Real-Time Performance Modeling of Link Layer Protocols for Multi-Layer Protocol Aggregation

Paul.J. Kuehn
Institute of Communication Networks and Computer Engineering (IKR)
University of Stuttgart
Stuttgart, Germany
paul.j.kuehn@ikr.uni-stuttgart.de

***Abstract*—A hybrid modeling method for real-time performance evaluation of link layer protocols is presented which makes use of task graph structures, Petri Net synchronization elements, general stochastic arrival/service processes, and channel error characteristics. The resulting models are analyzed exactly by probabilistic task aggregations leading to a separation of the protocol and queuing functions by which the protocol model is stepwise reduced to an aggregated frame transit time representation acting as a virtual service time T_x of a standard queuing model GI/G/n. The methodology is applied to two classical communication protocols: (1) the Stop-and-Wait (SW) Protocol and (2) the Selective-Repeat (SR) Protocol, both with positive Acknowledgements and Timeout Recovery (ACK/TO). The method is applied to the performance analysis of multi-layer architectures to reduce complexity and increase accuracy. The method is demonstrated for Networked Control Systems (NCS) where the link layer delay is embedded within the control loop by an equivalent stochastic phase and where delay threshold percentiles have to be guaranteed. The method is also a key to the performance evaluation of multi-layer protocol architectures where a lower layer subsystem is aggregated into a stochastic phase which can be inserted in the next higher protocol layer, applied layer by layer repeatedly.**

Keywords—Protocols, real-time performance, task graph, stochastic processes, queuing models, performance analysis, mean delays, delay percentiles, Service Level Agreements, Stop-and-Wait protocol, Selective-Repeat protocol.

1. INTRODUCTION AND STATE OF THE ART

Communication networks appear within a distributed control system as an "embedded system" between the remote physical plant and the controller ("Networked Control System", NCS), where sensor signals are sampled at the physical plant and transported through the communication network to the controller, become processed by the controller and where control commands are transported back to the physical plant to adjust its operation accordingly [1]. The real-time performance of the applied communication protocol affects the performance of the NCS directly and is subjected to service performance criteria prescribed by Service Level Agreements (SLAs). There exists a huge variety of communication protocols and corresponding studies on their performance, but comparatively few studies have addressed their real-time performance, i.e., methods to analyze and to guarantee latency percentiles.

A large number of publications has appeared during the last decades on link layer protocols addressing specific questions or conditions. For general descriptions we refer to standard literature, as, e.g., [2]. For specific studies we refer to the "Alternating Bit Protocol" (ABP) described by Stochastic Petri Nets (SPNs). Refined methods of these kinds have been repeatedly reported for **A**utomatic **R**epeat Re**Q**uest (ARQ) protocols with synchronizations between sending and receiving protocol engines by positive/negative acknowledgements (ACK/NAK) under burst or correlated errors using discrete-time analyses, soft error detection procedures or other specific model properties. From an extensive literature review we would like to draw some conclusions which motivate our approach of protocol analysis:

(1) Most analyses of SW protocols aim at state analyses yielding throughput and average delays but do not address the real-time delay performance. In almost all publications Timeout recovery is neglected, i.e., these protocols are subject to deadlock if a frame is lost through buffer blocking or when an ACK or NAK is lost. (2) The analyses of the SR protocol is more complex but shows a superior throughput and delay performance. The analyses are based either on the time for the successful transmission of a frame or by state analysis methods, mostly under idealized assumptions. State analyses are often based on discrete-time Markov Chains and run quickly into a high computational complexity. Studies on the effect of Timeouts and on the real-time performance are neglected in most analytical performance studies.

