Results of pilot project testbed modeling

K. Korcyl
CERN, Cracow
J. C. Vermeulen
NIKHEF, Amsterdam

Draft version 0.6 (incomplete), 16 April 2001

This note contains a description of a model of the ATLAS LVL2 pilot project testbed. The model has been implemented in Ptolemy as well as in Simdaq. Part of the model parameters have been taken from [2], the remaining parameters have been extracted from the measurement results. With these parameters already a large fraction of the measurement results can be calculated with simple formulas. A comparison between results calculated, Ptolemy and Simdaq model results and measurement results is presented.

1 Introduction

In the autumn of 1999 a number of measurements on the LVL2 pilot project testbed at CERN, running the reference software, has been made [1]. A model of this system has been set up in Ptolemy, a description of this model together with a partial set of measurement and modeling results can be found in [2]. In this note a larger set of measurement results is studied.

An assessment of the model parameters was made. Processing times and their dependence on event fragment size and on number of RoI requests and fragments per event were determined. This was done making use of testbed configurations for which the rate was determined by 100 % utilization of the component of interest (ROB emulator, supervisor or steering processor). The functioning of the steering processor has been studied in more detail for testbed configurations for which most actions in the system were completely sequential. By making use of the dependence of the results on event fragment sizes and on number of RoI requests and fragments per event the remaining parallelism could be understood. This understanding and detailed knowledge on all processing and transfer times made it possible to calculate the latencies and the event rates for these configurations. For all other configurations with two or more ROB emulators the rate was found to be completely determined by the fact that either the ROB emulators, the supervisor(s) or the steering processor(s) were utilized at 100 %. For all configurations the latency can be calculated directly from the rate, as the supervisor in all cases only outputs a new steering request after inputting a decision from a steering processor. Therefore the average latency is a multiple of the average time interval between the arrival of successive decisions, i.e. a multiple of the inverse rate. The multiplication factor is equal to the maximum number of outstanding events. For some of the configurations with one ROB emulator the rates could be calculated assuming 100 % utilization of the steering processor. For the remaining configurations with a single ROB emulator the differences between measured inverse rates and inverse rates calculated assuming 100 % utilization showed a systematic behavior. This probably can be explained from changes in the order of processing steps dealing with different events in the steering processor (TO BE CHECKED WITH COMPUTER MODELS).

The measurement results are compared to the results of a revised model implemented with Ptolemy [2] and a model implemented with Simdaq [3]. Both models use updated parameters obtained from the analysis described in this note.

2 The pilot project testbed

2.1 Configurations

The testbed system modeled consisted of 1 or 2 supervisors, 1 to 8 ROB emulators and 1, 2 or 4 steering processors. The configurations studied are specified in Table 1. The total number of configurations is 128. In the system a mix of 300 and 400 MHz Pentium II machines has been used for the nodes in the system. In Table 2 the processor speeds are specified.

Configuration	Number of steering processors	Number of supervisors	Number of ROB emulators (N)	Max. outstanding per processor (O)	Number of threads (T)	Event fragment size (S) (Bytes)
ctbed-s300-O-T-S	1	1	1	2-7	2,3,4	64,1024
ctbed-s400-O-T-S	1	1	1	2-8	2,3,4	64,1024
c-N-tbed-1-1-1-3-S	1	1	2-8	1	3	64,512,1024
c-N-tbed-1-1-4-3-S	1	1	2-8	4	3	64,512,1024
c-N-tbed-2-1-8-3-S	2	1	2-8	4	3	64,512,1024
c-N-tbed-4-2-8-3-S	4	2	2-8	4	3	64,512,1024

Table 1. Testbed configurations. The names used for labeling the configurations have been used for labeling the configuration and dump files of Simdaq. The values of "N", "O", "T" and "S" have to be to be taken from the four last columns.

Configuration	Steering processor	Supervisor	ROBIn
ctbed-s300-O-T-S	300	400	400
ctbed-s400-O-T-S	400	400	400
c-N-tbed-...	300	400	300

Table 2. Processor speeds for the configurations studied.

2.2 The supervisor

The sequence of actions for the supervisor is as follows :

- the supervisor sends to each processor steering requests, until the maximum number of outstanding events is reached ("O" in Table 1). Next requests are only sent if the number of outstanding events becomes lower than the maximum (due to the arrival of a decision sent by one of the steering processors). Events are assigned to each processor using a round-robin strategy. In the case of two supervisors it has been assumed in the Simdaq model that both assign events to all available processors.

- Additional steering requests (one per processor) are generated already if the maximum number of outstanding events is reached. If the supervisor receives a decision while it is generating a new steering request, it will first finish the generation of this request and then send out a new request. The generation time of an event is assumed to be the time at which output of the steering request starts. Once the output of a steering request has been started it cannot be interrupted by input of a decision or generation of a new steering request.
- The time stamps used for measuring the latency can not to be made exactly at the start of the processing associated with the output of a steering request and at the end of the processing associated with handling a decision handling. Hence, the latency needs to be corrected, see Section 3.

It was found that for the configuration with 2 steering processors, 1 supervisor, a maximum of 8 outstanding events, and 3 threads for 2 and 3 ROBs emulators and 64 Byte event fragments the supervisor is utilized at 100 %, i.e. determining the maximum event rate. From the rate measured (8206 Hz) it follows that per event 121.9 µs is available (400 MHz processor). The processing times specified in [2] are listed in Table 3. 37, 64 and 64.3 µs have been specified respectively (300 MHz processor). For a 400 MHz processor these numbers are scaled with a factor 0.75. The steering request generation time has been chosen to be 25.6 µs in stead of 27.75 µs (400 MHz processor, as used in all configurations) to give a total of 121.9 µs processing time per event. For 300 MHz processors this value would need to be scaled to 34.1 µs. These numbers may have to be revised, see the discussion in Section 3.2.

	Steering request generation time (µs)	Decision handling time (µs)	Time required for outputting steering request (µs)
300 MHz	34.1 (37in [2])	64	64.3
400 Mhz	25.6 (27.75 in [2])	48	48.25
Ptolemy parameters			
Simdaq parameter	RoIProcessTime	DecisionProcessTime	OutputProcessTime

Table 3. Processing times for the supervisor, see also section 3.

2.3 The ROB emulator

The ROB emulator is receiving RoIDataRequest messages from the steering processors and is responding to those requests by sending event fragment data together with a 32 Byte header to the steering processor that requested the data. Requests are queued until they can be serviced. It is assumed that no processor time is spent while queueing requests. The processor time spent in servicing a request is assumed to consist of a fixed part and a message size dependent part, see Table 4. The numbers for a 300 MHz processor have been derived from measurements results for the configuration with 2 supervisors, 4 steering processors and 4 ROB emulators and for fragment sizes of 64, 512 and 1024 Bytes, as initial results indicated a 100 % utilization of the ROB emulators. A straight line was fitted to the inverse of the rate measured as function of the message size, see Figure 1, which produced the parameters presented in Table 4.

