Evaluation of Communication Latency for Future Time-Critical Smart Grid Measurement and Control systems

Frank Ball*

Frank Ball Consulting
Oxford UK
frank_ball@ntlworld.com

Kashinath Basu

Department of Computing and Communication
Technologies
Oxford Brookes University
Oxford, UK
kbasu@brookes.ac.uk

Ali Maqousi and Tatiana Balikhina
Faculty of Information Technology
University of Petra
Amman, Jordan
amaqousi@uop.edu.jo; tbalikhina@uop.edu.jo

Abstract— **This paper focuses on the firm real-time requirements of Time-Critical Wide Area Measurement and Control systems for future Smart Grids. It outlines the findings from the first stage in an ongoing body of work aimed at developing models and techniques that will enable the performance evaluation of these systems to take place prior to their implementation. It then discusses the first phase in the second stage of this project that addresses the problem of evaluation in cases where the output from system's devices is variable. It then   presents some preliminary findings and outlines the direction for future work.**

Keywords- Smart Grid Control Systems; Time-Critical Applications; Performance Evaluation.

I. INTRODUCTION

In previous work [1][2] we have completed the first stage of an ongoing investigation into the problem of evaluating the temporal performance of the Smart Grid Synchrophasor Measurement and Control Systems, proposed to provide Wide Area Automated Control in the smart grids of the near future [3][4]. In contrast to other investigations in this area [5], that focus mainly on the evaluation of existing systems, our approach is to consider the evaluation of these future systems as part of their development which is to be carried out prior to their implementation. The investigation is motivated by the need to ensure that the temporal requirements of time-critical applications can be met prior to the system becoming operational, since failures in a grid control system can lead to very serious consequences [4]. The application level latency targets of these systems in are the order of 8-10ms, with a limit of 2ms for network communication delay [3][4], and given that these Systems operate over significant distances, propagation delay will be the major contributor to the latency. Due to the potentially short intervals available for forwarding and other network activities, the performance evaluation of these control system will need to be carried out at the microsecond level, [1]. Our investigation is focused on developing performance analysis and evaluation techniques that can advise and support the design and construction of future real-time devices, and the development of these time-critical control applications. Our aims are: firstly, to derive generic and parameterized models to support the performance evaluation of these systems; and secondly, to develop techniques and methods to evaluate the temporal performance of models based on specific proposed future systems. The major objective in the first stage of our investigation was to develop a processes for the evaluation Synchrophasor Measurement and Control Systems that are to be implemented using high performance equipment which has relatively deterministic behaviour. This stage is now complete and has provided a basic framework for further development. The objective of the second stage is to extend the evaluation process to accommodate variability in the output of the measurement and monitoring devices. The remainder of this paper is structured as follows: Section II. Presents the general operation of the fundamental components of a Synchrophasor Measurement and Control System, it also identifies the systems communications model, together with its resulting queuing network model; Section III, Outlines the findings from the first stage of the investigation; Section IV. Discusses the potential variability in the output from the monitoring devices, and presents work in progress toward accommodating this variability within the performance evaluation process; and finally, Section V, Concludes and discusses plans for future work.

II. SYNCHROPHASOR MEASUREMENT SYSTEMS

The proposal to extend automated control systems into the wide area requires a stepped change in overall performance, for which certain current practices for their design and implementation may no longer be sufficient. Motivated by this, researchers have put forward new guide lines and recommendations that will help to meet the new stringent latency requirement [3][4]. In line with these recommendations

our investigation is based on the following conditions being observed [1][2]: Strict Priority Queuing throughout; Forwarding decisions based on packet header only; Static routing for the time-critical classes; and no retransmission for error recovery, instead use redundant disjoint paths for reliability. Furthermore our investigation takes into account the guidelines to exploit the relatively small scale of a Synchrophasor Measurement System, and a priori knowledge of predicable traffic [3].

