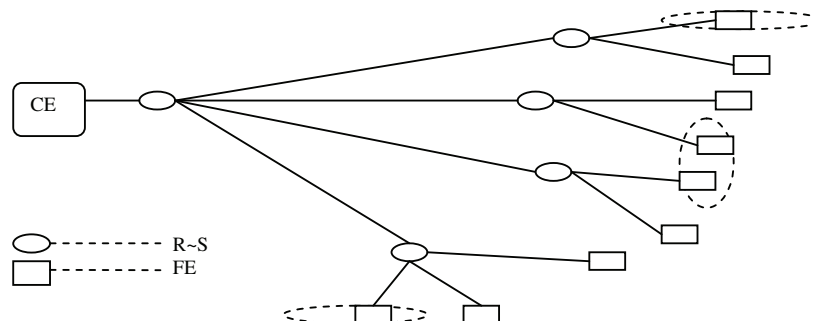


Fig. 3. Feedback set of multicast within ForCES.

4 Performance Testing

4.1 Fairness Testing of Reliable Multicast Transmission

Operating the reliable multicast testing procedures at FE and CE separately, the CE test program joins a multicast group through the IGMP protocol, of which the ID is 0xC0000000, the IP address is 233.4.4.4 and the members include three hosts, whose IP addresses are 10.20.0.59, 10.20.0.190, 10.20.0.229 respectively, and sends 1000 multicast protocol messages to 0xC0000000. On the other hand, setting the parameters

of the SmartBits600 to make it send a UDP message at a rate of 10 Kb/s within ForCES, and keep sending for 20 min. Then, in order to reflect that the process of the actual network status is turning good from bad, continue to send the random constructed UDP messages within ForCES at a rate of 50 Kb/s. The Wireshark software is used to capture the packets at the CE multicast protocol message port, and the result is shown in Fig. 4:

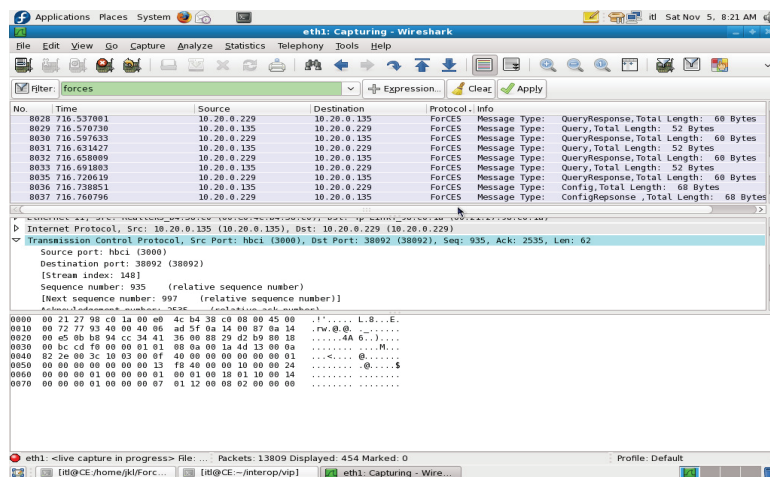


Fig. 4. Protocol message transmission within ForCES.

From the capturing result, during the first 20 minutes, the number of the multicast protocol message packets did not start from a small value, and soon increased to a more stable level in the initial stage, which is the result of that, CE adopted the sending window based on the history records and adopted the mimic AIMD mechanism which combines the congestion judgment and the trend of the network bandwidth, namely the adjustment parameters take use of the empirical values, which can make the reliable multicast sending window return to the optimal window size more quickly and smoothly. 20 minutes later, the size of the multicast messaging window will soon become a smaller level, and this size will be kept continuously after that. The reason why the sending rate is not excessive inhibition is that the congestion feedback mechanism based on the representative feedback set makes CE check the congestion and reduce the sending rate quickly.

