Congestion resolution and buffering in packet switched all-optical networks

Summary of Ph.D. thesis

Miklós Kozlovszky

Supervisor:
Prof. Tibor Berceli

Department of Broadband Infocommunication and Electromagnetic Theory
Budapest University of Technology and Economics
BUTE

Budapest, 2008

Introduction

Optical communication spectacularly conquers data networks. Electrical signals used in transmission gradually roll back, and optical signal based communication extends rapidly. The so-called "green networks" and the "green computing" initiatives are important new driving forces for all-optical switch development, where these new optical technologies are trying to save cost and massive amounts of power with "green optical links" in networks and in supercomputers using light instead of wires to transmit information. In case of optical networks, the main key points of packet switching and routing are how to read and process the routing information from the packet, how to delay the packets in optical domain and how to resolve contention happening in the network.
For contention resolution in optical switches, the real challenges are the speed efficiency and the optical signal buffering. Contention occurs, when the packets are being switched, and two or more packets are trying to leave the switch on the same wavelength of one port at the same time.
Contrary to electronical buffers; in the optical domain only limited solutions are available to store packets. The lack of optical RAMs is the key problem here, however there are some solutions (e.g.: slow light or signal buffering in optical delay lines) to solve this issue. In case of optical delay lines, besides the effective scheduling, also the size and attenuation of the complex buffer structure are important optimization parameters. This dissertation addresses the sizing problem of contention resolution and high-speed optimal packet switching in all-optical networks. It gives formal description (through discrete event simulation based models and analytical models) of optical buffering and proposes solutions for network elements communicating with light signals.
In the theme of Optical Packet Switching (OPS) - sometimes referred to as Photonic Packet Switching (PPS) – most of the published results are theoretical, and there are only a few of them validated. Validated results are focusing on Optical Burst Switching (OBS) in many cases, because this solution is more feasible on our technology level. Analytical models are using large number of simplification (e.g.: zero time TWCs, etc.) to model optical switching systems. The published algorithms are validated individually, and such separated analysis can not reveal and capture the underlying dependencies between the algorithms and the various systems elements. Modular and open model solutions are needed to research such complex and structured systems and predict their performance or assess their effectiveness. To give and build up such models, I have done an assessment of the feasible delay models, and I have implemented a discrete event based simulation model, which is able to help to assist research of all-optical switches with its large feature set and configuration parameter field. As most important literature resources to my research work, I have used works from Callegati [1,2], for the analytical models from: Telek [3,4], and German [6,7]. To create the Non-Synchronized Optical Switch Simulator (NSOSS) I have used the OMNeT++ [8,9] discrete event framework developed by A. Varga from BUTE. The numerical calculations and analytical models were implemented in Mathematica (Wolfram Inc.) and in SPNica [5]. Microsoft Excel (Microsoft Corp.) was used to do smaller calculations and most of the diagrams in the dissertation.
The basis and premise of my dissertation was the long history of research work in optical networks, which is leaded by Prof. Tibor Berceli at the Department of Broadband Infocommunication and Electromagnetic Theory of Budapest University of Technology and Economics, more precisely in the Optical and Microwave Telecommunication Laboratory, which developed numerous methods, solutions and equipments for Hungarian and international optical packet switching projects [S3, S6, S7, S9-S11, S21, S24].

Main objectives

Packet-switched services available today are using electronic switches; however this dissertation focuses on next generation networks, where the packet-switching is performed optically. Optical Packet Switching (OPS) is now a heavily researched hot topic in the field of optical communication, concentrating mostly on all-optical hardware, contention resolution, buffering, switching, protocols and network control. The main reason for researching the field of Optical Packet Switching was for me the vast of favorable property that optical packet switches can potentially provide in the near future. Optical Packet Switching technology can bring theoretically two orders of magnitude higher capacities, than electronic ones in the communication world. It will have improved utilization compared to circuit-switched or burst switched optical networks. Interesting driving forces are also the so-called "green networks" and the "green computing" initiatives, where new optical technologies are trying to save massive amounts of power with "green optical links" in networks and in supercomputers using light instead of wires to transmit information. The main objective of my research was to create an optical packet switch model, which gives accurate information about its operational parameters when the configuration (type of the used buffer structures, number of output/input ports, type of used routing or scheduling algorithm, etc.) of the model is changed.

My main target was to further develop the theoretical background of the optical packet switch models, enhance the sizing process of such equipments, and confirm the existence and usability of accurate all-optical packet switch models.

Further target of mine was to explore how the basic buffer structures can influence the overall performance of the switch, and gain information how the throughput and performance can be increased by changing certain configuration parameters (such as usage of TWCs, dedicated or shared buffering, etc.)

Due to the size and complexity of the problem, I have identified the following minor targets, to reach step-by-step my final goal:

- Exploration of the generic and focused literature about optical packet switching, congestion resolution and buffering in all-optical networks.
- Definition of the research field to separate later my research results from the existing scientific works. Classification of the published algorithms and architectures and analysis of the not published ones.
- Development of my own algorithms and solutions.
- Comparison of my algorithms and architectures to the already published ones, and selection of the optimal combinations or further optimization in some cases.
- Exploration of the generic and focused literature about analytical model building
- Development of the analytical models.
- Working out performance measurements to analyze the system, and simulations with discrete event simulator.
- Processing of the performance results, and distillation of the results into closed formulas and equations.
- Comparison of the analytical models with the discrete event simulation models, and validation of the performance results.
- Analysis of the system results, evaluation of the scope and sensibility of the results.

Methodology

As first step I have assessed the literature, and it turned out, that in optical packet switching the congestion resolution (due to the novelty of the topic) has very limited literature compared to optical

burst or circuit switching (I should note here, that in the last two years this tendency is changing rapidly). The main reason for that is the high price of the hardware elements, and the complexity of the problem, however the market leader equipment vendors and also the constantly increasing number of publications are showing clearly the actuality of this research theme.
I have chosen two separate parallel approaches to explore the sizing of all-optical networks (and switches), and their optimization problem during my research work; the discrete event simulation models, and the numerical calculation based on analytical models.
For the simulations I have defined and implemented a *Non-Synchronized Optical Switch Simulator (*NSOSS) [S25]. Besides the implementation of the simulator, for the analytical models I have developed active and passive type buffer structure models with stochastic Petri nets. I have compared the results coming from the two approaches to assess the accuracy of the models.
Beside the basic buffer models, I have also created and implemented scheduling algorithms, which I have further optimized during my research work. I have used for my research a large set of applications and utility tools, such as: OMNeT++ and ns-2 as basic frameworks for discrete event simulations, Matlab/Simulink, SPNica and LabView as generic application frameworks for numerical calculation and analysis. In general, most of the literature of this research field was available only in English; however some Hungarian and German publications were also helpful for me.

Used methods

- Discrete event simulations to model the system
- Stochastic Petri nets to define the analytical models and do numerical calculation
- Statistical result data analysis
- Validation of simulations and model results

New scientific results

THESIS 1.

Thesis 1. defines the implemented asynchronous optical switch model based on discrete event simulation, which is doing multi-level congestion resolution by the help of the in Thesis 2.1. defined delay buffers, deflection routing strategies, and tunable wavelength converters. The basic buffer structures and their scheduling algorithms have been validated with analytical models (specified in Thesis Group 4.). Relevant sections in the dissertation: 3/2, 3/3, 3/4. Related publications of the thesis are located in [S3, S6, S7, S9, S10, S11, S21, S24].