A detailed discussion on the operation of both existing and proposed future Synchrophasor Measurement Systems is presented in [1]. A Synchrophasor is a measurement of the amplitude and angle of a sinusoidal waveform (in this case the waveform of the power cycle) that is time-stamped using a UCT (Universally Coordinated Time) mechanism facilitated by GPS. These synchronized measurements provide a comprehensive picture of state of the power system and are taken by a Phasor Measurement Unit (PMU). This is a specialized device that periodically samples the power cycle at a particular item of equipment. It then calculates the synchrophasor measurement which is then encapsulated into a single fixed length message for transmission to a local Phasor Data Concentrator (PDC). This device then checks the validity of the messages before forwarding them via a WAN, to the Super Phasor Data Concentrator (SPDC), which in turn has a direct connection to the Controller. Therefore in each cycle the PDC outputs a batch of fixed length messages, with a typical size of 1000bits including protocol overheads. The length of the messages will be constant throughout the system but the size of the batches may vary for each PDC. The end-to-end latency of the system is defined as the time between the timestamp value of the message and completion of the control decision process [3][4]. The ultimate objective for synchrophasor based wide area control applications is to carry out the measurement-to-decision process within one power cycle [3][4], that in the case of a 60hz power cycle is 16.7ms, of which 8,35 is taken up by the sampling process.

If, for example, we chose a target of 2ms for the networking latency component then it is clear that all the message of a particular batch will have cleared the network before the transmission of the next batch begins. The PDCs prime function is that which is most vital to time-critical performance, and this is shown below in Fig 1.
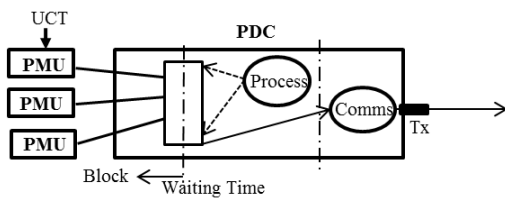

Figure 1. PDC Prime Function Basic Processing Model

Although a PDC may also be responsible for other functions, e.g. storing information to support historical analysis, it is important to ensure that the operation of these auxiliary functions cannot interfere with that of the prime function. One particular aspect of the prime function's basic operation that is very relevant to this discussion, is the PDCs use of a Waiting-Time parameter. It is this, which makes it possible to evaluate the network latency with some degree of isolation [1][2]. A PDC waits until all the messages it expect to receive in a given cycle have arrived, before starting to process them as a batch. To allow for message losses, the PDC employs a Waiting -Time parameter which is set to the time that the PDC expects the last message in that cycle to arrive. Once this time has expired, it starts to process the messages that have arrived in time and discards any messages from that cycle that arrive later. A similar process is employed by the SPDC, and in both cases a late arrival is equivalent to a loss. After processing the batch, the PDC starts to forward the messages and the rate at which they are sent will be limited by either the PDC's communication process or the link rate transmission process, whichever is the slowest. In the first stage evaluation process output at link rate was assumed for all PDCs. Determining the value on which to set the PDCs Waiting-Time parameter requires an evaluation of the PMU to PDC latency for each PDC. However, to simplify discussion we will assume that the waiting time is the same for all PDCs. In general, the Waiting-Time setting and start of output for each individual PDC would need to be calculated prior to the main evaluation process.

The Communication model for a Wide Area Synchrophasor Measurement System is that of a Many-to-One with synchronised sources, that produce output periodically. A simple example of which is shown below in Fig 2.
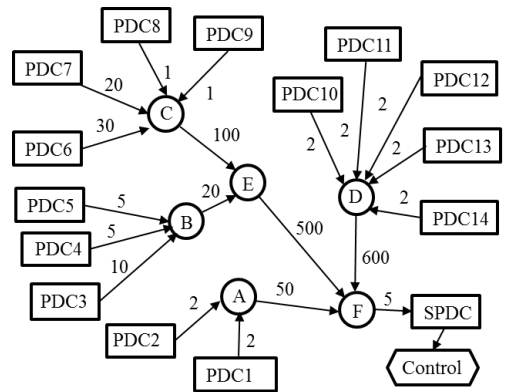

Figure 2. PDC Fan in to SPDC.

In this diagram the nodes A to F relate to the forwarding devices, and numbers assigned to the links represent the propagation delay of each link in micro seconds. Unlike streaming applications in which a message is produced periodically and transmitted immediately, each source creates a number of messages before the first message is sent. All the messages produced in each cycle are members of a set that needs to be received in total, within a given time period. However, the distance between source and destination can differ significantly between one source and another. Therefore, the messages from some sources will arrive well ahead of the deadline whilst those of other sources may arrive much closer to the latency limit. These factors need to be taken into account when considering the QoS requirements for this type of traffic.