This paper focuses especially on fundamental link layer, peer-to-peer ARQ protocols as the SW and SR protocols and their aggregated time characterization as basis for the representation of that layer within the next upper layer. In the remaining part of this paper the system models for the SW and the SR protocols are presented in Section 2. In Section 3 we develop an exact analysis by a task graph reduction method and reduce the problem finally to the solution of a standard queuing model of the type GI/G/1 or GI/G/n. Finally, an application example of a NCS was studied with an embedded SW protocol layer in Section 4. Section 5 concludes the paper.

2. MODELING OF PROTOCOLS

There exist quite different modeling techniques which were taylored for specific purposes. In our approach we will make use of elements known from *different* modeling techniques ("hybrid modeling"). We will use: (1) the task graph method which describes coherent sequences of operations by"tasks" known from Operating Systems theory representing the functional model structure, (2) task execution times and frame inter-arrival times modeled by random variables and corresponding stochastic distribution functions, (3) queuing elements for buffering of frames/tokens, (4) logical decisions by token arithmetics and state transitions from the method of Petri Nets, and (5) guiding of frames by decision symbols and their frequency of appearance by probabilities.

A. SW Protocol with ACK/Timeout Control

Timer control is required to avoid a protocol deadlock situation in cases a frame or an ACK is lost. For that case a Timer with value TO is set each time when a frame is transmitted. If the ACK arrives prior to the expiration of the Timer the Timer is stopped and the next frame can be scheduled for transmission. However, if the Timer expires prior to the arrival of the acknowledgement ("Timeout" event) the buffered frame is re-transmitted immediately irrespective whether the frame transmission had been successfully or not. If the ACK arrives after Timeout, the corresponding frame is sent again; the duplicated frame is detected through the frame sequence number and ignored by the Receiver.

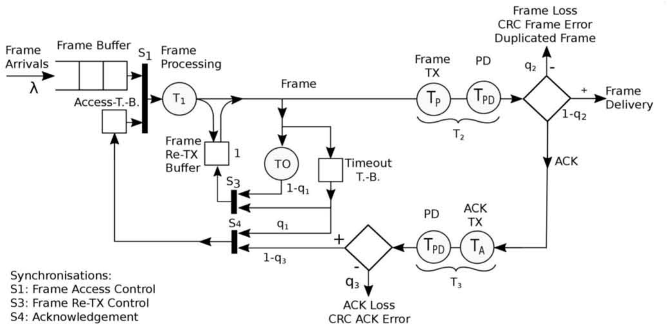

Figure 1. Performance Evaluation Model for the SW Protocol with ACK/TO Control

The detailed performance evaluation model for the SW protocol with ACK/TO control is shown in Figure 1. Arriving frames are buffered and become admitted for frame processing (time T_1) by opening gate S1 with the Access Token. The transmitted frame is buffered in the Frame Re-TX Buffer and the Timer is set with TO. If the frame arrives correctly it is delivered and an ACK is sent back to the sender. If the ACK arrives prior to the Timer expiration the Timer is stopped, the frame copy is deleted and an Access Token is generated. If the ACK is in error or lost the buffered frame is retransmitted upon Timeout. These alternatives are controlled by gates S3 and S4, respectively; the loss or error of a frame (an ACK) transmission and the required frame re-transmission is modeled by probabilities q_2 (or q_3). The case that the ACK arrives prior to Timeout is modeled by probability q_1; in that case the frame copy is deleted and the access token returned to the Access Token Buffer. In the SW protocol only 1 frame is in progress at a time (Send Window $W_S = 1$).