THESIS GROUP 2.

Congestion resolutions and avoidance in all optical packet switches

I have systematized the already published basic optical delay buffer structure models such as constant size (constant delay time) sequentially connected, (C-TOBS), degenerative parallel connected (D-POBS), recirculating sequentially connected (R-TOBS), and recirculating parallel connected (R-POBS) optical buffer structures (shown on Figure 1.). I have analyzed these basic models, and defined scheduling algorithms and closed formulas for their operational parameters.

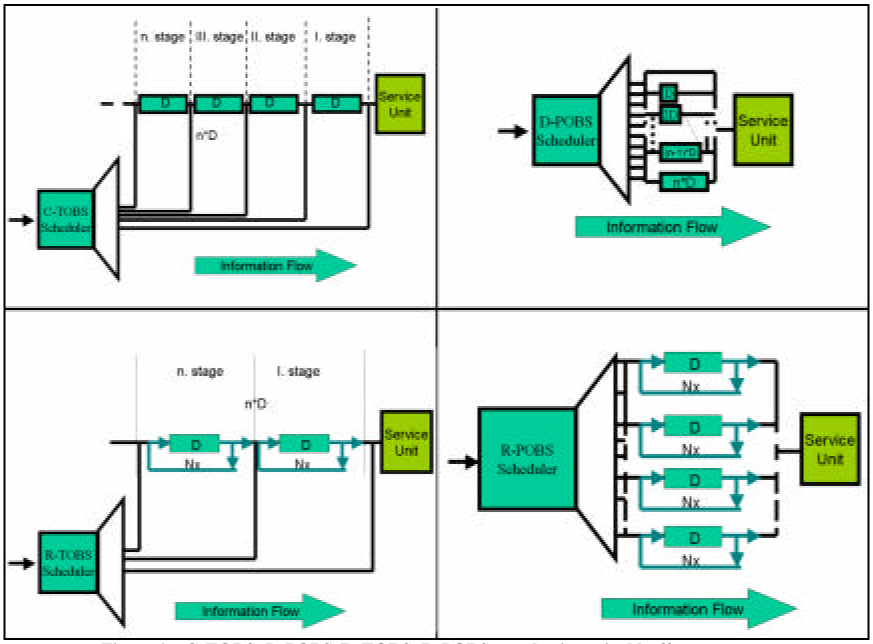

Figure 1. : C-TOBS, D-POBS, R-TOBS, R-POBS type basic optical buffer structures

Thesis 2.1.

It defines the optimized algorithms for the various buffer structures and TWCs. Thesis 2.1. **describes the following new algorithms for multi-level (wavelength and port level) congestion resolution:**

- **Scheduling algorithms for C-TOBS, D-POBS, R-xOBS structures**
- **Scheduling algorithm for Tunable Wavelength Converter (TWC)**
- **Optimized deflection routing algorithm**

The model in Thesis 1. contains the implementation of these algorithms. Relevant sections in the dissertation: 3/2, 3/3, 3/7.

Thesis 2.2.

I have investigated the characteristic of the C-TOBS, D-POBS, R-TOBS and R-POBS optical buffer structures. I have analyzed the relationship between the type of buffer structure, the number of buffers in the buffer structure and the performance of the structure. I have shown, that even a minimal amount of buffer outperforms the bufferless configuration (shown on Figure 2.).

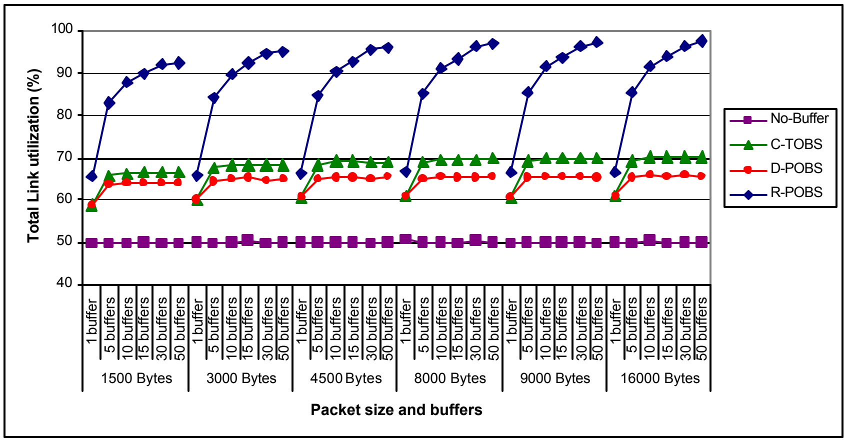

Figure 2. : Buffer performance results, changing parameters: packet size and number of buffers in the structure

I have also shown that the performance parameter of an all-optical packet switch equipped with C-TOBS, D-POBS, R-TOBS or R-POBS delay buffer structures can be defined by a closed formula.

Thesis 2.2 **gives closed formulas** (in case of 1500 Bytes constant packet size and optimized buffer size)**, how the performance characteristics of the C-TOBS, D-POBS, R-TOBS and R-POBS type buffer structures are changing, if the number of buffers is increased in the structure (Equations 2. - 5.) and gives also the root-mean-square deviation** (R.M.S.D) -also known as root mean square error- **of each formulas**. The R.M.S.D. is defined as:

$$RMSD = \sqrt{MSE(\Theta_1, \Theta_2)} = \sqrt{\frac{\sum_{i=1}^{n} [y_i - y(x_i)]^2}{n}} \quad (1)$$

, where MSE is a mean square error of the two kind of the simulated results and predicted values.

C-TOBS

$$P_{C-TOBS} = \frac{-4.7x \log[x] + 682.5x - 95.6}{10x} \quad (2)$$

R.M.S.D. < 0.211

D-POBS

$$P_{D-POBS} = \frac{-2.38x \log[x] + 651.5x - 64.8}{10x} \quad (3)$$

R.M.S.D. < 0.207

R-TOBS

$$P_{R-TOBS} = \frac{55.1x \log[x] - 3.7x^2 + 759.7x - 107}{10x} \quad (4)$$

R.M.S.D. < 1.108

R-POBS

$$P_{R-POBS} = \frac{62x\log[x] - x^2 + 748x - 96}{10x} \quad (5)$$

R.M.S.D. < 0.61

, where *P* is the performance (in % of the full utilization capacity) of a single wavelength and variable *x* means the amount of buffers in the buffer structure. Relevant section in the dissertation: 3/8.3.1.

Thesis 2.3.

I have investigated in the connection between C-TOBS, D-POBS, R-TOBS and R-POBS optical buffer structures, the number of buffers in the structure, and the utilization capacity. Thesis 2.3. **gives closed formula for various constant packet sizes (range 3000-16000 Bytes), the type of the buffer structure (C-TOBS, R-TOBS, D-POBS, R-POBS) and the performance /in % of the full utilization capacity/ (Equations 26.):**

$$P = \log\left[\frac{(L_K - L_D)}{L_D}\Delta P_{C/D/R-xOBS}\right] + P_{C/D/R-xOBS} \quad (6)$$

, where L_K is the new packet size (ranging from 3000 Bytes to 16.000 Bytes), L_D is the default packet size (1500 Bytes), $P_{C/D/RxOBS}$ is the performance equation of the buffer architecture at 1500 Bytes, and *?* $P_{C/D/RxOBS}$ is the empirical constant for each buffer structure (shown in the Table 1). These empirical constants were gained trough the regression analysis of the simulation results. Relevant section in the dissertation: 3/8.3.2.