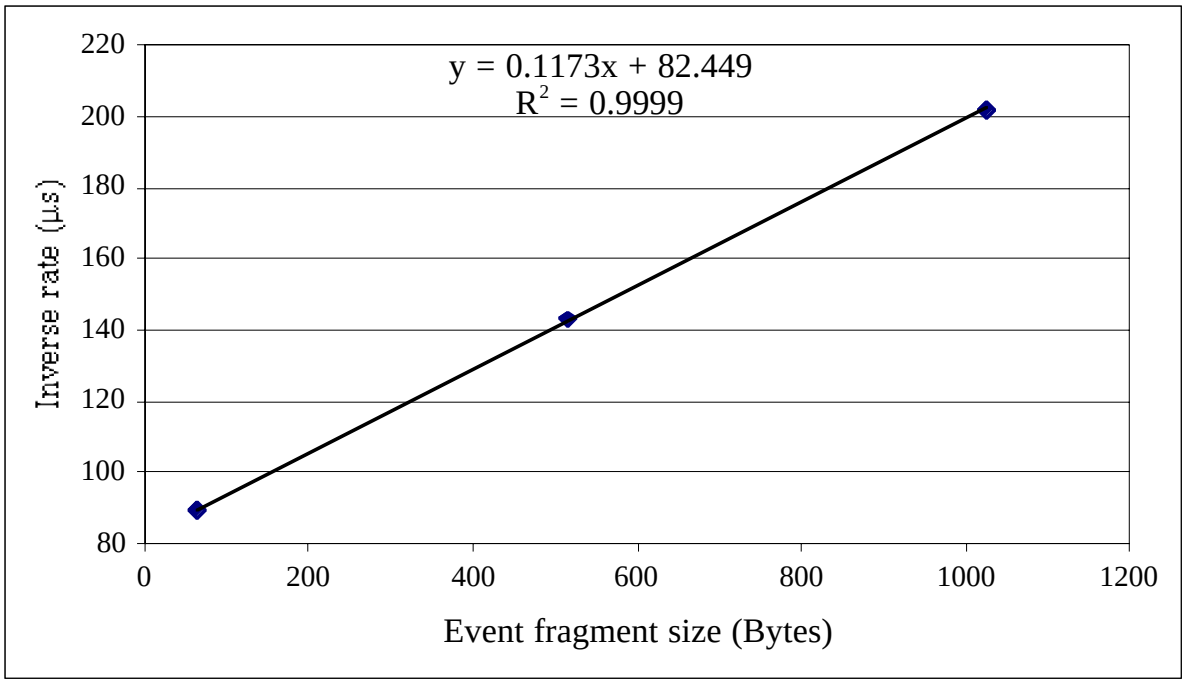

Figure 1. Inverse rate as function of message size for ROB emulator running at 100 % utilization. The measurement results are for the configuration with 2 supervisors, 4 steering processors and 4 ROB emulators.

	Processing time : size-independent part (µs)	Processing time per Byte (µs)
300 MHz	82.4	0.1173
400 Mhz	61.8	0.0880
Ptolemy parameters		
Simdaq parameter	ProcTimeAverage	ExtractByProcessorTransferSpeed

Table 4. Processing times for the ROB emulator.

2.4 The steering processor

The sequence of actions for the supervisor is as follows :

- A steering request is received from the supervisor, after which RoI data requests are generated and sent to the ROB emulators,
- After sending the RoI requests the processor waits for the arrival of the requested event fragments, unless there are further steering requests queued or if fragments from other events arrive,
- Upon arrival of an event fragment it is checked whether all event fragments requested have arrived. If not, the arrival of the remaining event fragments is waited for, unless there is other work for the processor to do,
- If a complete event has been received the event data is merged (including 32 Byte headers), after which a message for the supervisor is generated. In the final system this step would consist of feature extraction, taking a decision and outputting that decision.

The actions described above are executed by a single thread for a single event. The number of threads is either 1, 2, 3 or 4 and therefore also the maximum number of events being handled in parallel by a single steering processor is limited to 1, 2, 3 or 4. Once a thread is running it cannot be interrupted. Input data is handled by a special input thread, which can run only if the threads mentioned earlier are idle. The input thread passes the data to a thread that is not handling an event if a steering request has been received and to the thread waiting for the data received otherwise. Steering requests and fragment data received are queued without spending processor time until the input thread can handle the data. The model of the steering processor is described in more detail in [5].

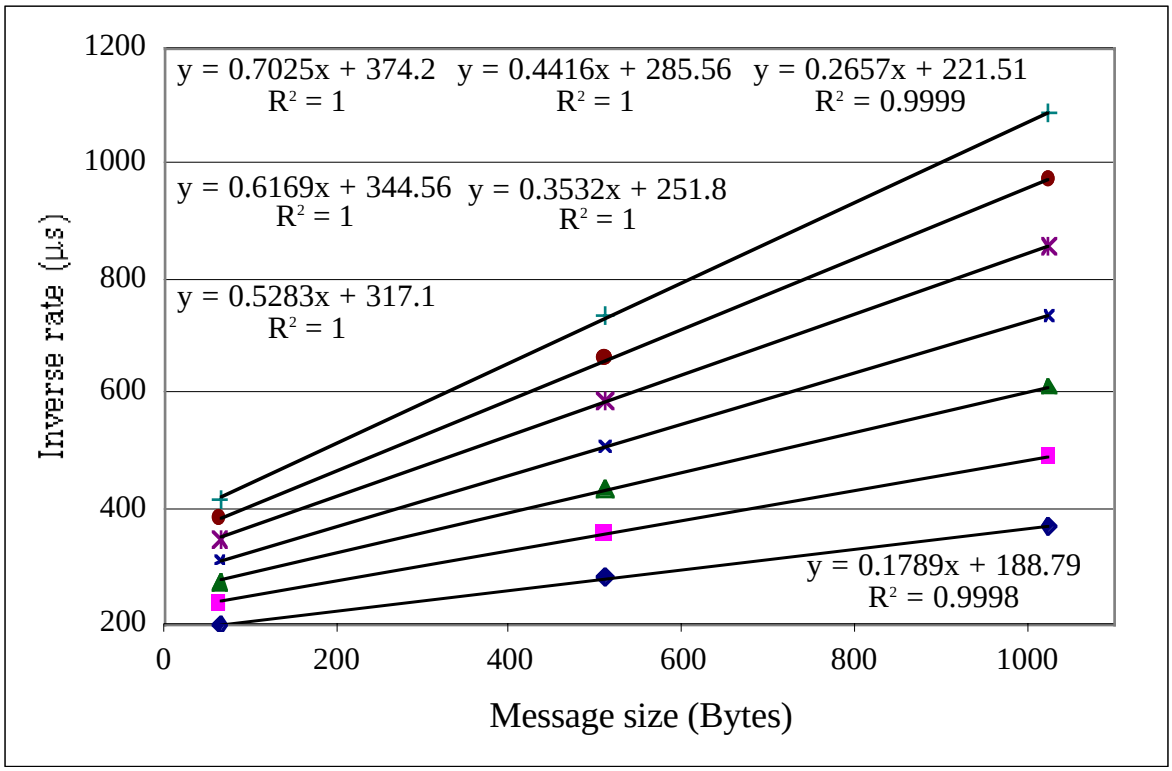

Figure 2. Inverse rate as a function of message size for 2 (lowest line) to 8 (highest line) ROB emulators for the configuration with one supervisor, one steering processor and at maximum four outstanding events. The formulas describe the lines fitted to the measurement results, the line with the lowest slope and offset corresponds to the configuration with 2 ROB emulators.

