QUEUEING SYSTEMS WITH FLOW ARRIVAL OF CUSTOMERS

S.A. Dudin

Belarusian State University,
4, Nezavisimosti Ave., 220030 Minsk, Belarus
dudin@madrid.com

Introduction. A standard assumption in queueing theory is that the input process is a point process defined by joint distribution of arrival epochs or inter-arrival times. At arrival epoch a batch of customers arrives into the system and it is admitted into the system depending on the system admission strategy and the system state. However, the typical feature of many nowdays telecommunication networks (IP networks particularly) is the flow (stream) arrival of customers. This means that the arrival of customers belonging to one entity is distributed in time. At the moment of a new flow arrival, the number of customers, which will arrive in this flow, and moments of their arrival are not known. The system manager has to decide whether or not to admit the

arriving flow taking into account that the rejection of a flow causes reduction of the system throughput and possible starvation of the servers in the future while the excessive admission of flows can cause the increase of the mean customer sojourn time and the jitter and loss probability of customers from admitted flows. One of possible mechanisms for admission control is token mechanism which assumes admission of a flow only in the presence of a free token from the finite pool of tokens, see [1]. Thus, the problem of admission strategy optimization is reduced to the optimal choice of the capacity of the pool.

Mathematical model and its analysis. The simplest model of the system with flow arrivals was investigated in [2]. The queueing system has a finite capacity. Service time has an exponential distribution. The customers arrive in flows according to a stationary Poisson arrival process. The total number of tokens in a pool is assumed to be fixed. If there is no token available, the batch is rejected. If the number of available tokens is positive and the system is not already full, this flow is admitted into the system and the number of available tokens decreases by one. It is assumed that one customer of a flow arrives at the flow arrival epoch and if it meets at least one free server, it occupies this server and is processed. If all servers are busy, the customer moves to a buffer, if it exists, and later it is picked up for the service. After admission of the flow, the next customer of this flow can arrive into the system in exponentially distributed time. If the system is not full, the customer is admitted into the system, otherwise, it is lost. The number of customers in a flow has geometrical distribution. If the exponentially distributed time since arrival of the previous customer of a flow expires and new customer does not arrive, it means that the arrival of the flow is finished. The token, which was obtained by this flow upon arrival, is returned into the pool of available tokens. Analysis of this system was implemented via consideration of a two-dimensional continuous time Markov chain having components describing the number of customers and busy tokens in the system at arbitrary time epoch.

In this presentation, results by [2] are extended to the systems with the Markov Arrival Process of flows, Phase type service time distribution, infinite buffer, possibility of a batch customer arrival within an admitted flow, arbitrary distribution of the number of customers in a flow, retrials of rejected customers from accepted flows, etc.

References

1. *A.A. Kist A.A., Lloyd-Smith B., Harris R.J.* A simple IP flow blocking model // Performance Challenges for Efficient Next Generation Networks. Proceedings of 19-th International Teletraffic Congress, 29 August – 2 September 2005, Beijing. 2005. P. 355–364.

2. *Lee M.H., Dudin S., Klimenok V.* Queueing model with times-phased batch arrivals // Managing Traffic Performance in Converged Networks – The Interplay of Convergent and Divergent Forces. Proceedings of the 20th International Teletraffic Congress, 17–21 June 2007, Ottawa. 2007. V. 4516. P. 716–730.