B. SR Protocol with ACK/Timeout Control

The main disadvantage of SW protocols is overcome by increasing the Sending Window to $W_S > 1$. Frames are distinguished by a cyclic frame numbering from 0 to (W_S-1), where the actual number of frames being in progress is maximally W_S to guarantee a definite frame identification. The corresponding protocols fall into the class of "Sliding Window" protocols, where the actual (i.e., open) window size is the difference between W_S and the number of still unacknowledged frames. The sender is allowed to send only as many frames as the open window indicates: With each frame sent out the window size is reduced by 1 and with the reception of an ACK the window size is increased by 1. Within the SR protocol only frames in error are selected for re-transmission. Frames are cyclically numbered by integers running from 0 to 2^m-1. The Receiver uses a window mechanism with a "Receiver Window" $W_R > 1$, i.e., up to W_R frames can be accepted which might have been delayed differently or have arrived out of sequence to reduce the required number of frame re-transmissions at the cost of frame buffering and frame re-ordering before delivery at the receiver site. If the frame numbers are out of a power of 2, i.e., $0,1,\ldots,2^m-1$, the maximum window sizes are limited by $Ws = W_R = W = 2^{m-1}$ to guarantee a definite frame identification [2].

Figure 2 presents the Performance Evaluation Model for our version of an SR protocol with ACK/TO Control and is an extension of the model in Fig.1. The differences are: (1) At the gate S1 up to W frames can be admitted to take part in the transmission processes, controlled by the initial marking of the Access-Token-Buffer with capacity of W Tokens. (2) The "mutual exclusion" control for either, Frame Re-TX or Token Release for a new frame admission is extended to the number of frames which are currently admitted. For each frame a copy is placed in the Frame Re-TX Buffer immediately after

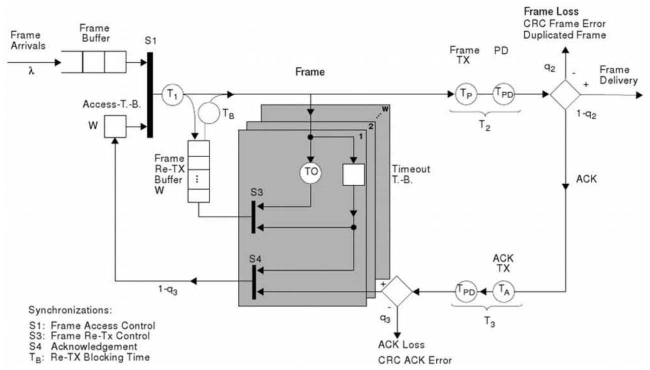

Figure 2. Performance Evaluation Model for the SR Protocol with ACK/TO Protocol

admission; it resides there until its successful transmission is acknowledged. The times T_X of frames being in transit are independent of each other and are identically distributed; this is indicated in Figure 2 by the parallel shaded boxes, one for each activated Frame transfer. *Note:* If a frame is admitted already during an ongoing frame transmission a collision could occur; this could be avoided by a blocking phase T_B as indicated in Fig. 2; in that case successive transit times would be dependent on each other and aggravate the analysis.

C. Modeling Transit Frame Time as Part of a Queuing Model

The transit time T_X of a frame is defined as the time between its admission for transmission until the instant when the Token is returned to the Access Token Buffer. This time can be represented by a task graph in Fig. 3, where the task durations $T_0 = T_P + T_A + 2 \cdot T_{PD}$ and $T_R = \{T_2+T_3 | T_2+T_3 < TO\}$ and where $E[T_P + T_A + 2 \cdot T_{PD}] < TO$ to avoid a protocol life-lock. The task graph of Fig. 3 can be stepwise reduced to one equivalent stochastic phase T_X [4] which serves as service time within a standard queuing system of the type GI/G/1 for the SW and GI/G/W for the SR protocol, c.f. Fig. 4.