Initial analysis showed that for a number of configurations the steering processors probably were running at 100 % utilization. The configuration with 1 supervisor, 1 steering processor (400 MHz), 2 - 8 ROB emulators and a maximum of 4 outstanding requests has been used for determining relevant parameters. Figure 2 shows the inverse rate as function of the size of the messages received. By dividing the slope by the number of ROB emulators sending a message an average of 0.0884 µs per Byte received is obtained. This number determines the speed with which data is moved in the steering processor and is in agreement with the 0.0880 µs per Byte found for the 400 MHz ROB emulator for the speed with which data is moved (see Section 2.3). The latter value has been used for further calculations[1] and in the simulations.

1. This value resulted in the data points in Figure 9 to be somewhat better in agreement with a linear dependence on the number of ROB emulators than would be the case for a value of 0.0884 µs / Byte

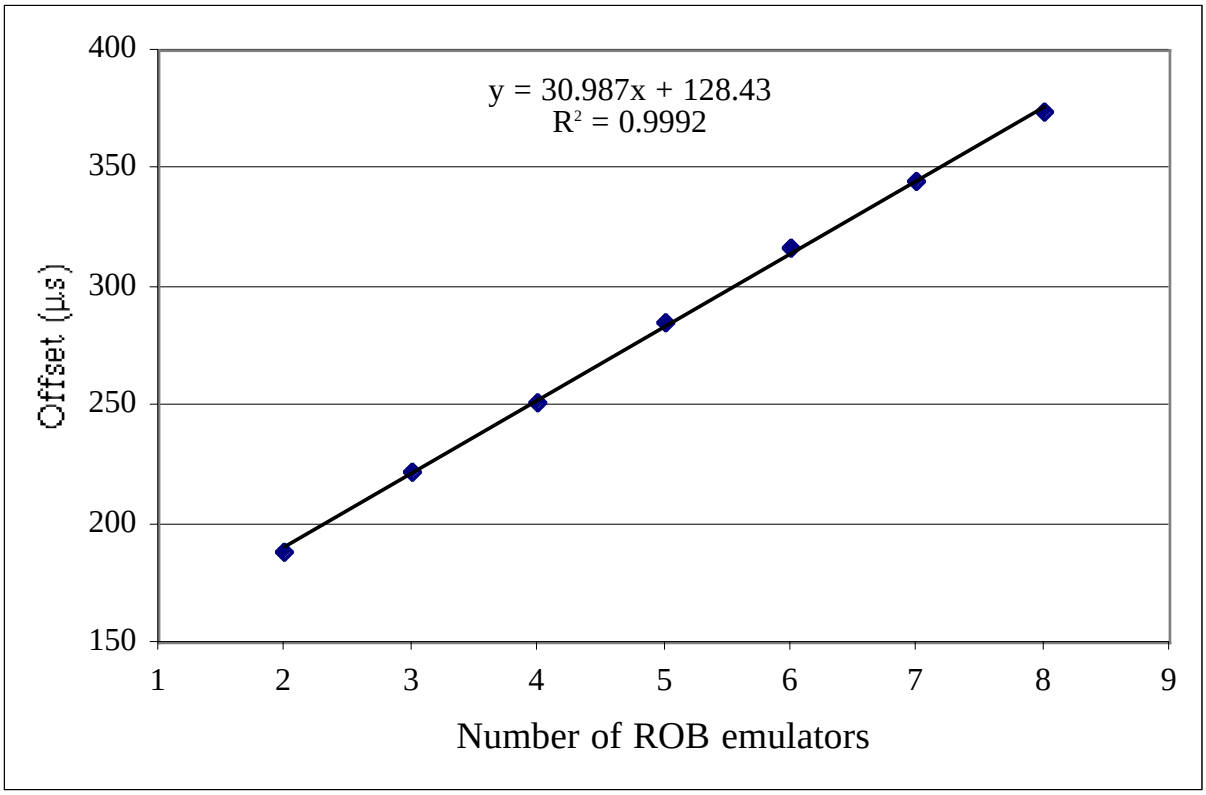

Figure 3. Offset found in Figure 2 as function of the number of active ROB emulators

Figure 3 shows the offsets of Figure 2 as a function of the number of active ROB emulators. The slope of the line fitted represents the extra processing time required for handling an additional ROB emulator for empty event fragments (i.e. only consisting of a header). It therefore consists of two components : the extra time needed to generate an additional RoI request and the extra time needed to handle an additional (empty) event fragment.

The measurement results on latency for the configuration with one supervisor, 2 - 8 ROB emulators, one steering processor and one outstanding event (in stead of four outstanding events) also provide the possibility to extract relevant parameters. As there can be only one outstanding event in this configuration the latency can be calculated from the times spent in the processors and in the network and switch. Table 5 contains an overview of the processing and transport times contributing to the latency. For data transfer via the network links also network protocol data should be taken into account. These give a fixed contribution to the time needed for transferring a message across a link. The protocol requires a minimum data size of 46 Bytes and 26 additional Bytes. Additionally there is a minimum gap of 12 Bytes (0.96 µs for the 100 Mbit links used) between successive frames transmitted via the same link. For the message sizes used see Section 2.6.

In step 13 the processing time spent is equal to N times the process time for inputting one event fragment. However, the latency only depends on the time spent for the last event fragment, if input of all previous fragments overlaps with the transfer of event fragment data via the input link and if the time needed for transferring a fragment via the link is longer than the processor time spent for input of the data (see Figure 4). For the shortest event fragments this condition could be invalid. A linear relation between the latency, with transfer times for protocol overhead and inter-frame gaps subtracted, and the event fragment size (including the 32 Byte header) should be observed if the condition is met. If not, a systematic deviation should be observed. In

Figure 6 does not indicate a significant deviation from a linear relation. The time needed for transferring a single Byte has been taken to be 0.08 µs, the nominal value for 100 Mb/s Ethernet. In Figure 7 the differences between measured values and values calculated from the lines fitted and described by the relations in Figure 6 are plotted. There is no clear systematic deviation.

The processing of step 5 associated with receiving a steering request and sending the first RoI request (which determines when the first event data arrives in the processor) may or may not depend on the number of ROB emulators N. The time of sending the second and following requests does not affect the latency, if the time interval between sending requests to different ROB emulators is shorter than the time needed for an event fragment to be transported via the link between switch and steering processor (see Figure 4). This is due to the fact that event fragments from the different ROB emulators have to pass one after another via this link. The behavior of the latency, with transfer times for protocol overhead and inter-frame gaps subtracted, plotted as function of N would then for 64 Byte event fragments differ from the behavior of 512 or 1024 Byte fragments if this condition is not met. Again Figure 7 shows that the behavior for 64 Byte event fragments for different N does not clearly differ from the behavior for the other event fragment sizes, so it is assumed that the condition mentioned is met.

Step	Time spent in step	Action	Latency depends on
1	Supervisor	output steering request	fixed time
2	Network link to switch	transmit 56 + 26 Bytes	link speed
3	Switch	delay 6 µs	fixed time
4	Network link to steering processor	transmit 56 + 26 Bytes	link speed
5	Steering processor	input steering request, generate and output RoI requests	N and / or fixed time
6	Network link to switch	transmit 44 + 2 + 26 Bytes	link speed, inter-frame gap
7	Switch	delay 6 µs	fixed time
8	Network link to ROB emulator	transmit 44 + 2 + 26 Bytes	link speed
9	ROB emulator	output E + 32 Bytes after 78.7 + 0.1173 * E µs	fixed time and E
10	Network link to switch	transmit E + 32 + 26 Bytes	link speed and E + 32 + 26
11	Switch	delay 6 µs	fixed time
12	Network link to steering processor	transmit N * (E + 32 + 26) Bytes	link speed, N * (E + 32 + 26), inter-frame gap
13	Steering processor	input N fragments, merge N * E Bytes, process and output decision	fixed time, N and E
14	Network link to switch	transmit 28 + 18 + 26 Bytes	link speed
15	Switch	delay 6 µs	fixed time
16	Network link to supervisor	transmit 28 + 18 + 26 Bytes	link speed
17	Supervisor	input and set time stamp	fixed time

Table 5. Processing and data transport steps contributing to the latency. N is the number of ROB emulators, E is the event fragment size (in Bytes).