Buffer structure type	?$P_{C/D/R-xOBS}$	Root Mean Square (R.M.S.D)
C-TOBS	$6.29654\ e^{\frac{-0.74}{x}}\ x^{0.1}$	R.M.S.D. < 0.963
D-POBS	$\frac{2.58571\, e^{\frac{0.18}{x}}}{x^{0.047}}$	R.M.S.D. < 0.963
R-TOBS	$0.139457\ e^{\frac{2.55}{x}+0.07x}\ x^{1.59}$	R.M.S.D. < 1.671
R-POBS	$e^{\frac{1}{5}+\frac{0.41}{x}}\ x^{0.65}$	R.M.S.D. < 1.229

Table 1.: ?$P_{C/D/R-xOBS}$ values of the various buffer structures

Thesis 2.4.

I have analyzed from various aspects the characteristics of the buffer structure and wavelength converter based congestion resolution. I have realized both single and multi wavelength simulations. In the multi wavelength simulations, I have used TWCs with zero (hypothetical model) and non-zero tuning time. I have investigated into the following system parameters:

Signal level changes in case of various buffer structures, and amount of buffers in the structure

I have analyzed how the type of delay structures is influencing the signal attenuation of the packets. It turned out, that the signal attenuation is influenced by the type of the buffer structures and the number of used buffers.

Thesis 2.4. states that **the optical packet travelling through active buffers experiences more attenuation, than through passive buffers**.

The x-POBS (parallel connected) **type buffer structures** (where x can stand for D or R) **have better control on the signal attenuation, than the x-TOBS** (sequentially connected) **type structures** (where x can stand for C or R).

The R-POBS type buffers are producing (recirculating and parallel connected) **better attenuation values then the R-TOBS type buffers, and the attenuation of R-POBS type buffers is upper bounded** (in case of identical working conditions).

Utilization in case of various buffer delay time and buffer type

I have investigated how the type of the buffers and their delay time influences the total utilization on the wavelength.

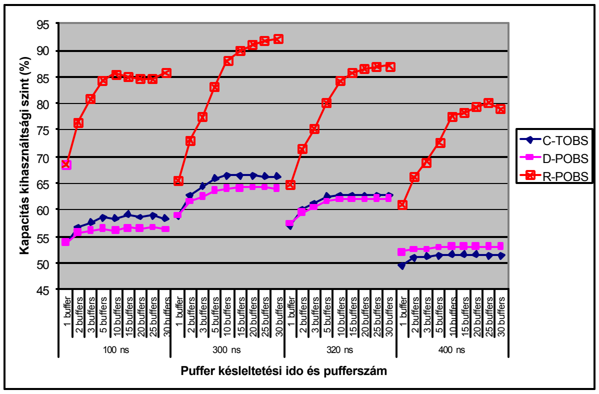

Figure 3.: Buffer delay time vs. performance /packet size=1500 Bytes (300 ns at 40Gbps)/

Thesis 2.4. states that (in case of constant packet size and buffer delay time) **the performance degradation is not symmetric, and deviation from optimum towards smaller delay gives better link utilization performance, than if the delay time is longer then the optimum.**

In this scenario (shown on Figure 3.) smaller buffers (delay time is shorter with 100 *ns*) are performing significantly better than bigger (delay time is longer with 100 *ns*) ones. This statement holds both for parallel and sequentially connected buffer structures.

Optimal number of Tunable wavelength Converters (zero and non-zero tuning time)

I have investigated how Tunable Wavelength Converters (TWCs) with various tuning time parameters can influence the total utilization of the port in case of different amount of buffers in the buffer structures. Thesis 2.4. shows, that **the performance enhancement due to the usage of a TWC is decreasing if the number of buffers on the wavelengths is increasing** and this effect is proportional to the tuning time of the TWC.

Optimal number of non-zero-time Tunable wavelength Converters

In real case scenario TWCs have non-zero tuning times, which are comparable to the system speed and to packet service time. I have investigated the optimal number of TWCs. Buffer sizes in the delay structure were homogenous and optimized for 1500 Bytes. The TWCs have constant tuning times, independently from the actual wavelength. Tuning time difference between two wavelengths was not influenced by any other parameters. Thesis 2.4. **gives a range of the optimal number of shared Tunable Wavelength Converters per port for the optimal utilization of the port (Equation 7.).**

$$N_{TWC} = \left[\frac{N_W}{3};\frac{N_W}{2}\right] \quad (7)$$

where N_{TWC} is the optimal number of TWCs used for the wavelength conversion, N_w is the number of wavelengths on the port. The equation states that one can use minimum the ⅓, maximum the ½ number of wavelengths as an optimal number of required TWCs. The range depends mainly on the number of wavelengths per port and if the tuning time of the TWC is less than the packet processing time (in case of average packet size=1500 Bytes → 300 ns at 40Gbps), the tuning time does not influence the equation. Available TWCs have ~200 ns average tuning time, which means that the equation is appropriate for sizing such equipments. The range in the equation is large enough also for zero tuning time TWCs. Outside the optimal range, additional TWC negatively influences the overall system performance. For this range it is assumed that the tuning times of all the TWCs are constant (200 ns). Relevant sections in the dissertation: 3/8.1.3, 3/8.1.4, 3/8.2. and 3/8.3.5.
Related publications of the thesis group are located in [S3, S6, S7, S9, S10, S11, S21, S24].

THESIS GROUP 3.

End-to-end communication performance

I have investigated into the network communication performance characteristics of the second and third layer of the OSI reference model if the traffic is flowing through all-optical packet switches. Different utilization levels have been used for measuring systematically how the parameters of the physical layer influence the performance of higher layer communication protocols (such as UDP, TCP, SCTP) and the average packet delay.

Thesis 3.1.

I have investigated how the Packet Delay Time depends on three key factors; namely on the buffer structure type, on the amount of buffers in the structure and on the actual utilization of the switch, later on referred to as System Load (in case of end-to-end /point-to-point/ UDP traffic). Ethernet (100Base-X) with 100 Mbps maximum end-to-end communication speed was assumed. From the results it turned out, that in case of UDP based point-to-point communication (packet size was constant 1500 Bytes, and puffers were optimized to this size), Parallel Optical Buffer structures (such as R-POBS and D-POBS) keep the average Packet Delay lower than Tandem Optical Buffer Structures, even at high system load. R-POBS type buffer structure was chosen as a reference model structure because this is the most interesting buffer structure from performance and effectiveness point of view. The overall performance of such active buffer structure is greater than any other investigated buffer structure. Thesis 3.1. **gives closed formula (in case of UDP type end-to-end traffic) about the relation between the factors: Average Packet Delay (APD) on an individual wavelength in an output buffered optical switch, and System Load (link utilization) of the switch, which has R-POBS type buffer structures with a certain amount (*x*) of buffers. (Equation 8.).**

$$APD(x) = 6.155*10^{-4} + \mathrm{A}^2\left(1.4*10^{-10} - 1.6*10^{-8}e^{-x}\right) + 1.1*10^{-13}A^3x \quad (8)$$

where *APD* is the Average Packet Delay (in sec)**, *x* is the number of buffers in the Buffer Structure** (it is a dimensionless number)**, and *A* is the utilization of the switch or System Load** (in % of the full utilization capacity). Recommended usage range of the equation (due to accuracy) is 1 - 20 buffers, however this is not a real limitation. Nowadays the number of used buffers is usually less than 5 in buffer structures. The equation has $MRS_{RPOBS-APD} < 2.326030 \cdot 10^{-5}$ in the full System Load range. Relevant section in the dissertation: 4/2.3.1.