Time-critical applications generally require a hard deadline, however, in certain cases some missed deadlines may be acceptable. some missed deadlines and losses may acceptable [2][3][7]. However, the term "some" is a rather vague parameter and as an alternative we propose the use of probabilistic hard real-time, that offers a more precise definition, i.e. for a latency bound T, a condition, such as the following, must hold:

$$P(t > T) \leq n10^{-x} \qquad (1)$$

This allows the application to choose an allowable value for loss, which can then be related directly to high percentiles of delay distributions during the evaluation process. Also the timescale to which this condition should apply may vary depending on the particular application. For example, in some cases it may acceptable to have several missed deadlines close together, provided that the number of missed deadlines over a longer period are within the given limit. Conversely, other applications may require significant spacing between missed deadlines, in addition to an overall limit within a longer period. For the case of synchrophasor measurement system it has been recommended that this guarantee should apply to each individual message, and not to the aggregation of the batch in each cycle [3].

Finally, it is important to note that the SPDC also uses a waiting time parameter. The SPDC's setting could be determined from the results of an evaluation, by using the overall worst case delay, in situations where that delay would satisfy the probabilistic hard-real-time requirements.

The process for evaluating the latency of the paths between the PDCs and the SPDC is based on an open Queuing Network Model (QNM). A simplified example of which is shown below in Fig 3.
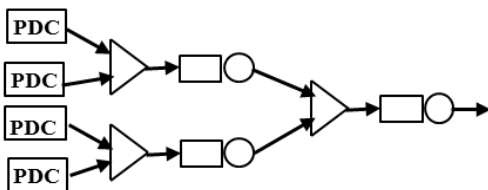

Figure 3.   Queuing Network Model.

In the first stage of our investigation the evaluation was based on the assumption of best case equipment i.e. deterministic line rate output from the PDCs, line rate switching in the forwarding devices, and pre-emptive strict priory queuing at the output queue. In this case the forwarding function was modelled as being a multiplexer with a pipelined fixed delay of M μs, followed by the use of pre-emptive strict priority queuing to ensure that the traffic of lower classes could not impose any additional delays on the server. Therefore the resulting queues were of the class D/D/1 (deterministic arrivals; deterministic service times; and a single server) with cyclic short term active periods. Due to the deterministic properties of this case, the input and output functions were the same for each cycle, and a basic network calculus approach was sufficient to carry out the evaluation. For this second stage

of the investigation we have removed the assumption of deterministic output from the PDCs. and replaced it with the assumption of stable non-deterministic output with known variability. In this case the queuing model becomes that of the class G/D/1 (generally distributed arrivals; deterministic service times; and a single server) with cyclic short term active periods for which the input function varies between each cycle. Due to the short term queuing behaviour, analytical solutions based on classical queuing theory will not be applicable. Basic network calculus was useful in the case of deterministic PDC output since the arrival functions were the same for each traffic cycle. However, variability in PDC output means that the arrival and output functions will vary for each cycle, both in duration and arrival pattern. Therefore, simulation is probably the most appropriate option for evaluation.

III.   OUTCOMES OF THE FIRST STAGE EVALUATION PROCESS

The evaluation processes we developed in the first stage of our work [1][2] is based on the conditions outlined above in section II. It also based on a link rate of 500Mb/s, the data rate supported by the fibre optic carriers contained within an OPGW (Optical Ground Wire Systems) cable. Since these cables provide both protection and communications and can be installed on high voltage pylons it quite likely that they will be used for communication at this level in the grid. This evaluation process offers only a limited solution to the problem of evaluating the temporal performance of Synchrophasor Measurements Systems. However, it still offers some useful properties to the evaluation process in general. In particular, this evaluation process is relatively straight forward and does not require any extensive processing. Therefore, it provides a relatively quick method for carrying out preliminary evaluations that could help to identify source that may be the most vulnerable to missed deadlines. This could be potentially useful, for example, when deciding on the routing paths for the fan-ins to the SPDC. Experimental evaluations carried out using the deterministic evaluation process have shown that sources with the longest paths (in propagation delay) do not always have the highest overall latency, and that interaction in the queues also has an important influence.