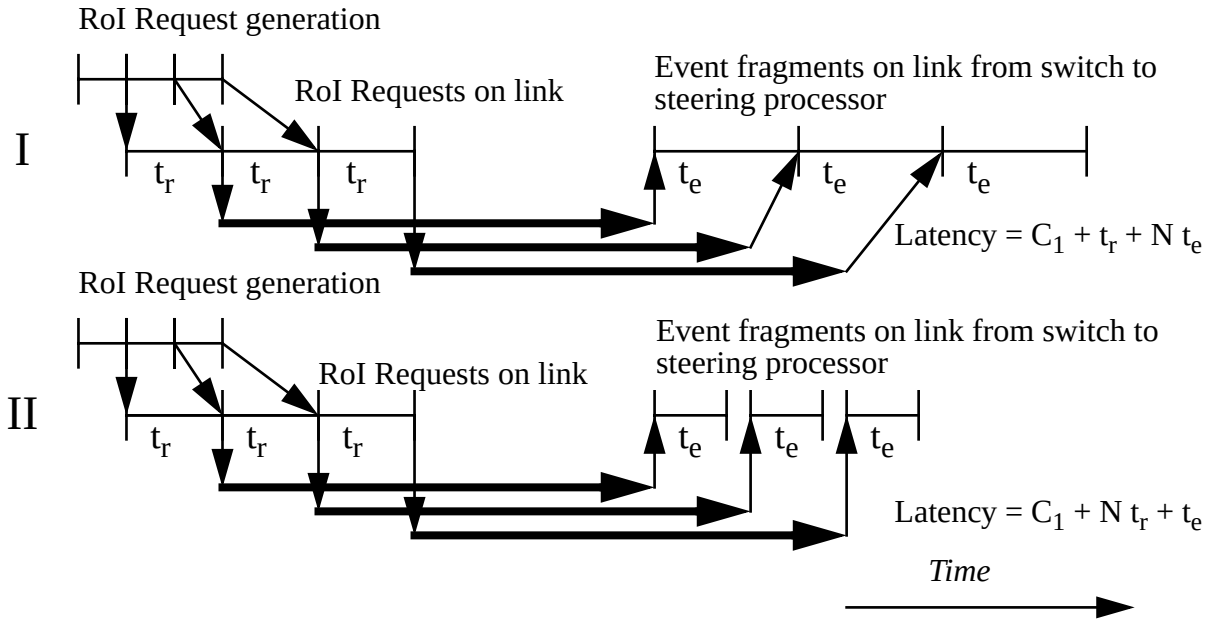

Figure 4. Diagram showing how the relative size of the link transfer times of a RoI request (t_r) and of an event fragment (t_e) affect the latency. For this figure it is assumed that t_r is longer than the time required for generating a RoI request. The thick arrows indicate the time needed for transferring a request via the switch and the network link to the ROB emulator, the processing time in the ROB emulator and the time needed for transferring the response of the ROB emulator via the output link and via the switch to the link to the steering processor. $t_r < t_e$ (I) or $t_r > t_e$ (II) result in a different dependence of the latency on N, the number of ROB emulators. The RoI request generation time should be used in the formulas for the latency in stead of t_r. if the RoI request generation time is larger than t_r..

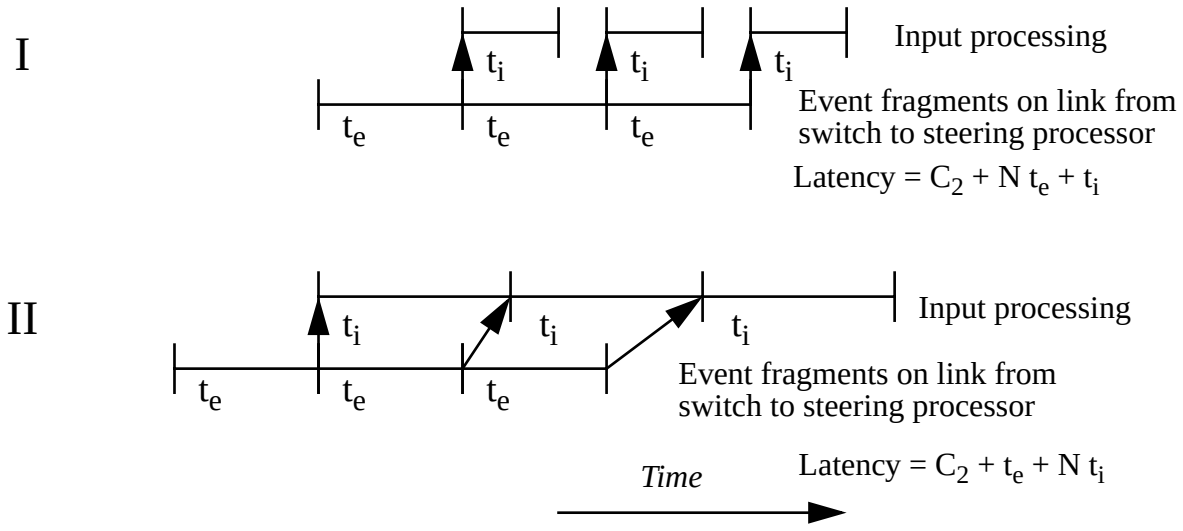

Figure 5. Diagram showing how relative sizes of the link transfer time of an event fragment (t_e) and of the input processing time of the steering processor (t_i) affect the latency. $t_i < t_e$ (I) or $t_i > t_e$ (II) result in a different dependence of the latency on N, the number of ROB emulators