Thesis 3.2.

I have investigated into TCP based end-to-end communication performance, and have compared it with the UDP based results. The maximum possible utilization between the endpoints was 100 Mbps. I have analyzed how the end-to-end communication performance is changing in case of various System Load (link utilization) of the all-optical switch (pre-defined utilization levels:10%, 50%, 80%, 90%), different types of buffer structures and the number of buffers in the buffer structure.
Thesis 3.2 **states that in case of TCP based end-to-end communication:**

- **The recirculating, active type (R-TOBS and R-POBS) delay structures have better communication performance, than the passive structures.**
- **If the utilization of the link is less than 50% additional buffers in the structure can compensate the performance degradation.**

Furthermore I have confirmed the already known statement that the TCP based end-to-end communication performance depends more on the utilization parameter of the link, than the UDP based end-to-end communication.
Thesis 3.2 **states that in case of UDP based end-to-end communication small amount of buffers in the buffer structure can increase significantly the end-to-end communication performance (up to 70% of the total throughput can be achieved), even in case of high link utilization.**
Thesis 3.2 **states that in case of TCP based end-to-end communication, only large number of buffers can increase the end-to-end communication performance, however the achieved throughput is always smaller than UDP based end-to-end communication performance** (in case of identical traffic conditions). This means for TCP based communication, optical switches should work with deeper buffer structures compared to UDP based communication, however even with deeper buffer structure (many additional buffers) the full point-to-point communication performance cannot be achieved under heavy load due TCP's flow and congestion control mechanism. This drives to the conclusion, that even bufferless congestion resolution can give similar performance (this is valid strictly only for TCP traffic) as buffered congestion resolution if the network traffic is high. Relevant sections in the dissertation: 4/2.3.2. and 4/2.3.3.

Thesis 3.3.

I have investigated into TCP and SCTP end-to-end communication performance. From the initial measurements it turned that SCTP is very stack dependent. The BSD protocol stack was much more tuned for SCTP than any other operating system and SCTP was able to outperform TCP on BSD in many scenarios. However the majority of HPC clusters are using Linux as base operating system, so it was obvious, to choose Linux OS as measurement platform (software and hardware configuration of the testbed is detailed in the dissertation). I have analyzed and gave performance results (results gained from cluster performance measurements) of single and multi-rail measurements within cluster environment over Fast Ethernet and Gigabit Ethernet.
Thesis 3.3 states, that

- **The out-of-the-box single stream performance of SCTP is comparable with TCP on Fast Ethernet links** (1500-2000 Bytes as payload)**.** TCP is better performing than SCTP (less than 3% difference has been measured between the two protocols /SCTP is slower/ on Fast Ethernet).
- **The slower handshake of SCTP does not influence significantly the connection establishment time in cluster environment, even in case of large number of connections.**

Due to the additional leg in the SCTP connection establishment (it uses 4-way handshake during connection setup), each SCTP connection experiences average $^1/_2$ RTT delay during connection establishment. The added leg provides SCTP associations with better defense against DoS attacks.

- **In case of SCTP based traffic one-to-one mode using the same applications and function calls, the sender side's CPU is working approximately twice more intensive thanin case of TCP based traffic.**

Relevant sections in the dissertation: 4/3. and Appendix C.
Related publications of the thesis group are located in [S1-S9, S15, S18, S20, S22-S24].

THESIS GROUP 4.

Delay buffer structure analytical models

During my analytical modeling work, I have defined and described various types of buffer structures with Petri net. My main target was with the analytical models to verify the basic buffer structures of the discrete event simulation based models. I have defined two types of bufferless models. In the first model the system handles variable size packets and the service time is non-deterministic. The second model assumes constant packet size and (as a consequence) deterministic service time. Both models are working with exponential packet interarrival time. Furthermore I have created the stochastic Petri net based analytical models of the buffer structures defined in the second thesis group. I have compared the analytical models with the discrete event simulation based models (from Thesis 1.), and compared and validated the models and the two model building methodologies with a common performance parameter. The results of the comparison are detailed in the dissertation 5./7. section, where I draw the conclusion, that the developed analytical models and the developed discrete event simulation models are comparable, and analytical models validating successfully the basic discrete event models. As comparison parameters, the average effective service rate of the analytical models and the average performance of the service units were compared. Root Mean Square (RMS) was calculated for the results to evaluate the numerical difference of the models. The difference between the performance results coming from the two modeling approaches was analyzed. The RMS values are lower in case of the paired constant packet size discrete even simulation models and analytical models, than with variable sized packet models. With constant packet sizes the RMS values are below 2.5 and with variable packet sizes the RMS values are below 3.1 for the whole performance result set. This difference shows convincing accuracy if we recall the initial limiting assumptions of the analytical models (e.g.: guard time elimination, etc.). The two independent model building methodologies provide comparable results regarding the performance of the basic buffer structures, however analytical models have strong limitations due to the state space explosion if the modeled system becomes complex. The solvable problem size with analytical models is more limited, than with discrete event simulations. Problems, which are successfully modeled and analyzed with discrete event simulations in the second and in the third thesis groups (buffer structures with 30-50 buffers, multi-wavelength simulations, etc.) are out of scope of the buffer structure scale analytical models. However the developed analytical models were able to validate successfully the basic discrete event simulation based models.

Thesis 4.1.

Thesis 4.1. defines the basic analytical model of the passive delay buffer structure with help of stochastic Petri net. The basic passive optical buffer structure model contains only a service unit and a single passive buffer, however it can be a basis model of both two passive structures (constant length buffer and degenerated type buffer) integrated in Tandem (TOBS) or Parallel (POBS) Optical Buffer

Structure. The models were created both with constant and variable sized packets (Figure 4.). Relevant section in the dissertation: 5/6.1.

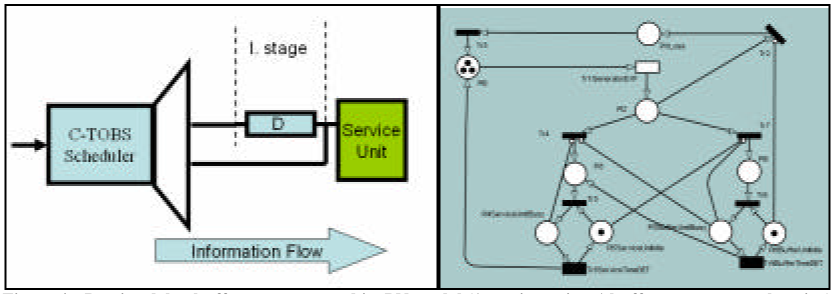

Figure 4.: Passive delay buffer structure and its PN model (1 service unit + 1 buffer, constant packet size)

Thesis 4.2.

Thesis 4.2. defines the basic analytical model of the active delay buffer structure with help of stochastic Petri net. The basic active optical buffer structure model contains only a service unit and a single active buffer; however it can be a basis model of both an active structure integrated in Tandem (TOBS) or Parallel (POBS) Optical Buffer Structure. The models were created both with constant and variable sized packets (Figure 5.). Relevant section in the dissertation: 5/6.2.