IV.   ADDRESSING VARIABILITY IN PDC OUTPUT

Although these is no principled reason why PDCs could not be designed to produce deterministic output, from a pragmatic perspective some degree of variability is inevitable. However, it will important to ensure that; any variability is stable; does not have a too wide a range; and is made known. Ideally, this information should be provided in the form of a probability distribution. One possible way of obtaining such information could be through the results of compliance tests [1] where in addition to ensuring correctness, which is currently the main consideration, output performance is also measured and evaluated.. We intend to investigate this possibility in greater detail in future work. However, since at this stage of the investigation the main objective is to develop an evaluation process, rather than evaluate an existing system, in the examples presented below we will assume values that we believe could represent typical and appropriate output

characteristics for PDCs that have been specifically designed for time-critical operations.

A probability mass function (pmf) is a function that gives the probability that a discrete random variable is equal to some value. One example of a pmf that might represent the output of a PDC designed for time-critical operation, is shown below in Fig. 5.
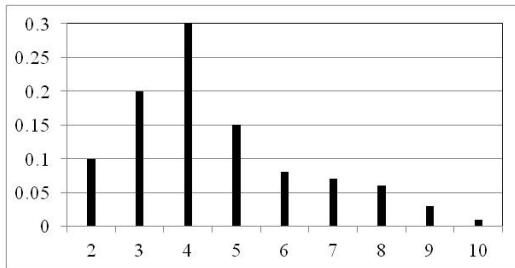

Figure 4.   PDC Fan in to SPDC.

In this example the discrete unit is the microsecond, however, if greater accuracy is required a finer granularity would be used. Also both the range and shape of the distribution have been chosen arbitrary in line with the comments made in the previous paragraph. Given that the pmfs for the PDCs have been made available, then together with knowledge of the Fan-in to the SPDC, we can construct a Discrete Event Simulation (DES) based on the QNM introduced in section 2.3. However, since we are interested in studying the effects of relatively rare events that could result in deadlines being missed, rather than the evaluation of average performance, runtimes may need to be quite long. Therefore, runtime reduction is an important consideration. Although modelling the system as a QNM and using a building block approach may allow some degree of parallel processing, other possibilities for runtime reduction also need to be considered. One obvious step toward run time reduction would be to eliminate sources and nodes close to the SPDC whose output will have transited the final forwarding device (Node F in figure 4) long before that of the more distant nodes starts to arrive. Node A in figure 4 which serves PDCs 1 and 2, is an examples of where elimination would be possible.

Another possible option is to focus on the variability in burst duration rather than that of the individual message inter-arrival times. In each cycle the PDCs will transmit a fixed number of fixed length messages, a typical length for which is 1000bits. However, the inter-message interval time of each will vary in accordance with the pmf, and the burst durations will be the sum of a number of random variables taken from the pmf. Therefore, given the pmf for the PDC output process, and a number of messages n, then the pmf of the burst durations can be derived by the n fold convolution of the PDCs pmf. Although the convolution process requires the assumption of mutual independence, which may not be valid in all cases, biased convolution approaches have been developed and shown to be effective for evaluating the performance of Probabilistic Hard Real-Time System in other areas of application [6]. However, to simplify discussion, in this paper we will focus on the case were the requirement of mutual

independence holds. The number of messages sent from a PDC depends on how many PMUs it serves, with 20 being a typical number. The pmf for the burst durations from a PDC serving 2 PMUs is shown below in figure 5.
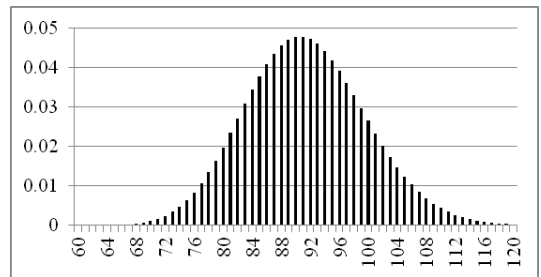

Figure 5.   pmf of burst durations of 20 messages

As would be expected with a series of self-convolutions the probability mass has moved toward the central value and although the maximum possible range of burst duration would be 40 to 200μs, in this case the reachable range lies between 60 and 120μs. Also in the cases where a number of bursts arrive simultaneously at the multiplexer and are formed into a joint arrival process for the queue, the pmf for the resulting burst duration can be formed using conditional probabilistic arguments. In this case the conditions required to form a joint burst duration of d, are that there must be at least one bust of duration d, with the duration of all the others to being less than or equal to d.