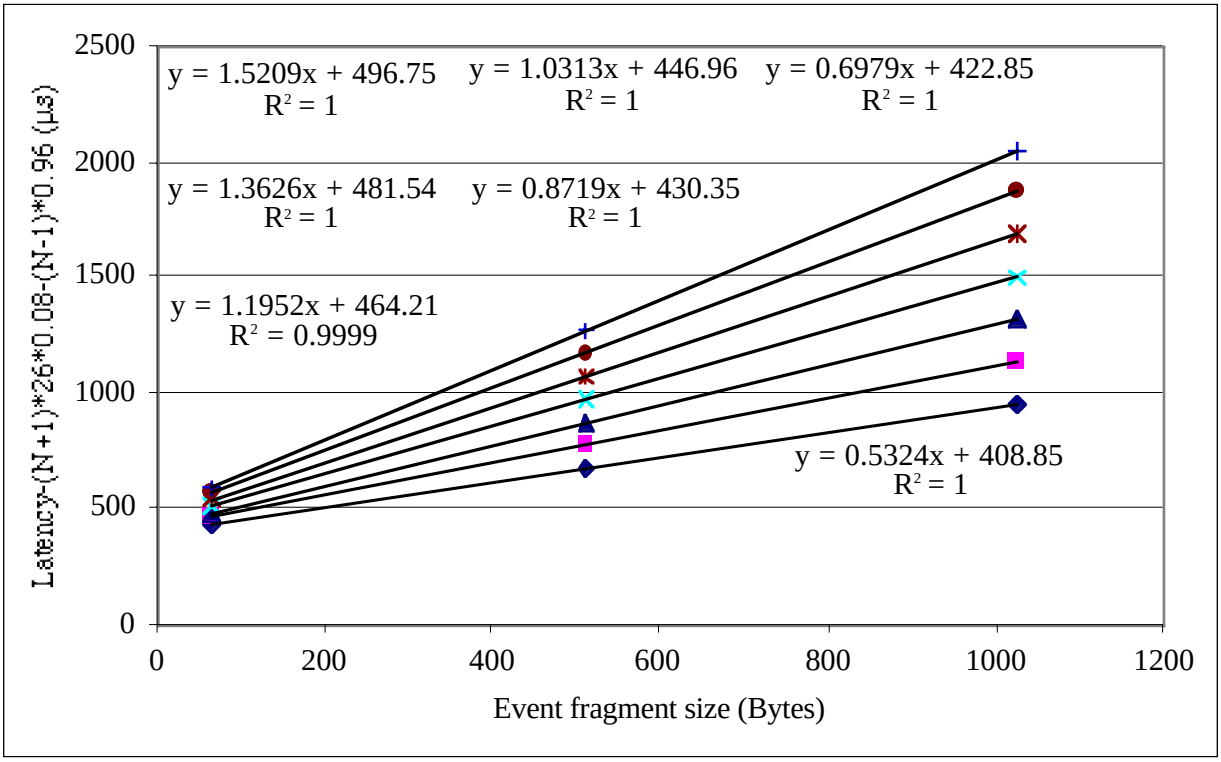

Figure 6. Latency, with transfer times for protocol overhead and inter-frame gaps subtracted (N is equal to the number of ROB emulators), as function of message size (32 Byte header included) for 2 (lowest line) to 8 (highest line) ROB emulators for the configuration with one supervisor, one steering processor and at maximum one outstanding event. The formulas describe the lines fitted to the measurement results.

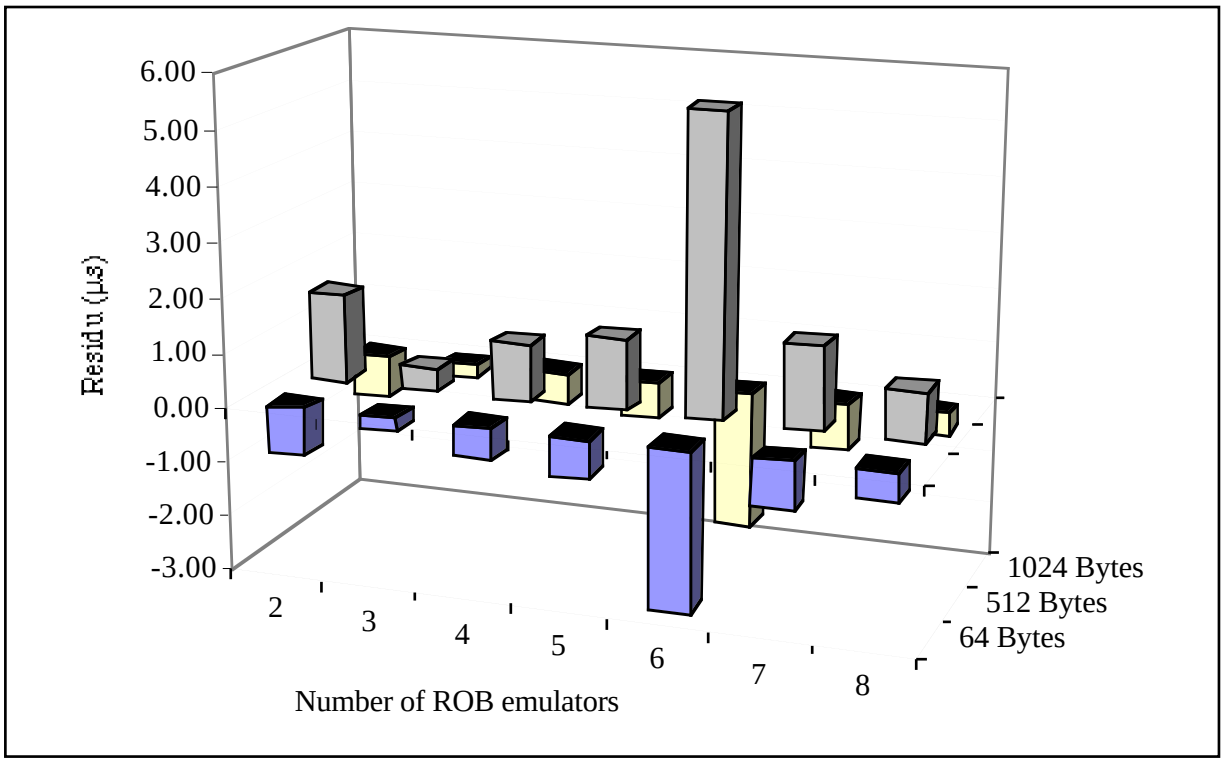

Figure 7. Differences, for the latency with transfer times for protocol overhead and inter-frame gaps subtracted, between values calculated from the equations of the fitted lines in Figure 6 and values measured.

From the discussion above and from Table 5 it is expected that the dependence on the number of ROB emulators N and on the event fragment size E of the measured latency is completely determined by :

1. the time needed for moving the data in the ROB emulator (E * 0.1173),
2. the time needed for merging the data in the steering processor (N * E * 0.088),
3. the time needed to transport the event data across the network links ((N + 1) * (E + 32 + 26) * 0.08)
4. N - 1 times the inter-frame gap ((N - 1) * 0.96).

This expectation can be checked by subtracting the contributions of these items from the measured latencies and plotting the results. Figure 8 show that these do not depend on the fragment size. However, there is a clear dependence on N. The slope of the line fitted in Figure 9 indicates that for each additional ROB emulator the latency increases with 11.845 µs. This can only be due to steps 5 and 13 in Table 5, i.e. from an increase in the contribution to the latency by the processing in the steering processor. From the analysis of the event rate for 100 % utilization of a steering processor it was found that the processing time spent per event by the steering processor increases with 30.987 µs for each additional ROB emulator (see Figure 3) if the event fragment size is kept the same. From Figure 7 and Figure 4 it can be concluded that the processing time spent on input of a single event fragment can not be longer than the time to transfer a 96 Byte fragment (64 + 32 Byte header) via a network link. The transfer time is (64 + 32 + 26) * 0.08 + 0.96 = 10.72 µs. The same is true for the processing time spent for each additional RoI request in parallel with outputting of the RoI requests (see Figure 4), i.e. the maximum time spent is also 10.72 µs. This leaves 30.987 - 21.44 = 9.55 µs spent per ROB emulator, which can only be spent before the first RoI request is output (otherwise Figure 7 would show a dependence on event fragment size) and / or after input of all event fragments. FURTHER DICUSSION BASED ON MEASUREMENTS AND COMPUTER MODELS TO BE INCLUDED HERE.

In Table 6 an overview of the parameters is presented, together with their names as used in the Ptolemy and Simdaq models. The time needed for feature extraction have been taken from [4], the merging speed is taken from the analysis in Section 2.3 and is identical to the speed specified in [4]. The times depending on N have been taken from the analysis in this section. TABLE NEEDS FURTHER WORK.