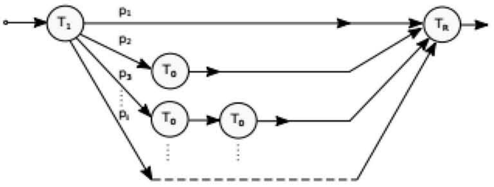

Fig. 3. Task Graph for Frame Transit Time Tx of tSW and SR Protocols with ACK/TO

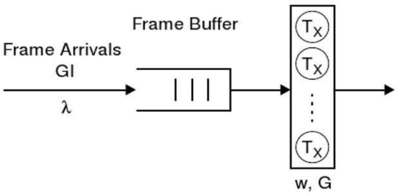

Fig.. 4. Queuing Model GI/G/w

3. PERFORMANCE ANALYSIS

Most protocol analyses in the literature follow a state analysis of the whole protocol system which results in rather complex models which are then analyzed under simplifying assumptions as Markovian traffic processes or discrete-time analysis methods. In this paper we will *separate* the analysis into: (1) The probability distribution for T_X as the *"Virtual Frame Transmission Time"*, and (2) Use of T_X as service time of a standard queuing model of the type GI/G/n as a total system model. By this way we can carry through the analysis exactly without specializing assumptions on stochastic processes based on the system parameters given in Table 1.

Table 1. Input and Derived Parameters

Frame Length	L_P	arbitrary distribution
ACK/NAK Length	L_A	arbitrary distribution
Frame Transmission Time	T_P	PDF $f_P(t)$, LT $F_P(s)$
ACK/NAK Transmission Time	T_A	PDF $f_A(t)$, LT $F_A(s)$
Propagation Delay Time	T_{PD}	PDF $f_{PD}(t)$, LT $F_{PD}(s)$
Frame Processing Times	T_1	(usually constant)
Channel Transmission Rate	R	
Timeout Parameter	TO	
Timeout Recovery Probability	q_1	
CRC Frame Error Probability	q_2	
CRC ACK/NAK Error Probability	q_3	
Bit Error Rate	BER	
Frame Forwarding Delay	$T_2 = T_P + T_{PD}$	
ACK/NAK Feedback Delay	$T_3 = T_A + T_{PD}$	
Virtual Frame Transmission Time	T_X	
Max. Frame Throughput Rate	λ_{max}	
PDF Probability Density Function	LT	Laplace Transformation

A. SW Protocol with ACK/TO Control

We will assume that *no* frame losses occur. At any time either 0 or 1 frame is in the transit phase. A successful cycle happens when a frame and its ACK are correctly transmitted which occurs with probability $(1-q_1)(1-q_2)(1-q_3)$; an unsuccessful cycle happens with its failure probability $q_F = 1 - (1-q1)(1-q_2)(1-q_3)$. Successive cycles happen independently of each other; thus, a successful frame transmission after n cycles occurs with a geometric probability distribution p_n, from which the average number of cycles $E[N]$ is derived:

$$p_n = P\{N=n\} = q_F^{(n-1)}(1-q_F) \qquad n = 1,2,... \tag{1a}$$

$$E\{N\} = 1/(1-q_F). \tag{1b}$$

where $$q_1 = P\{T_2+T_3>TO\} = 1 - \int_{t=0}^{TO} f_2(t) \otimes f_3(t)\, dt \tag{1c}$$

The virtual frame transmission time for exactly *n* frame transmissions is constituted from the frame processing time T_1, *(n-1)* Timeout periods *TO*, and a residual time $T_R = T_x(n)$ for the successful frame transmission shorter than the Timeout period:

$$T_x(n) = T_1 + (n-1)\cdot TO + (T_2 + T_3 | T_2 + T_3 \leq \mathrm{TO}). \tag{2a}$$

In this paper we will assume worst-case conditions for maximum frame sizes, constant propagation delay and frame processing times. Regarding the rules for conditional PDFs the expression for the LT $\Phi_x(s)$ and the PDF $f_X(t)$ of the total frame transmission time T_X: are

$$\begin{aligned}\Phi_x(s) &= \exp(-t_1 s)\cdot\sum_{n=1}^{\infty} p_n \exp(-n[t_P + t_A + 2t_{PD}]s) \\ &= \frac{(1-q_F)\exp(-[t_1 + t_P + t_A + 2t_{PD}]s)}{1-q_F\exp(-[t_P + t_A + 2t_{PD}]s)} \\ &= \frac{(1-q_F)\exp(-[t_1 + t_0]s)}{1-q_F\exp(-[t_0]s)}\end{aligned} \tag{2b}$$

$$f_x(t) = (1-q_F)\sum_{i=1}^{\infty} q_F^{i-1}\,\delta(t-[t_1 + it_0]) \tag{2c}$$