4.2 Heterogeneity Testing of Reliable Multicast Transmission

Operating the reliable multicast testing procedures at FE and CE separately, the CE test program joins a multicast group through IGMP protocol, of which the ID is 0xC0000000 and the IP is 233.4.4.4, and the members include three hosts, of which the IP is 10.20.0.59, 10.20.0.190, 10.20.0.229 respectively, and sends a multicast protocol message with a length of 8000 byte to 0xC0000000. On the other

hand, the UDP message constructed by SmartBits600 transports at the rate of 200 Kb/s, to simulate the actual network environment, the send status is shown in Fig. 5:

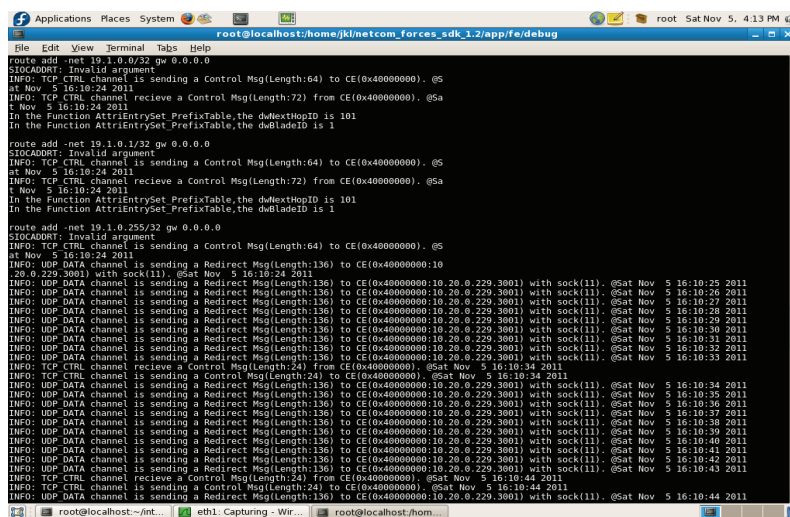


Fig. 5. Feedback situation of the feedback represents.

From Fig. 5, during the multicast process, FE sent session messages in the form of UDP to respond to ACK messages periodically; the periodic time was set nearly as long as the cycle timer of the session message, but the feedback ID had a slightly difference. As a conclusion, the feedback mechanism based on the feedback representative set has a more sensitive feedback characteristic and a less feedback oscillation than those based on the worst receiver.

5 Conclusion

On the basis of current research status of multicast congestion control, because of the lack of the rate regulatory mechanisms within the original ForCES, this paper puts forward a kind of rate adjustment mechanism of CE based on an improved AIMD mechanism, which meets the demands of fairness within ForCES. Through the research of the heterogeneity of the receiving port, this paper also designs a congestion feedback mechanism based on the feedback representative set and this mechanism can also meet the intra protocol fairness better.

Acknowledgments

This work was supported in part by a grant from the National Basic Research Program of China (973 Program) (No.2012CB315902), the National Natural Science

Foundation of China (No.61402408,61379120,61170215), Zhejiang Leading Team of Science and Technology Innovation (No.2011R50010-03, 2011R50010-12, 2011R50010-20), Zhejiang Provincial Key Laboratory of New Network Standards and Technologies (NNST) (No.2013E10012).

References

1. Halpern, J., Salim, J.H.: “Forwarding and Control Element Separation (ForCES) Forwarding Element Model”, RFC 5812, Mar. 2010, <http://www.ietf.org/html.charters/forces-charter.html>
2. Rizzo, L.: PGMCC: A TCP-friendly Single-rate Multicast Congestion Control Scheme[C], in Proceedings of ACM SIGCOMM, 2000.
3. Fairhurst, K.: “TFMCC Protocol Behaviour in Satellite Multicast with Variable Return Path Delays” IEEE 2006.
4. McCanne, S., Jacobson, V., Vetterli, M.: Receiver-driven Layered Multicast, in: Conference proceedings on applications, technologies, architectures, and protocols for computer communications, vol. 26, ACM, SIGCOMM, 1996, pp. 117–130.
5. Ma, H., Meng, X., Ma, Z., Li, J.: Layered Multicast Congestion Control by Particle Swarm Optimization in Heterogeneous Environment [J]. The Journal of China Universities of Posts and Telecommunications, 2011, 18(3):85–91.
6. Zhai, M., Wu, G., Gu, G.: Internet Congestion Control and Reliable Multi-point Delivery Progress [J]. Computer Research and Development, 2000, Vol. 37, No. 1.
7. Bhattacharyya, S., Towsley, D., Kurose, J.: The Loss Path Multiplicity Problem in Multicast Congestion Control [C]. Proceedings of IEEE INFOCOM, New York, March 1999, vol. 2, pp. 856–863.
8. Xie, F., Feng, G.*, Siew, C.K.: Study on Nominee Selection for Multicast Congestion Control [J]. Computer Communications 29(2006) 1458–1469.