The parameters in Table 6 apply for configurations with 2 - 8 ROB emulators. To check the validity for the configurations with a single ROB emulator again measurement results can be used for which 100 % utilization of the steering processor can be assumed. Unfortunately measurements have only been done for two different event fragment sizes : 64 Bytes and 1024 Bytes. For this check the measurement results shown in Table 7 (configuration with a 400 MHz steering processor) have been used to obtain the parameters presented in Table 8. The offset of 139.5 µs should be equal to the sum of all times not depending on the event fragment size specified in Table 6. This sum is equal to 128.43 + 30.987 µs, so there is 19.9 µs missing for the configuration with a single ROB emulator. EXPLAIN FROM REFERENCE SOFTWARE ?

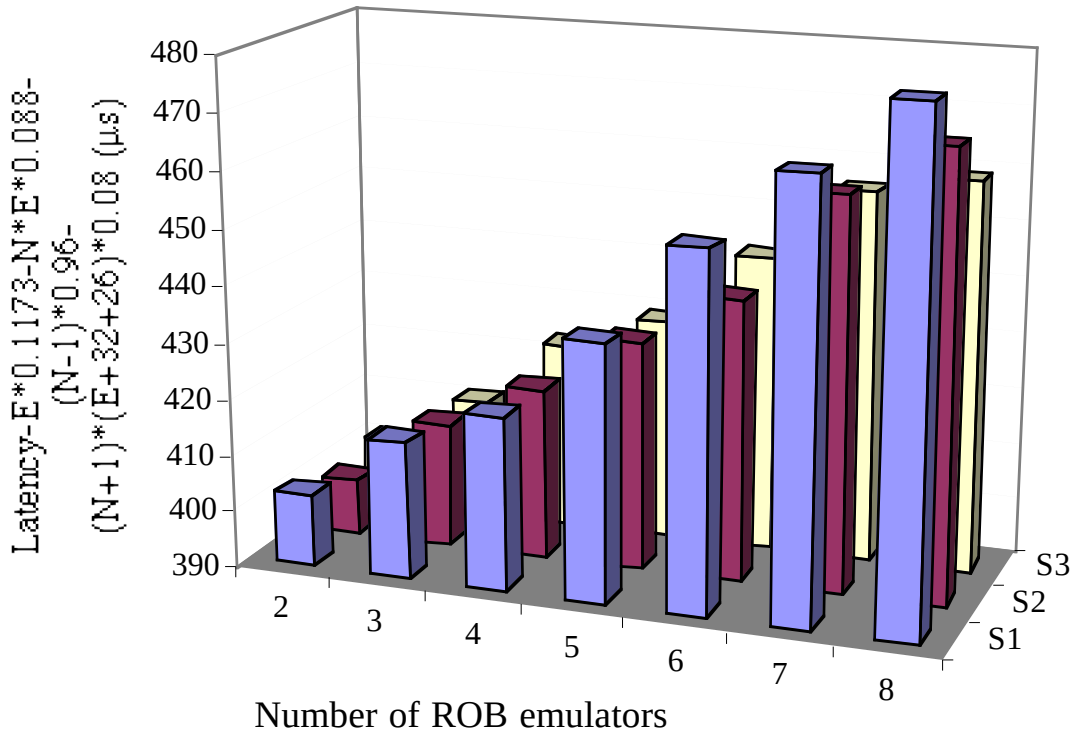

Figure 8. Latency, with times for moving data in the ROB emulator, for merging data in the steering processor, for transferring inter-frame gaps and for transferring the event data via the network links (taking into account that transfers between ROB emulators and switch proceed in parallel) subtracted, as function of the number of ROB emulators and of event fragment size.

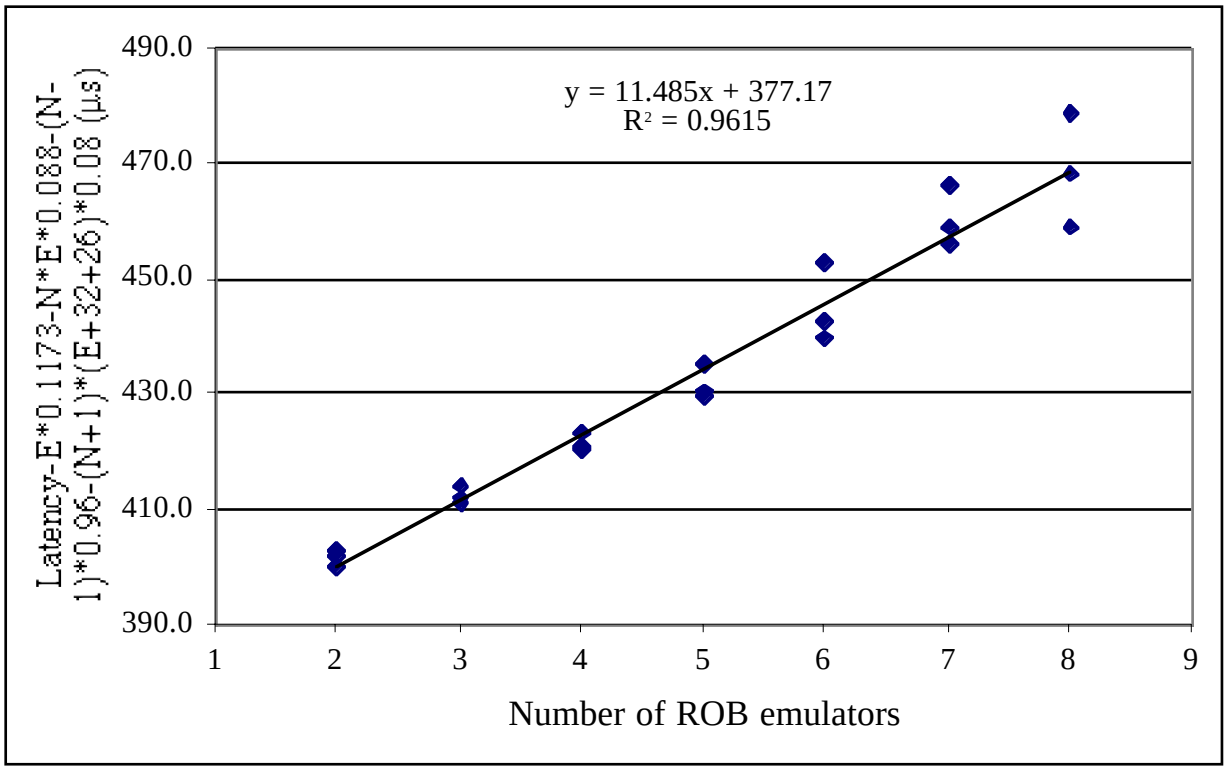

Figure 9. The quantity plotted in Figure 8 as function of the number of ROB emulators only, with a line fitted to the data points.