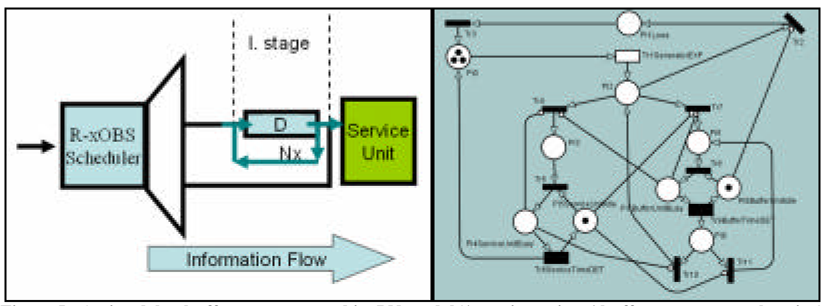

Figure 5.: Active delay buffer structure and its PN model (1 service unit + 1 buffer, constant packet size)

Thesis 4.3.

Thesis 4.3. **defines the analytical models of the active and passive multi-buffer delay structures with help of stochastic Petri nets.** These active and passive parallel and tandem multi-buffer structures are previously defined in the second thesis group. In all the analytical models homogeneous TOBS and POBS type buffer structures have been assumed. The following higher level buffer structures have been defined:

- **The analytical model of the C-TOBS multi-buffer delay structure with constant packet size and exponential interarrival time** has been defined with help of stochastic Petri net (Figure 6.).

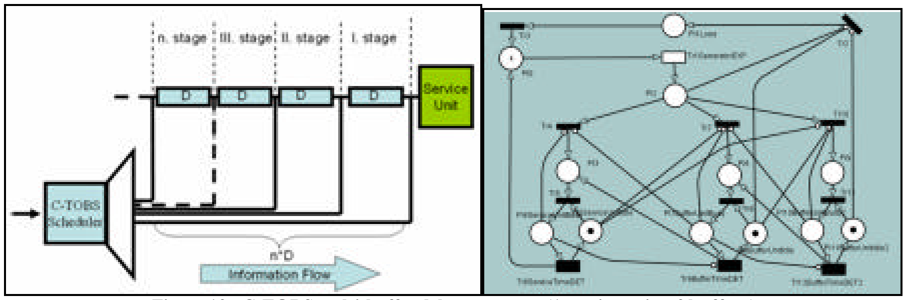

Figure 12.: C-TOBS multi-buffer delay structure (1 service unit + 2 buffers)

- The analytical model of the D-POBS multi-buffer delay structure with constant packet size and exponential interarrival time has been defined with help of stochastic Petri net (Figure 7.).

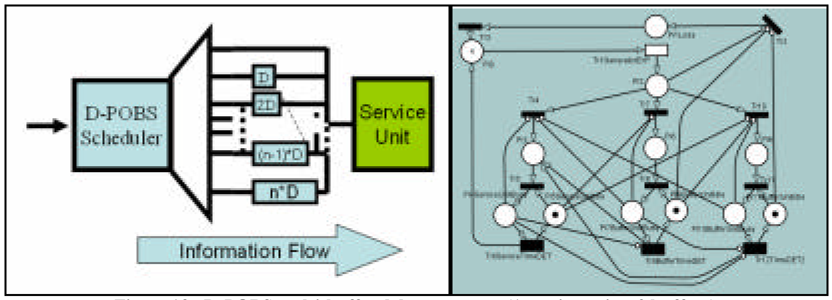

Figure 13.: D-POBS multi-buffer delay structure (1 service unit + 2 buffers)

- The analytical model of the R-TOBS multi-buffer delay structure with constant packet size and exponential interarrival time has been defined with help of stochastic Petri net (Figure 8.).

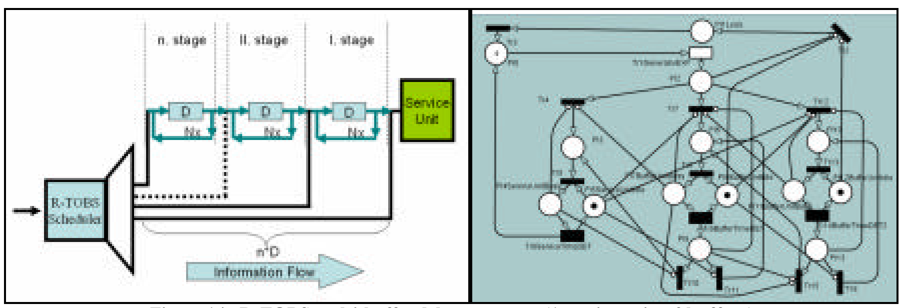

Figure 14.: R-TOBS multi-buffer delay structure (1 service unit + 2 buffers)

- The analytical model of the R-POBS multi-buffer delay structure with constant packet size and exponential interarri val time has been defined with help of stochastic Petri net (Figure 9.).

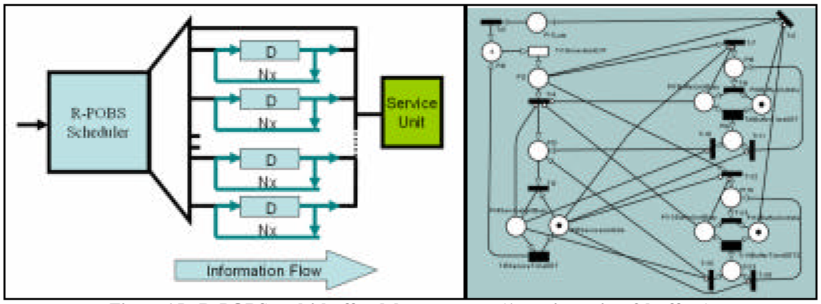

Figure 15.: R-POBS multi-buffer delay structure (1 service unit + 2 buffers)

Relevant section in the dissertation: 5/6.3.
Related publications of the thesis group are located in [S3, S6, S24].

IV. Dissemination of the results

The results described in the theses are published in conference proceedings, journals, and demonstrated at several scientific forums and exhibitions. The topic of the dissertation is tightly connected to the research work, which took/takes place on the Department of Broadband Infocommunication and Electromagnetic Theory at Budapest University of Technology and Economics.
In the European FP6 IST LABLES project -besides the development of the routing control interface- the pre-sizing of the optical delay buffers was done by the help of these results [S7, S10, S22].The developed models and sizing method were integrated into the technical paper of an ETIK project, which was done together with Ericsson. The NSOSS software package (the discrete event simulation based implementation of the all-optical switch model) is available for the research community via the OMNeT++ community site. The simulation results of the developed software package were disseminated successfully in many forums (both in papers and on conferences).
The ClusterPerf measurement software (also developed by myself) has been used mostly at CERN (European Organization for Nuclear Research-Switzerland), and also at FermiLab-ban (Fermi National Accelerator Laboratory-USA), for the measurement of local, high-speed/high-performance distributed datacommunication networks. The ProtocolPerf measurement application has been used at Budapest Tech for SCTP based communication measurements, and for some protocol optimization tasks. The ptATCP software was used in lot of TCP based application communication performance measurements at CERN. This software as a vital module of the XDAQ framework is operating -from the second half of year 2008- in the LHC experiment as part of the CMS DAQ network's software repository. The main task of the application is to provide asynchronous TCP based communication layer between the data processing applications.
My future research targets on this field will be the investigation into performance parameters in case of complex traffic patterns, and network performance evaluation of all-optical switches. My numerical results, related to sizing and congestion resolution of all-optical switches can be applied to build such switches with optimized resource consumption. I hope that the developed discrete event simulation based all-optical switch model will work as a useful research tool for the research community and my communication protocol measurement tools will continue to help effectively engineers and researchers in their work, to evaluate HPC clusters and protocols.