The pmf for the case of five simultaneously arriving busts with durations compliant to the pmf in figure 5, is shown below in figure 6, together with the original distribution from which it was formed. As would be expected due to the required conditions to form a combined burst of duration d, it can be seen the resulting distribution has shifted to the right of the distribution from which it was formed. In this case the reachable range of burst duration lies between 80 and 130μs. and the number of messages arriving within this interval will be the sum of the messages from each of the individual burst
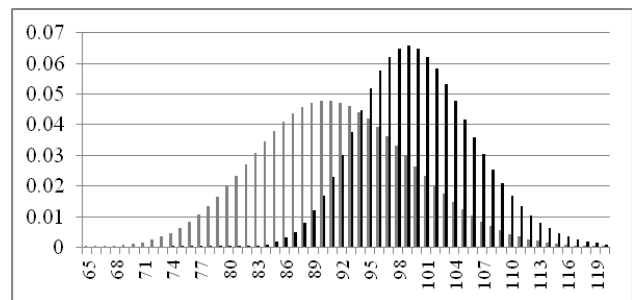

Figure 6.   pmf for 5 simultaneously arriving bursts

In the case shown above the simultaneously arriving bursts all belong to the same distribution and all have the same number of messages. However, it is also possible to derive a combined burst pmf in the case where both the original distribution and the number of messages for each burst differ. The same conditions hold, and the basic principle used to form the distribution remains the same, however, the process of

formation requires a few more steps. In every cycle of transmission a burst comprising a fixed number of messages will arrive at the queue at a particular time. The duration of the burst will vary randomly between each cycle, in accordance with a specific pmf. From the perspective of the queue an important factor is the utilisation of the server, which may also vary at each cycle

The server utilisation u_i for a particular burst b_i is given by:

$$u_i = ns/d_i \qquad (2)$$

where n = the number of messages; s = the messages service time; and d_i is the duration of b_i.

The message service time s is given by :

$$L/R \qquad (3)$$

: where L is the length of the message in bits; and R is the line transmission rate in bits/second.

In our investigation we are assuming a message length of 1000bits (which is typical in current systems); and a transmission rate of 500Mb/s, based on the capacity offered by OPGW fibre optic carriers, as discussed in section 3, therefore, in this case s=2μs. In cases where the queue server utilisation is less than one, the server will be inactive at some time during the burst duration. Otherwise it will be active for a period greater than, or equal to, the arrival burst duration. When this is the case the output function will become deterministic. If we consider the case shown in figure 7, where the reachable range is from 80 to 130μs, the utilisation for both extremes are: 100x2/80 =2.5; and 100x2/130=1.54. Therefore, for every cycle the output function burst from the queue will be deterministic with a constant duration of 200μs. An example of where this situation would occur is shown at node D in figure 4. In this case simulation of node D and all its sources could be replaced by a constant output function. Although this is obviously a special case, cases similar to this may not necessarily be rare. For example, this situation could occur where a number of PDC are located in a small geographical area and are all attached to the same local forwarding device.

However, if we consider a case of 2 simultaneously arriving burst from distribution 5, then the reachable range will be from 70 to 120μs and the total number of messages will 40. In this case the range for which the condition ui >=1 holds, now becomes 70 to 80μs. In this case the probability of a deterministic output is quiet small. Although, in a different cases where the range covers a significant proportion of the probability space, then some reduction in simulation effort might be possible

However, for the cases where required condition does not hold. or only holds partially, within the reachable range, the evaluation process we need to consider how use the information provided by the burst length pmf. Unfortunately, knowing only the burst length is insufficient, since for any given burst length, individual bursts may follow different trajectories. An example of 5 different trajectories relating to bursts of the same length, and each with same number of messages is shown below in Fig 7.
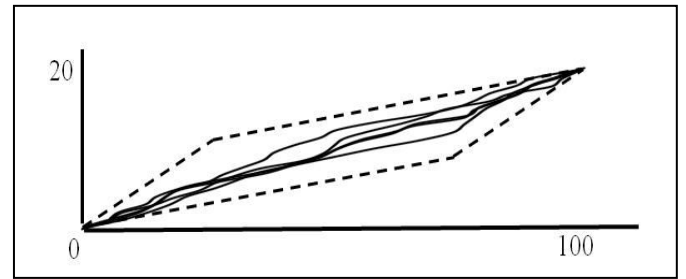