	Receive steering request	Handling of steering request	Generation of RoI data requests	Receive event fragment	FEX, global + generation of decision
300 MHz	40.0	45.9	40.0+N*14.3[a], N*6.4[a]	N * 14.3	45.3+N*6.4+ N*E*0.01173
400 MHz	30.0	34.4	30.0+N*10.7[a.], N*4.8[a.]	N *10.7	34.0+N*4.8+ N*E*0.0880
Ptolemy parameter	receive_event_ time	ROBroiData Request_ prepare_time, subsequentRoiRequest		receive_roi_ time	fex_process_ time, steerin-gRoi ReceiveFloat
Simdaq parameter	RcvSelroir Processtime	RoIInput ProcessTime	RoIFormulate ProcessTime, RoIFormulate PerRequest ProcessTime	RcvSelevent Fragment ProcessTime	ProcTime Average, MergeSpeed

a. Only for N ≥ 2, see text

Table 6. Parameters for the steering processor, N is the number of ROB emulators, E is the event fragment size. These parameters are to be used for the configurations with 2 - 8 ROB emulators NEEDS FURTHER WORK

Configuration	Event size	Measured rate (Hz)	Inverse rate (µs)
c-tbed-s400-8-3-64	64	6875	145.5
c-tbed-s400-8-3-1024	1024	4360	229.4

Table 7. Measured rate for the configurations indicated (see Table 1) with a single ROB emulator.

	Calculated	Slope fixed
Slope	0.0874	0.088
Offset (µs)	139.9	139.5

Table 8. Calculated parameters for the inverse rate as function of event size (including 32 Byte header) for the configurations specified in Table 7.

The missing 19.9 µs is close to the 21.4 µs, the maximum time available for generating a RoI request or input of an event fragment in parallel with transfer across a link, -> CONCLUSION ? The same analysis was done for the configuration with a 300 MHz steering processor, see Table 9 and Table 10. The offset of 184.7 µs is smaller than the expected value of (128.43 + 30.987)*4/3 = 212.6 µs, so there is 27.9 µs missing, which is close to 4/3 * 19.9 µs..

Configuration	Event fragment size	Measured rate (Hz)	Inverse rate (µs)
c-tbed-s300-5-2-64	64	5231	191.2
c-tbed-s300-8-3-1024	1024	3270	305.8

Table 9. Measured rate for the configurations indicated (see Table 1) with a single ROB emulator.

	Calculated	Slope fixed
Slope	0.1194	0.1173
Offset (µs)	183.5	184.7

Table 10. Calculated parameters for the inverse rate as function of event size (including 32 Byte header) for the configurations specified in Table 9.

2.5 Network connections and the switch

Fast Ethernet has been used in the testbed. The switch is a store and forward switch without contention for the traffic generated in the testbed. For the parameters, taken from [2] see Table 11.

Transport time per Byte for one link (µs)	Delay in switch per message (µs)
0.080	6.0

Table 11. Parameters for network connections and switch.

2.6 Messages

Table 12 contains an overview of the message types and associated sizes. The message sizes are different from those specified in [2] and have been obtained from [4]. For each message 26 Bytes of protocol overhead has to be added if link transfer times are calculated, while messages originally shorter than 46 Bytes are padded to that length. Also the inter-frame gap of 0.96 µs has to be taken into account if messages are sent immediately after each other via the same link.

Source	Destination	Message	Size (Bytes)
Supervisor	Steering processor	SteeringRequest	56 + 26
Steering processor	ROB emulator	RoIDataRequest	44 + 2 + 26
ROB emulator	Steering processor	RoIData	32 + fragment size + 26
Steering processor	Supervisor	Decision	28 + 18 + 26

Table 12. Overview of message types and associated sizes.

3 Comparison of measurement results to results computed from the parameters

3.1 Configuration with one supervisor, one steering processor, 2 - 8 ROB emulators and one outstanding event

Table 14 contains a overview of measured and calculated latencies for the configuration with one supervisor, one steering processor, 2 - 8 ROB emulators and one outstanding event. Calculated values have been obtained by adding all contributions to the latency as listed in Table 13. The contribution to the latency by the steering processor consists of a fixed time (128.43 µs, see Figure 3), a contribution due to merging the data (N * E * 0.088 µs), a contribution dependent on N (N * 11.485 µs, see Figure 9) and another fixed time. The last contribution is due to the fact that only the time spent by the processor before the output of the first RoI request and after arrival of the last event fragment contributes to the latency (see Figure 4). Its value is equal to the slope of the line in Figure 3, the part of the total processing time increasing with N, minus the value of the slope of the line in Figure 9, the part of the processing time increasing with N that never is spent in parallel with transport of data across the network.

Step	Time spent in step by	Action	Contribution to latency (µs)
1	Supervisor	output steering request	48.25
2	Network link to switch	transmit 56 + 26 Bytes	6.56
3	Switch	delay 6 µs	6.00
4	Network link to steering processor	transmit 56 Bytes	6.56
5	Steering processor	input steering request, generate and output RoI requests	Included in formula specified for step 13
6	Network link to switch	transmit 44 + 2 + 26 Bytes	5.76
7	Switch	delay 6 µs	6.00
8	Network link to ROB emulator	transmit 44 + 2 + 26 Bytes	5.76
9	ROB emulator	output E Bytes after 78.7 + 0.1173*E µs	82.45 + E * 0.1173
10	Network link to switch	transmit E Bytes	(E + 32 + 26) * 0.08
11	Switch	delay 6 µs	6.00
12	Network link to steering processor	transmit N * E Bytes	N * (E + 32 + 26) * 0.08 + (N - 1) * 0.96
13	Steering processor	input N * E Bytes, merge N * E Bytes, process and output decision	128.43 + (N - 1) * 11.485 + N * E * 0.088 + 30.99, includes contribution from step 5
14	Network link to switch	transmit 28 + 18 + 26 Bytes	5.76
15	Switch	delay 6 µs	6.00
16	Network link to supervisor	transmit 28 + 18 + 26 Bytes	5.76
17	Supervisor	input and set time stamp	48.00

Table 13. Processing and data transport steps contributing to the latency. N is the number of ROB emulators, E is the event fragment size (in Bytes).

Event size (Bytes)	Number of ROB emulators	Measured latency (µs)	Calculated latency (µs)	Error, no correction (µs)	Relative error with correction (%)	Error with correction (µs)
64	2	451	458.8	7.8	-0.9	-4.2
	3	478	486.6	8.6	-0.7	-3.4
	4	500	514.4	14.4	0.5	2.4
	5	530	542.3	12.3	0.1	0.3
	6	563	570.1	7.1	-0.9	-4.9
	7	592	598.0	6.0	-1.0	-6.0
	8	620	625.8	5.8	-1.0	-6.2
512	2	687	697.7	10.7	-0.2	-1.3
	3	790	800.8	10.8	-0.2	-1.2
	4	889	903.9	14.9	0.3	2.9
	5	990	1007.0	17.0	0.5	5.0
	6	1090	1110.1	20.1	0.7	8.1
	7	1200	1213.2	13.2	0.1	1.2
	8	1300	1316.3	16.3	0.3	4.3
1024	2	962	970.7	8.7	-0.3	-3.3
	3	1148	1159.9	11.9	0.0	-0.1
	4	1337	1349.0	12.0	0.0	0.0
	5	1520	1538.1	18.1	0.4	6.1
	6	1710	1727.2	17.2	0.3	5.2
	7	1900	1916.3	16.3	0.2	4.3
	8	2080	2105.4	25.4	0.6	13.4

Table 14. Measured and calculated latencies for the configuration with one supervisor, one steering processor, 2 - 8 ROB emulators and at maximum one outstanding event.