Publications of the author in the field of the Ph.D. thesis

International journal papers in foreign language

[S1] **M. Kozlovszky**, T. Berceli and L. Kutor; Analysis of SCTP and TCP based communication in high-speed clusters, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 559, Issue 1, 1 April 2006, pp. 85-89.

[S2] **M. Kozlovszky**; A TCP/IP transport layer for the DAQ of the CMS experiment, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 534, Issues 1-2, 21. November 2004, pp. 125-129.

Hungarian journal papers

[S3] **Kozlovszky M.**, Berceli T.; Ütközésfeloldási stratégiák modellezése optikai hálózati kapcsolókban, Híradástechnika, 2007 július, pp. 20-25.

International conference papers

[S4] **M. Kozlovszky**, T. Berceli; Analysis of SCTP and TCP based communication in high-speed clusters, X. International Workshop on Advanced Computing and Analysis Techniques in Physics Research, ACAT 2005, May 22- 27 2005, DESY, Zeuthen, Germany.

[S5] **M. Kozlovszky**; Developing a transport layer component for the DAQ framework of the CMS Experiment, IX. International Workshop on Advanced Computing and Analysis Techniques in Physics Research, ACAT 2003, December 2003, KEK, Tsukuba, Japan.

[S25] **M. Kozlovszky,** T. Berceli, V. Kozlovszky, "NSOSS: the Non-Synchronized Optical Switch Simulator", ACM International Conference Proceeding Series, Proceedings of the 2nd international conference on Performance evaluation methodologies and tools, NSTools 2007, Nantes, France, Workshop Session:Simulation models, Article No. 15, ISBN:978-963-9799-00-4.

Hungarian conference papers

[S6] **Kozlovszky M.**, Berceli T., Kozlovszky V., Ütközésfeloldási stratégiák teljesen optikai alapú hálózati kapcsolók esetében, 2007 május, Tavaszi Szél 2007 Konferencia, Budapest, Hungary

[S7] **Kozlovszky M.**, Dr. Berceli T., Dr. Bíró J.; Optimalizált SCML alapú csomagküldés optikai hálózatokban, HTE05, 2005 május, Budapest, Hungary.

[S8] **Kozlovszky M.**; Hálózati kommunikációs problémák és megoldások nagysebességu adatfeldolgozó hálózatokban, HTE04, 2004 május, Budapest, Hungary.

Hungarian conference papers (in foreign language)

[S9] **M. Kozlovszky**, T. Berceli, G. Kovács, V. Kozlovszky, Measurements on optical buffering models made from fiber delay lines, 2007 május, 12th Microcoll Conference, Budapest, Hungary.

[S10] **M. Kozlovszky**, T. Berceli; Optical delay buffer optimization in packet switched optical network, CSCS2006-The Fifth Conference of PhD Students in Computer Science, June 27-30, 2006, Szeged, Hungary.

Technical Reports

[S11] T. Cinkler, P. Koppa, P. Maák, P. Richter, A. Hámori, M. Serényi, A. Báder, G. Kovács, **M. Kozlovszky**, T.Berceli, F. Kárpát, A. Dér, S. Kökényesi; Optical Packet Switching– Feasibility Study and Project Proposal, July 2006, Budapest, Hungary.

[S12] CMS – The TriDAS Project Technical Design Report, Volume 2:Data Acquisition and High Level Trigger; 2002 December, Geneva, Switzerland.

Other thesis related publications

[S13] V. Brigljevic, G. Bruno, E. Cano, S. Cittolin, S. Erhan, D. Gigi, F. Glege, R. Gomez-Reino Garrido, M. Gulmini,J. Gutleber, C. Jacobs, **M. Kozlovszky**, H. Larsen, I. Magrans de Abril, F. Meijers, E. Meschi, S. Murray, A. Oh, L. Orsini, L. Pollet, A. Racz, D. Samyn, P. Scharff-Hansen, C. Schwick, P. Sphicas, J. Varela; FEDkit: a design reference for CMS data acquisition inputs, LECC 2003, Amsterdam, Netherland, 2003 Sept. 29-Oct. 3.

[S14] V. Brigljevic, G. Bruno, E. Cano, S. Cittolin, S. Erhan, D. Gigi, F. Glege, R. Gomez-Reino, M. Gulmini, J. Gutleber, C. Jacobs, **M. Kozlovszky**, H. Larsen, I. Magrans, F. Meijers, E. Meschi, S. Murray, A. Oh, L. Orsini, L. Pollet, A. Racz, D. Samyn, P. Scharff-Hansen, S. Shepelevich, P. Sphicas, and C. Schwick; An Uniform XML-Based Approach to Manage Data Acquisition Hardware Devices, 2003 September.

[S15] S. Aziz, L. Berti, V. Brigljevic, G. Bruno, E. Cano, A. Csilling, S. Cittolin, S. Erhan, D. Gigi, F. Glege, M. Gulmini,J. Gutleber, C. Jacobs, **M. Kozlovszky**, H. Larsen, M. Litmaath, I. Magrans, G. Maron, F. Meijers, E. Meschi, S. Murray,V. O'Dell, A. Oh, L. Orsini, L. Pollet, A. Racz, D. Samyn, P. Scharff-Hansen, P. Sphicas, C. Schwick, I. Suzuki,N. Toniolo, and L. Zangrando; A Simulation of an Event Building Network for the CMS High Energy Physics Experiment, SCI2003, Orlando, Florida, USA, July 27-30, 2003.

[S16] V. Brigljevic, G. Bruno, E. Cano, S. Cittolin, A. Csilling, D. Gigi, F. Glege, R. Gomez-Reino, M. Gulmini, J. Gutleber, C. Jacobs, **M. Kozlovszky**, H. Larsen, I. Magrans de Abril, F. Meijers, E. Meschi, S. Murray, A. Oh, L. Orsini, L. Pollet, A. Racz, D. Samyn, P. Scharff-Hansen, C. Schwick, P.Sphicas V. ODell, I.Suzuki L. Berti, G. Maron, N. Toniolo, L.Zangrando, A.Ninane, S.Erhan, S. Bhattacharya, J. Branson; The CMS EventBuilder, June 2003.

[S17] V. Brigljevic, G. Bruno, E. Cano, S. Cittolin, A. Csilling, D. Gigi, F. Glege, R. Gomez-Reino, M. Gulmini, J. Gutleber, C. Jacobs, **M. Kozlovszky**, H. Larsen, I. Magrans, F. Meijers, E. Meschi, S. Murray, A. Oh, L. Orsini, L. Pollet, A. Racz, D. Samyn, P. Scharff-Hansen, C. Schwick, P. Sphicas, L. Berti, G. Maron, G. Rorato, N. Toniolo, L. Zangrando, M. Bellato, S. Ventura, S. Erhan; Run Control and Monitor System for the CMS Experiment, CHEP-2003-THGT002, Jun 2003. 8pp. ,2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California, 24-28 Mar 2003.

