System Resource Allocation Algorithm for Multicast Service in NGN

Seng Kyoun Jo[1], Won Ryu[1], Seong Gon Choi[2] and Jun Kyun Choi[3],
([1]ETRI, 161 Gajeong-dong, Yuseong-gu, Daejeon 305-700, Korea, Tel: +82-42-860-6461, Email: skjo@etri.re.kr
[2]Chungbuk National University, 12 Gaeshin-Dong, Heungduk-Gu, Chungbuk 361-763, Korea,
[3]ICU, 119 Munji-Dong, Yuseong-Gu, Daejeon 305-732, Korea)

Abstract

It is necessary to control service requests with definite criteria in order to ensure end-to-end guarantee of Quality of Service (QoS) in the constraints limited system capacity. For including the aspect of system operation in the performance index, we define a system cost function as 'GAIN' with the terms 'reward' and 'penalty' respectively. Thus we find the allocation set which maximizes the system GAIN and the service guaranteed conditioning. For the performance analysis, we show that the system can be modeled as M/G/m/m with limited capacity, investigate GAIN under various traffic conditions in context of NGN realization.

1 Introduction

Multimedia service mentioned as a representative service in NGN, usually takes up lots of resources because of its characteristic such as the amount of data transferred and specified user groups, etc. Hence for supporting these kinds of service such as IP TV and Video-on-Demand (VoD), multicast technology is usually deployed in order to make the best utilization of limited resources. Deploying multicast for convergent solutions from QoS-guarantee perspective to user demands for specified user groups is an issue that needs to be effectively addressed for NGN realization.

To address mentioned issue, we propose a system resource allocation algorithm for supporting multimedia services efficiently using multicast technology based on limited system capacity. Considering the aspect of the performance of the system, which has the authority for controlling service requests, three classes of service priority are assumed in this paper.

We describe briefly network architecture, system modeling in chapter 2 and cost model in chapter 3 and propose system resource allocation algorithm. In chapter 4, we perform the performance analysis to show the optimal system resource allocation set and how it is dependent with traffic conditions.

2 Network Architecture and System Model

For the multicast service in NGN, it is assumed there are one service provider and multiple receivers. Service request is usually controlled by the ingress system. However, in a practical integrated NGN deployment scenario, the egress system instead of the ingress one would execute admission control under the assumption that transmission between the ingress and egress system is reasonably guaranteed. This is compatible with traditional models of ingress -controlled admission control.

For the one service, we start with three kinds of service requests: Guaranteed Service (GS), Premium Service (PS) and Best Effort Service (BES) in the order of descending priority (GS >PS >BES). Based on the priority, service degradation should be considered in this paper.

For the egress system modeling, we assume that the egress it has the authority to limit the client requests per class and allocates the total system capacity N into n_{GS}, n_{PS} and n_{BES} sub-capacities corresponding to GS, PS and BES respectively. Within

89-955301-4-6 98560

a class, admitted service requests occupying system capacity can be serviced equally by means of the multicast copy function performed in egress system. Accordingly, all service requests are assumed to be Poisson processes with rate λ and service time is generally distributed [1, 2] with mean $1/\mu$. Each request requires one unit capacity directly mapped to the unit client. Based on the above assumptions, the system can be modeled as a M/G/m/m queueing system [3] which can maximally service N clients considered to comprise three parallel subsystem such as $M/G/n_{GS}/n_{GS}$, $M/G/n_{PS}/n_{PS}$ and $M/G/n_{BES}/n_{BES}$. $M/G/n_{GS}/n_{GS}$ queue is for GS clients only with the arrival rate equal to λ_{GS}, service rate is $i\mu$ where there are i GS clients being served, and the number of GS multicast sessions equal to n_{GS}. Subsystems for PS and BES can be applied as the GS queue structure except that there can be service requests more than one class according to the service degradation policy.

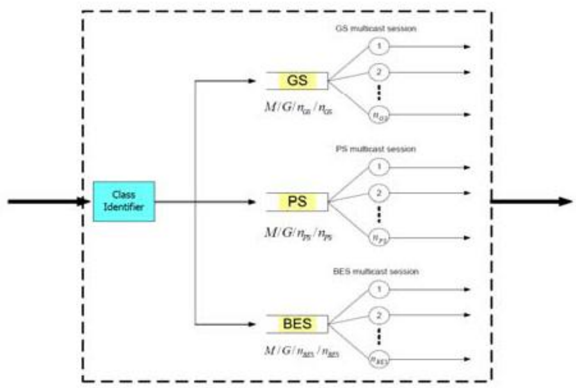

Fig. 1 Egress System Modeling

Fig. 1 illustrates the egress system modeling comprising of three sub-systems which administrate admission control for service request.

3 Cost Model

The performance index being considered takes both 'rewards' and 'penalties' of clients into consideration. It is termed as the system total pay-off rate and referred to as 'system GAIN'. It is first introduced in [4, 5] as the average amount of reward received by the system per unit time of serving a request.

As the aspect of system, reward and penalty is assigned to a value when the client is serviced successfully or rejected. The value is expressed as the weighting factor by letting the reward and penalty of BES be absolute value '1'. For example, if the system on average services N_X clients per unit time and rejects M_X clients per unit time, then the system GAIN is

$$\sum N_X R_X - \sum M_X L_X \tag{1}$$

R_X and L_X denoting the reward and penalty (loss) of X where X = GS, PS and BES.