From (2c) the ordinary moments of T_x follow as

$$E[T_X] = t_1 + t_0 + \frac{\mathrm{TO}}{1-q_F}\cdot q_F \tag{2d}$$

$$E[T_X^n] = (1-q_F)\cdot\sum_{i=1}^{\infty} q_F^{i-1}\cdot[t_1 + t_0 + (i-1)\mathrm{TO}] \tag{2e}$$

The maximum frame throughput rate λ_{max} follows as reciprocal from the average frame transfer time as

$$\lambda_{max} = 1/E[T_X] = 1/t_X \tag*{(2d)}$$

With the results for T_X we can find all other performance values in principle from the queuing system GI/G/1. Exact results are known for the cases M/G/1 and GI/M/n [3,6]. The characteristic performance values for the M/G/1 delay system with negative-exponentially distributed frame arrivals (M) for the probability of waiting *W*, the first and second ordinary moment of waiting times E[$T_W{}^i$], i = 1,2, and the squared coefficients of waiting times of arriving frames c^2_w and delayed arrivals c^2_D are (*Note:* in (3c,d) T_H must be replaced by T_X):

$$W = \mathrm{P}\{T_W > 0\} = \rho = \lambda\cdot\mathrm{E}[T_X] \tag{3a}$$

$$E[T_W] = \frac{\lambda E[T_H^2]}{2(1-\rho)} \quad , \qquad E[T_W^2] = \frac{\lambda^2 E[T_H^2]^2}{2(1-\rho)^2} + \frac{\lambda E[T_H^3]}{3(1-\rho)} \tag*{(3b,c)}$$

$$c_w^2 = E[T_w^2]/E[T_w]^2 - 1 \ , \quad c_D^2 = W(c_w^2 + 1) - 1 \tag*{(3d,e)}$$

The waiting time DF can be well approximated by a Weibull DF from which the waiting time percentile follows

$$p = [1 - W(t_{Th})] \tag*{(3f}$$

Note: We prefer the waiting time characteristics for the *delayed* arrivals $T_D = \{T_W | T_W > 0\}$ as it has a smaller coefficient of variation and T_D is the better metric for SLA indicating the individually experienced delay. The Weibull DF depends only on the first moment and on the coefficient of variation *c* with $0 \leq c \leq \infty$ and is tabled in [5]. As the DF of delays in the M/G/1 queuing model are rather complex we will approximate the DF of delayed frames by the coefficient of variation c_D using the Weibull DF from which we can find the threshold of delay t_{Th} which is exceeded with percentile *p*.

For purposes of an equivalently aggregated model for the link layer protocol, represented by an M/G/1 queue, into the *next higher layer* we need further quantities of the M/G/1 queue: the Flow (or Sojourn) Time T_F and the output process, characterized by the inter-departure time T_O . In a queuing model M/G/1 the flow time T_F is the sum of the two independent random variables waiting time T_W and holding (service) time T_H: $T_F = T_W + T_H$, i.e., the PDF of T_F follows as convolution of the PDFs of T_W and T_H resulting in $\Phi_F(s) = \Phi_W(s)\cdot\Phi_H(s)$, from which we conjecture that

$$E[T_F] = E[T_W] + E[T_H], VAR[T_F] = VAR[T_W] + VAR[T_H] \tag{3g}$$

The output process is only known for the models M/M/1 and M/D/1, but for the queuing model M/G/1 the second moment of T_O is exactly known [8], represented by the squared coefficient of variation