[S18] E. Cano, S. Cittolin, A. Csilling, S. Erhan, D. Gigi, F. Glege, J. Gutleber, C. Jacobs, **M. Kozlovszky**, H. Larsen, I. Magrans, F. Meijers, E. Meschi, S. Murray, A. Oh, L. Orsini, L. Pollet, A. Racz, D. Samyn, P. Scharff-Hansen, P. Sphicas, C. Schwick, T. Strodl; CMS Data to surface Transportation Architecture, Sept. 2002. Prepared for 8th Workshop on Electronics for LHC Experiments, Colmar, France, 9-13 Sep 2002. Published in Colmar 2002, Electronics for LHC experiments 285-288.

[S19] **Kozlovszky M.**, dr. Eged B.; TCP/IP mérések (segédlet), 2003, Budapest, Magyarország.
http://wit.mht.bme.hu/education/description/lab-tcpip.pdf

[S20] **Kozlovszky M.**, dr. Eged B.; GPRS mérések (segédlet), 2001, Budapest, Magyarország.
http://wit.mht.bme.hu/education/description/lab-gprs.pdf

[S21] **M.Kozlovszky**; Blocking vs. Nonblocking I/O Models, 2003 September, CERN, Geneva, Switzerland.

[S22] **M. Kozlovszky**, G. Kovács, T. Berceli; Subcarrier Multiplexed Label (SCML) based routing within a packet switched optical network, 2005 június, PWCOM2005, Gothenburg, Sweden.

[S23] **M. Kozlovszky**, M. Pierié XDAQ ptATCP User Manual V1.2., 2004 10. 26., Geneva, Switzerland.
http://xdaq.web.cern.ch/xdaq/Documents/User%20Manuals/Applications/XDAQ%20ptATCP%20User%20Manual/XDAQ%20ptATCP%20User%20Manual1.2.pdf/XDAQ%20ptATCP%20User%20Manual1.2.pdf

[S24] H. Larsen, **M. Kozlovszky**; XDAQ ptATCP User Manual V1.1.,2003 10. 10., Geneva, Switzerland.
http://xdaq.web.cern.ch/xdaq/Documents/User%20Manuals/Applications/XDAQ%20ptATCP%20User%20Manual/XDAQ%20ptATCP%20User%20Manual1.1.pdf/XDAQ%20ptATCP%20User%20Manual1.1.pdf

Other publications

[S26] **M. Kozlovszky**, K. Karóczkai, I. Márton, A. Schnautigel, P. Kacsuk, G. Hermann, R. Harrington, D. Martin, C. Winsnes, T. Strodl; User oriented Grid testing, 6th Austrian-Hungarian Workshop on Distributed and Parallel Systems, September 21-23, 2006, Innsbruck, Austria.

[S27] P. Kacsuk, **M. Kozlovszky**; P-GRADE Portal: Towards a User-friendly Grid Environment, GCCP 2005, November 29-December 2. 2005, Bratislava, Slovakia, ISBN 80-969202-1-9.

[S28] The ANNOTATOR sequence analysis suite; G. Schneider, M. Wildpaner, F. Leitner, B. Eisenhaber, **M. Kozlovsky**, W. Kubina, S. Maurer Stroh, M. Novatchkova, A. Schleiffer, S. Tian, F. Eisenhaber, ECCB05/JBI Computational Biology, 28th September -1st October 2005, Madrid, Spain.

[S29] Molnár B., Galamb O., **Kozlovszky M.**, Dinya E., Gyorffy B., Tulassay Z.;Vastagbél-biopsziák automatikus analízise mRNS-expressziós profilok statisztikai kiértékelésével, Magyar Belgyógyász Társaság 41. Nagygyulése, 2006. November 09-11. Budapest, Hungary (Magyar Belorvosi Archívum pp. 90-91/ ISSN 0133-5464).

[S30] **M. Kozlovszky**; Biotechnológia oktatás a BMF Neumann Informatikai Karán, IX. Országos Neumann Kongresszus, 2006. június 27–29. Széchenyi Egyetem, Gyor, Hungary.

[S31] M. Wildpaner, G. Schneider,S. Tian, **M. Kozlovszky**, A.Schleiffer, F. Eisenhaber; The Large Scale Sequence Annotation System, 2004, GEN-AU Meeting, Vienna, Austria.

[S33] K. Sandor, **M. Kozlovszky**, V. Kamaras, L. Ficsor, V. S. Varga, B. Molnar, Porting a 3D image registration application to multi-core environment, Spring Simulation Multiconference, Proceedings of the 2008 Spring simulation multiconference (HPCS 2008), Ottawa, Canada, pp. 379-384, 2008, ISBN:1-56555-319-5.

[S32] **Kozlovszky M.**, Drótos D., Karóczkai K., Lovas R., Márton I., Schnautigel A., Balaskó Á., Tóth A.;SEE-GRID, a Dél-európai grid infrastruktúra, Híradástechnika 1945 Vol. LXII.,pp. 7-11, 2007 november.

[S33] K. Sandor, **M. Kozlovszky**, V. Kamaras, L. Ficsor, V. S. Varga, B. Molnar, Porting a 3D image registration application to multi-core environment, Spring Simulation Multiconference, Proceedings of the 2008 Spring simulation multiconference (HPCS 2008), Ottawa, Canada, pp. 379-384, 2008, ISBN:1-56555-319-5

[S34] **M.Kozlovszky**; SEE-GRID-SCI's training model in South Eastern Europe; Zero-In eMagazine (BELIEF-II), 16. January 2009.

[S35] **M. Kozlovszky**, A. Balasko, A. Varga; Enabling OMNeT++-based simulations on Grid Systems, SIMUTools 2009, Roma, Italy, 6. march 2009 (accepted paper)

Reference list

1. Referenced paper (Co-author)

[S13] V. Brigljevic, G. Bruno, E. Cano, S. Cittolin, S. Erhan, D. Gigi, F. Glege, R. Gomez-Reino Garrido, M. Gulmini,J. Gutleber, C. Jacobs, **M. Kozlovszky**, H. Larsen, I. Magrans de Abril, F. Meijers, E. Meschi, S. Murray, A. Oh, L. Orsini, L. Pollet, A. Racz, D. Samyn, P. Scharff-Hansen, C. Schwick, P. Sphicas, J. Varela; FEDkit: a design reference for CMS data acquisition inputs, LECC 2003, Amsterdam, Netherland, 2003 Sept. 29-October 3.

Reference

(H1) ref.[5]Performance of the CMS Silicon Tracker Front-End Driver G. Iles, R. Bainbridge, D. Ballard, I. Church, E. Corrin, J.A. Coughlan, C.P. Day, C. Foudas, E.J. Freeman, J. Fulcher, W.J.F. Gannon, G. Hall, R.N.J. Halsall, J. Leaver, M. Noy, M. Pearson, M. Raymond, I. Reid, G. Rogers, J. Salisbury, S. Taghavi, I.R. Tomalin, O. Zorba, 10thWorkshop on Electronics for LHC Experiments and Future Experiments, Boston, 13-17 September 2004.

2. Referenced paper (Co-author)

[S17] V. Brigljevic, G. Bruno, E. Cano, S. Cittolin, A. Csilling, D. Gigi, F. Glege, R. Gomez-Reino, M. Gulmini, J. Gutleber, C. Jacobs, **M. Kozlovszky**, H. Larsen, I. Magrans, F. Meijers, E. Meschi, S. Murray, A. Oh, L. Orsini, L. Pollet, A. Racz, D. Samyn, P. Scharff-Hansen, C. Schwick, P. Sphicas, L. Berti, G. Maron, G. Rorato, N. Toniolo, L. Zangrando, M. Bellato, S. Ventura, S. Erhan; Run Control and Monitor System for the CMS Experiment, CHEP-2003-THGT002, Jun 2003. 8pp. ,2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California, 24-28 Mar 2003.