4 Performance Analysis

Having described details of our revenue based resource allocation algorithm, we now turn to evaluating its performance using numerical analysis.

A. System GAIN

Now we analyze the performance index system GAIN. For the calculation of degraded service request, we consider the loss model of M/G/m/m system. Where the average arrival rate is λ and service time is X, the M/G/m/m loss model and its equilibrium state probabilities are defined in [3] as

$$p_k = \frac{\frac{\rho^k}{k!}}{\sum_{n=0}^{m} \frac{\rho^n}{n!}} \qquad k = 0, 1, \cdots, m \tag{2}$$

where $\rho = \lambda/\mu$ and the blocking probability $P_B = p_m$. Using (1) and (2), we can derive the GS, PS and BES sub-system GAINs and total system GAIN by summing each sub-system GAINs.

B. System GAIN

Fig. 2 shows the example of system GAIN under the condition where system capacity N is 100, request arrival rates of GS, PS and BES are equally 40, mean service time is 1, the reward and penalty ratio of BES:PS:GE are equal to 1:2:4 and 1:2:3 respectively. Then we find the optimal allocation set as (43, 45, 12) which maximizes the revenue and its corresponding system GAIN is 203.08.

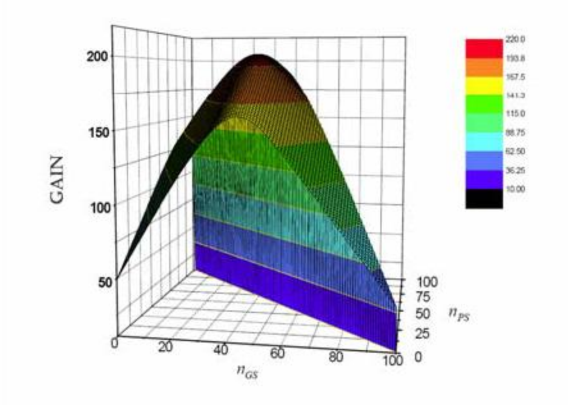

Fig. 2 Distribution of System GAIN

In addition, we investigate the optimal allocation pattern when the traffic intensity changes. Fig.3 shows the result under the condition traffic load is light as λ_{GS} (GS service request arrival rate) increases. Notice that light traffic load is assumed when the service request arrival per each class is equal to 40. Also the heavy traffic load denotes when the service request arrival per each class is equal to 80 respectively. The resource for GS is linearly intended to increase according to λ_{GS}. However, the resource for PS and BES is reducing slightly. For the case of heavy traffic condition pictured in Fig.4, the resource for GS is also linearly risen according to λ_{GS}. On the other hand, the BES resource is not allowed to service due to the reward ratio. Economically says, the opportunity cost for BES is much smaller than one of GS and PS class. Thus the resource for BES is apt to be neglect as λ_{GS} increases when the traffic load is heavy.

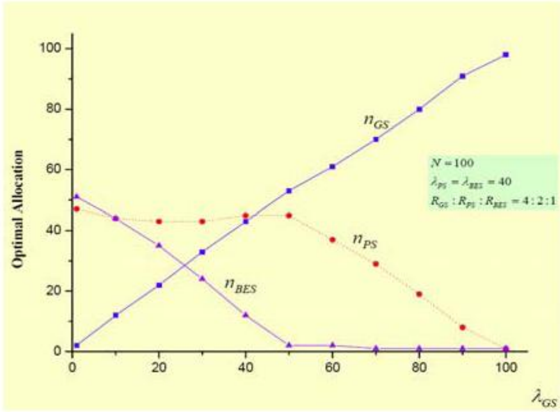

Fig. 3 Optimal Allocation Set for Heavy Traffic Load

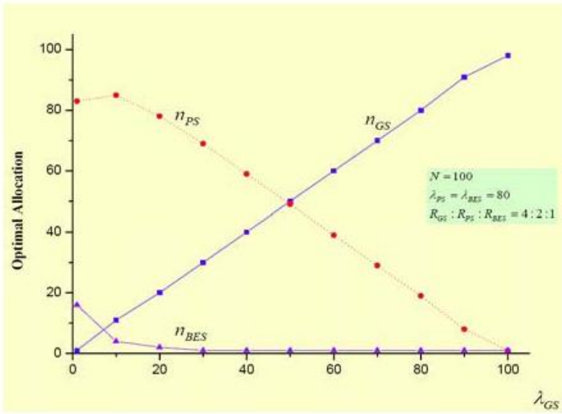

Fig. 4 Optimal Allocation Set for Light Traffic Load

5 References

1. M. Crovella, C. Lindemam and M. Reiser: Internet Performance Modeling: the state of the art at the turn of the century, Performance Evaluation, Vol. 42(2-3), Jan 2000, pp. 91-108
2. B. Jesus, Martinez Jose M and C. Guillermo: Traffic Analysis on Multimedia Data Service, Proceeding of SPIE, Vol. 2917, Nov 1996, pp. 316-327
3. D. Grosz and S. Kraus.: Fundamentals in Queueing Theory, 2nd Edition, Wiley (1985)
4. Ing-Ray Chen and Chi-Ming Chen.: Threshold-Based Admission Control Policies for Multimedia Servers, The Computer Journal, Vol. 39,1996, pp. 757-766
5. Ing-Ray Chen and Tao-Hung His.: Performance analysis of admission control algorithms based on reward optimization for real-time multimedia servers, Performance Evaluation, Elsevier Science, Vol. 33, Number 2, July 1998, pp. 89-11