Figure 7. Examples of Burst Trajectories

In this particular example the burst duration in 100μs and the number of messages is 20. However, any other values would show similar behaviour. Although shown on the same diagram, each of the 5 trajectories would belong to a different cycle. Each burst will follow the same general trend line but take different paths either side of, and sometimes crossing, that trend line. However, because both the burst duration and the number of messages are the same for each burst, there is a limit to how far the trajectory can vary in either direction from the trend line. The Parallelogram shown by the dotted lines in figure 7 represents the maximum possible deviation from the trend line. The upper bound can by derived by calculating the maximum number of consecutive messages with the minimal inter-arrival time that can be accommodated, under the condition that the burst duration remains unchanged. It is, therefore, also limited by the maximum possible inter-arrival time. The lower bound is the complement of the upper bound. We have carried out some preliminary experiments using these bounds which have shown that for a certain service rate that is slightly higher than the burst trend rate, in the case of the upper path, the burst duration of the output function will be the same as that of the arrival function. However, in the case of the lower path the burst duration of the output function was shown to be significantly longer. However, these theoretical bounds are virtually unreachable. In the example shown figure 7, the probability of achieving the maximum number of consecutive message with the minimal inter-arrival time alone, is in the order of 10^{-10}. However, we have evaluated a number individual burst trajectories that tended toward the lower path, whist still remaining some distance from it. In each case there was still an extension to the burst duration, although of lesser significance. Therefore, the next step in our investigation will be to develop a method for deriving reachable worst case trajectories. Currently, we are focussing on the use of Monte Carlo Simulation (MCS) to derive the reachable worst case trajectories, which is an appropriate choice given that MCS is commonly used in the study of trajectories in other areas. This part of our work is still in the set up stage and as yet no results have been produced. However, in parallel with the development of the MCS, we are also considering the viability of using trajectories in conjunction with a modified version of our first stage evaluation process, as a potential alternative for deriving worst case delay bounds.

V.   Conclusion and Future Work

The work we have presented in this paper is part of an ongoing investigation, and follows on from the completion of

first stage of this work. The paper has outlined the findings from the first stage with particular emphasis on those aspects most relevant to the next stage of the investigation. Following this, the paper has presented an analysis of the expected variability in the output from PDCs. It has then discussed the usefulness of probability distribution functions and the use of convolution techniques, together with other methods, for deriving higher level distributions from the original device output distribution. Following on from this, the paper then considered how certain properties could be exploited to reduce simulation run times. It has also identified that in addition to the burst's length, it's trajectory must also be considered. In the immediate future we will continue to study the properties of burst trajectories and their potential use in a modified version of the first stage evaluation process. The main goal of this work is to find whether or not it is possible to develop a hybrid analytical/MCS method for evaluation that could replace the use of DES.

[1] F. Ball and K. Basu, K., "Performance Evaluation of Time-Critical Smart Grid Applications", Proceeding of INC, July 2016, Frankfurt, DE

[2] A. Maqousi, T, Balikihn, K. Basu and F. Ball,"Network level Quality of Service (QoS) challenges for Smart Grid Measurement and Control systems", Proceeding of IT-DREPS, (Dec 2017), Amman Jordan.

[3] D.E. Bakken, A.Bose, C.H. Hauser, D.E. Whitehead, and G.C Zweigle, G.C, "Smart Generation and Transmission With Coherent, Real-Time Data", Proceedings of the IEEE, Vol 99, No 6. 2011

[4] K.C. Budka, J.G, Deshpande, T.L. Doumi, M. Maddan and T. Mew,."Communication Network Architecture and Design Principles for Smart Grids", Bell Labs Technical Journal, vol. 15 No. 2, 2010.

[5] K. Zhu,.M. Chenine, L. Nordström, S. Holmström, and G. Ericsson, .(2013)" Communication Delay in a Utility TCP/IP Network", International Journal of Emerging Electric Power Systems, 14, (4), pp 341–350.

[6] G. Bernat,. A Colin and S.M. Petters,(2002), "WCET Analysis of Probabilistic Hard Real-Time Systems", 23rd IEEE Real-Time Systems Symposium, RTSS, 2002.
.