Reference

(H2) ref.[7]F. Drouhin et al.; The CERN CMS Tracker Control System, 6 January 2004, CMS Note.

3. Referenced paper (Co-author)

[S18] E. Cano, S. Cittolin, A. Csilling, S. Erhan, D. Gigi, F. Glege, J. Gutleber, C. Jacobs, **M. Kozlovszky**, H. Larsen, I. Magrans, F. Meijers, E. Meschi, S. Murray,A. Oh, L. Orsini, L. Pollet, A. Racz, D. Samyn, P. Scharff-Hansen, P. Sphicas, C. Schwick, T. Strodl; CMS Data to surface Transportation Architecture, Sept. 2002. Prepared for 8th Workshop on Electronics for LHC Experiments, Colmar, France, 9-13 Sep 2002. Published in Colmar 2002, Electronics for LHC experiments 285-288.

Reference

(H3) ref.[15] The Front-End Driver card for the CMS Silicon Tracker Readout, J.A. Coughlan, S.A. Baird, K.W. Bell, E. Corrin, J.A. Coughlan, C.P. Day, C. Foudas, E.J. Freeman, W.J.F. Gannon, G. Hall, R.N.J. Halsall, J. Salisbury, A.A. Shah, S. Taghavirad, I.R. Tomalin, Electronics WORKSHOP, COLMAR, 9-13 September 2002.

References

[1] An Ge, Franco Callegati, and Lakshman S. Tamil; On Optical Burst Switching and Self-Similar Traffic, IEEE Communications Letters, vol. 4, no. 3, march 2000.

[2] F. Callegati, W. Cerroni, G. Corazza, Scheduling Algorithms for a Slotted Packet Switch with either Fixed or Variable Length Packets, Kluwer Acad. Publ. Photonic Network Communication, 8:2,163-176,2004.

[3] Telek M., Horváth A., Számítógép rendszerek teljesítményelemzése, BME jegyzet, 2000, Budapest, Hungary

[4] M. Telek, A. Bobbio, L. Jereb, A. Puliafito, and K. Trivedi, ``Steady state analysis of Markov regenerative SPN with age memory policy,'' in 8-th Int. Conf. on Modeling Techniques and Tools for Computer Performance Evaluation, Lecture Notes in Computer Science (H. Beilner and F. Bause, eds.), vol. 977, pp. 165-179, Springer Verlag, 1995.

[5] R.German; Markov regenerative stochastic petri nets with general execution policies:suipplementary variable analysis and a prototype tool,Elsevier Performance Evaluation 39 (2000) pp.:165-188.
[6] A. Zimmermann, J. Freiheit, R. German, G. Hommel: Petri Net Modelling and Performability Evaluation with TimeNET 3.0. 11th Int. Conf. on Modelling Techniques and Tools for Computer Performance Evaluation (TOOLS'2000), LNCS 1786, pp. 188-202, ISBN 3-540-67260-5, Springer-Verlag, Schaumburg, Illinois, USA, 2000
[7] Reinhard German, Christian Kelling, Armin Zimmermann an Günter Hommel, TimeNET A Toolkit for Evaluating Non-Markovian Stochastic Petri Nets,Report 1994-19, Die Professoren des Fachbereichs Informatik der Technischen Universitaet Berlin, Technische Universitaet Berlin, 1994
[8] "The OMNeT++ Discrete Event Simulation System," Andras Varga. In the Proceedings of the European Simulation Multiconference (ESM'2001). June 6-9, 2001. Prague, Czech Republic.
[9] A. Varga. OMNeT++ Object-Oriented Discrete Event Simulation System User Manual. URL reference http://www.omnetpp.org/doc/manual/usman.html, 2006.

Abbrevations

Rövidítés	Teljes név
AON	All-Optical Networks
ATM	Asynchronous Transfer Mode
API	Application Programming Interface
APD	Average Packet Delay
ASIC	Application Specific Integrated Circuit
AVBL	Average Virtual Buffer Length
AWG	Arrayed Wave Guide
BER	Bit Error Rate
CERN	European Organization for Nuclear Research
CIOQ	Combined Input Output Queuing
CMS	Compact Muon Solenoid
CPO	Coherent Population Oscillations
C-TOBS	Constant Buffer in Tandem Optical Buffer Structure
CWDM	Coarse Wavelength Multiplexing
DAQ	Data Acquisition
DET	Deterministic Time Transition
DES	Discrete Event Simulation
DLOB	Dual-Loop Optical Buffer
D-POBS	Degenerate Buffers in Parallel Optical Buffer Structure
DPSK	Differential Phase-Shift Keying
DSPN	Deterministic and Stochastic Petri Net
DR	Deflection Routing
DWDM	Dense Wavelength Multiplexing
ECS	Electronical Circuit Switching
EIT	Electro-magnetically Induced Transparencies
EPS	Electronical Packet Switching
EXP	Exponentially Distributed Time Transition
FB	Feedback Buffering
FCFS	First Come First Served
FDL	Fiber Delay Line
FPGA	Field-Programmable Gate Array
Gbps	Gigabits per second
HOL	Head-Of-Line
IATime	Inter arrival time distribution
IP	Internet Protocol
IQ	Input Queuing
LAN	Local Area Network
LHC	Large Hadron Collider
MAN	Metropolitan Area Network
Mbps	Megabits per second
MRGP	Markov Regenerative Stochastic Processes

NIC	Network Interface Card
NSOSS	Non-Synchronized Optical Switch Simulator
OBS	Optical Burst Switching
OCS	Optical Circuit Switching
O/E/O , O-E- O	Optical-Electrical-Optical
OPS	Optical Packet Switching
OTN	Optical Transport Network
OQ	Output Queuing
PFBS	Port Level Feedback Buffering
PLP	Packet Loss Probability
PLD	Packet Loss Distribution
PN	Petri Net
POBS	Parallel Optical Buffer Structure
PPS	Photonic Packet Switching
PS	Packet Switching
QoS	Quality of Service
RAM	Random Access Memory
RFC	Request For Comments
RMSD	Root-Mean- Square Deviation
RPR	Resilient Packet Ring
R-xOBS	Recirculating Buffers in Parallel / Tandem Optical Buffer Structure
R-TOBS	Recirculating Buffers in Tandem Optical Buffer Structure
R-POBS	Recirculating Buffers in Parallel Optical Buffer Structure
SFBS	Switch Level Feedback Buffering
SC	System Configuration
SCML	Subcarrier Multiplexed Label
SCTP	Stream Control Transmission Protocol
SCWP	Scattered Wavelength Path
SDH	Synchronous Digital Hierarchy
SHWP	Scattered Wavelength Path
SOA	Semiconductor Optical Amplifiers
SOHO	Small Office/Home Office
Tbps	Terabits per second
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TOBS	Tandem Optical Buffer Structure
TWC	Tunable Wavelength Converter
UDP	User Datagram Protocol
VBL	Virtual Buffer Length
WAN	Wide Area Network
WDM	Wavelength Division Multiplexing
WFB	Wavelength Level Feedback Buffering