$$c_O^2 = c_A^2 + \rho^2(c_H^2 - c_A^2) \tag{3h}$$

B. SR Protocol with ACK/TO Control

With the identical DF for T_X in case of the SR protocol of Section 2B the analysis of this protocol can be performed by the queuing model GI/G/n, where n indicates the window size of the SR protocol . This model is more difficult to analyze but we know excellent approximate solutions for the M/G/n queuing model from [7]

$$W = P(T_W > 0) = W|_{M/M/n} \tag{4a}$$

$$E[T_W] = \frac{1 + c_H^2}{2} E[T_W]|_{M/M/n} \tag{4b}$$

$$c_D^2 = 2\rho - 1 + \frac{4(1-\rho)E[T_H^3]}{3(c_H^2 + 1)^2 E[T_H]^3} \tag{4c}$$

C. Verification of Analytical Results by Simulation

Under the indicated assumptions our analytical results are exact, except for the full DF of the waiting times for which we choose the Weibull DF fitted with the exact first and second moments. Numerical results have been tested for the two basic Link-Layer protocols SW and SR with ACK/TO Control which showed an excellent accuracy and will be presented for some representative example cases for the following example parameters: $t_1 = 0.1$ ms, $t_0 = 1.0$ ms, $q_F = 0.1$, $\lambda/n = 0.1,..,0.75$ 1/ms $n = 1$ (SW), $n = 4$ (SR).

Table 2. Comparison of Analytic Results (Regular Font) with Simulations (Italic Font)

λ /n 1/ms	0.10	0.30	0.50	0.70	0.75	0.10	0.30	0.50	0.70	0.75
ρ	0.121	0.363	0.605	0.848	0.903	0.127	0.380	0.633	0.887	0.950
Protocol	**SW with ACK/TO Control**					**SR with ACK/TO Control**				
$E[T_T]$ ms	1.267	1.267	1.267	1.267	1.267	1.267	1.267	1.267	1.267	1.267
	1.267	*1.267*	*1.267*	*1.267*	*1.267*	*1.267*	*1.267*	*1.267*	*1.267*	*1.267*
$E[T_w]$ ms	0.108	0.455	1.283	5.812	14.12	0.000	0.023	0.168	1.247	3.312
	0.108	*0.457*	*1.290*	*5.856*	*15.09*	*0.000*	*0.026*	*0.177*	*1.267*	*3.337*
c_W	3.581	1.948	1.412	1.104	1.044	30.39	4.741	2.173	1.259	1.108
	3.592	*1.951*	*1.413*	*1.108*	*1.046*	*29.73*	*4.664*	*2.142*	*1.249*	*1.103*
$E[T_D]$ ms	0.851	1.198	2.026	6.556	14.86	0.231	0.300	0.507	1.639	3.715
	0.853	*1.210*	*2.043*	*6.660*	*14.35*	*0.274*	*0.352*	*0.550*	*1.684*	*3.761*
c_D	0.867	0.907	0.946	0.984	0.993	0.867	0.907	0.946	0.984	0,993
	0.872	*0.911*	*0.949*	*0.988*	*0.995*	*0.812*	*0.832*	*0.880*	*0.963*	*0.983*
$E[T_F]$ ms	0.894	1.241	2.069	6.599	14.90	0.787	0.809	0.954	2.033	4.098
	0.894	*1.244*	*2.076*	*6,642*	*15.13*	*0.787*	*0.813*	*0.964*	*2.054*	*4.124*
c_F	0.574	0.722	0.842	0.944	0.967	0.670	0.665	0.672	0.815	0.905
	0.575	*0.724*	*0.844*	*0.916*	*0.971*	*0.670*	*0.665*	*0.670*	*0.812*	*0.902*

D. Determination of Threshold Percentiles

For real-time applications guarantees have to be given that a certain delay threshold t_{Th} is exceeded only with a prescribed small probability ("percentile") p. We will approximate the complementary DF of delays of delayed arrivals by the first and second moment only and apply for that case the Weibull DF whose two parameters can be derived from the two moments. The numerical determination requires, however, an iterative procedure where the Gamma Function is involved. In the Delay Tables [5] the complementary DF of delayed arrivals is plotted for a wide range of coefficients of variation c_D from which the threshold values t_{Th} can be determined. In Table 3 the threshold values are given for three percentiles 0.05, 0.02 and 0.01 and five different load cases.

4. MULTI-LAYER PROTOCOL AGGREGATION EXAMPLE

We consider now a Networked Control System (NCS) application layer example where an Actuator A (Plant) is interconnected with the Controller C through a Network (bidirectional channel) operated by the SW protocol with ACK/TO Control as presented in Section C, c.f. Fig.5. The complete function of the Network N is aggregated by the DF of the Flow time T_F in either direction represented by its Laplace Transform (LT) $N(s) = \Phi_F(s)$, respectively. The accuracy of our method of aggregating the whole link layer into a stochastic delay phase T_F representing the Network module within Fig. 5 has been verified successfully [9].

Table 3. Delay Thresholds t_{Th} for Three Different Percentiles p

λ 1/ms	0.10	0.30	0.50	0.70	0.75
t_{Th} ms, p = 0.05	2.1	3.9	5.9	19.7	42.2
t_{Th} ms, p = 0.02	2.4	4.2	8.9	24.7	57.8
t_{Th} ms, p = 0.01	2.8	5.0	9.2	29.8	67.8

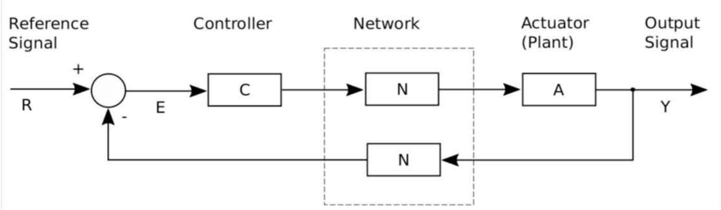

Fig. 5. Model of a NCS Interconnected by a Network N in both directions

5. CONCLUSIONS

A novel approach for an integrated analysis of communication protocols has been presented. By this method exact and closed-form solutions can be derived under quite general process assumptions and have been applied to two link-layer protocols.

REFERENCES

[1] A. Gupta, M.-Y. Chow, "Networked Control System: Overview and Research Trends ", IEEE Trans..Ind. Electronics, Vol. 57, No.7, July 2010, pp. 2527-2535.

[2] A. Leon-Garcia, I. Widjaja, "CommunicationNetworks - Fundamental Concepts and Key Architectures", 2nd Edition, McGraw-Hill.

[3] H. Kobayashi, B.L. Mark, "System Modeling and Analysis - Foundations of System Performance Modeling", Pearson Education Inc., Upper Saddle River, N.J., 2009.

[4] P. J.Kuehn, "Performance and Energy Efficiency of Parallel Processing in Data Center Environments", Springer Lecture Notes LN 8945,2015, pp.17 -33.

[5] P. Kuehn, "Tables on Delay Systems",I nstitute of Switching and DataTechnics, Unversity of Stuttgart, 1976.

[6] B.W. Gnedenko, D. König, "Handbuch der Bedienungstheorie II" (in German), Akademie-Verlag Berlin,1984.

[7] W. Whitt, "Approximations for the GI/G/m Queue", J. Production and Operations Management, Vol. 2, No. 2, Spring 1993,pp.114 -161.

[8] T. Makino, "On a Study of Output Distributions", J. of thge Opns. Res. Soc.of Japan 8 (1966), 1968, pp. 109 - 133.

[9] P.J. Kuehn, S. Scholz, S. Cao, F. Li, "Performance Modeling of Networked Control Systems", Paper accepted for:the 9th Int. Congress on Ultra Modern Telecommunications and Control Systems (ICUMPT), 2017, Nov. 6 - 8, 2017.