The calculated values of the latency are on average 13.1 µs larger than the measured values. This number is consistent with the correction of 12.0 µs required to remove a systematic deviation between measured values and values for the latency calculated from the rate,see the next section. If this correction is subtracted from the calculated values small errors result (at maximum 1 %). The systematic difference is probably due to the times at which the time measurements in the supervisor are made : the first measurement will be done some time after the start of the processing associated with outputting a steering request and the second measurement will be done some time before the end of the processing associated with receiving a decision from the steering processor.

In Table 15 a comparison between measured and calculated rates is presented. For the calculation of the rates from the measured latencies the latter were increased with 12.0 µs, and the result was inverted. Again only small errors are found, the largest being equal to 1.3 %.

Event size (Bytes)	Number of ROB emulators	Measured rate (Hz)	Rate from corrected measured latency (Hz)	Rate from calculation (Hz)	Relative error in calculation (%)	Error in calculated inverse rate (µs)
64	2	2155	2159.8	2179.7	1.1	-5.3
	3	2039	2040.8	2055.0	0.8	-3.8
	4	1948	1953.1	1943.8	-0.2	1.1
	5	1833	1845.0	1844.1	0.6	-3.3
	6	1732	1739.1	1754.0	1.3	-7.2
	7	1651	1655.6	1672.4	1.3	-7.7
	8	1584	1582.3	1598.0	0.9	-5.5
512	2	1429	1430.6	1433.3	0.3	-2.1
	3	1246	1246.9	1248.8	0.2	-1.8
	4	1109	1109.9	1106.3	-0.2	2.2
	5	992	998.0	993.1	0.1	-1.1
	6	901	907.4	900.8	0.0	0.2
	7	825	825.1	824.3	-0.1	1.1
	8	762	762.2	759.7	-0.3	4.0
1024	2	1026	1026.7	1030.1	0.4	-3.9
	3	860	862.1	862.2	0.3	-2.9
	4	740	741.3	741.3	0.2	-2.4
	5	649	652.7	650.2	0.2	-2.7
	6	578	580.7	579.0	0.2	-2.9
	7	520	523.0	521.8	0.4	-6.8
	8	476	478.0	475.0	-0.2	4.6

Table 15. Measured and calculated rate for the configuration with one supervisor, one steering processor, 2- 8 ROB emulators and at maximum one outstanding event.

3.2 Configuration with one supervisor, one steering processor, 2 - 8 ROB emulators and 4 outstanding events

In Figure 10 measurement results for the latency for the configuration with 1 supervisor, 1 steering processor, 2 - 8 ROB emulators and 4 outstanding events are plotted as function of the number of ROB emulators. The average of the offsets of the lines is 525.8 µs. This is equal to 4.03 times 128.43 µs, the value of the part of the processing time of the steering processor which is independent of the number of ROB emulators and the event fragment size (see Figure 3). The average of the slopes of each line divided by the part of the processing time depending on the number of ROB emulators and the event fragment size (30.987 + E * 0.088) is equal to 3.99. It therefore seems to be reasonable to assume that the latency is equal to 4 times the processing time spent by the steering processor per event and that the rate is equal to 4 times the inverse of the latency. The factor of 4 is caused by the fact that the maximum number of outstanding events is 4, together with the 100 % utilization of the steering processor. It can be understood by considering that the supervisor, after the system has stabilized after starting up, only sends a new steering request if a decision is received. This is due to the steering requests being sent faster than decisions arrive, in any case during the start-up phase and hence always the maximum number of

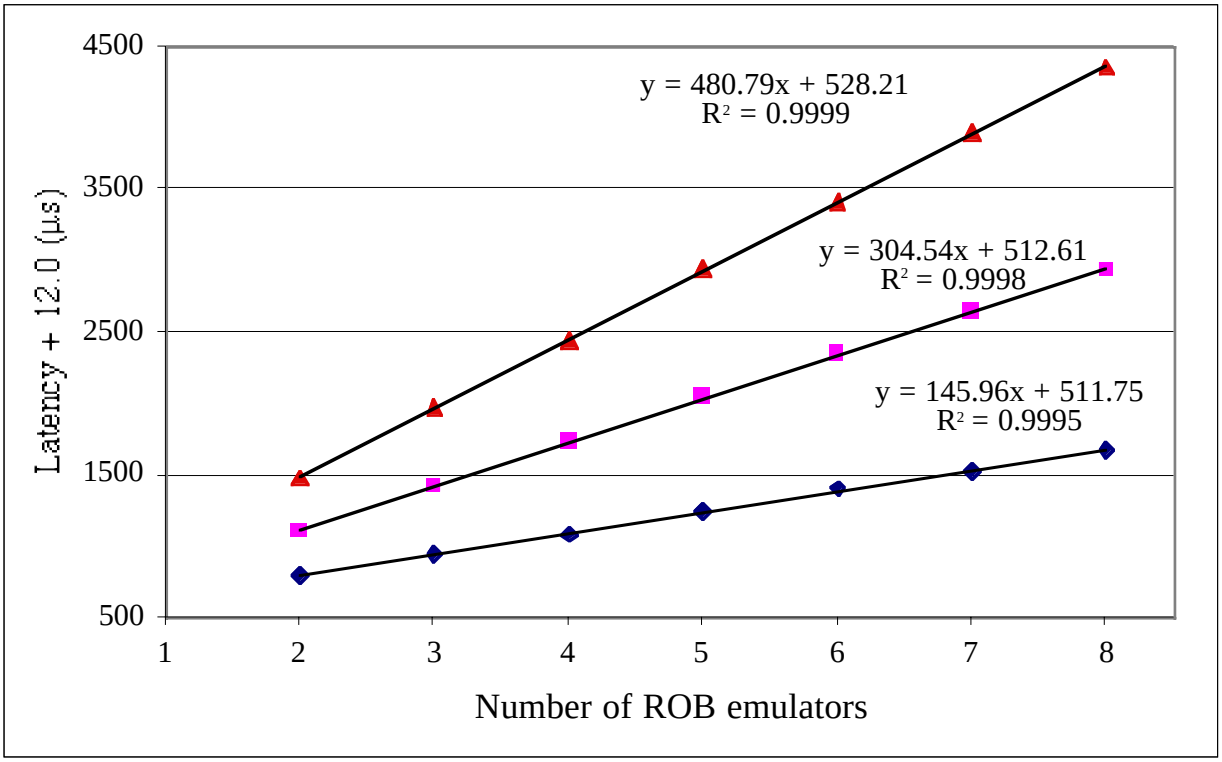

Figure 10. Latency for 64, 512 and 1024 Byte fragments for the configuration with one supervisor, one steering processor and at maximum four outstanding events. The formulas describe the lines fitted to the measurement results.

steering requests will have been sent. In other words, the output of a steering request is synchronized with the arrival of a decision of an event of which the id is 4 lower than the event for which the steering request is sent. Event decisions arrive with a rate which is determined by the inverse of the processing time spent per event by the steering processor, as its utilization is 100 %. Hence the latency, i.e. the interval between sending a steering request and the arrival of the decision in the supervisor, will be equal to 4 times the inverse rate measured and hence to 4 times the processing time spent per event by the steering processor.

In the configurations with 2 - 8 ROB emulators and more than one steering processor it may occur that the utilization of the ROB emulators or of the supervisor becomes 100 %. In that case the same reasoning applies, with the processing time spent per event of the component which is utilized at 100 %. The rates and latencies for these configurations and the configuration considered in this section hence are determined completely by the component with 100 % utilization and the number of outstanding events. However, one correction needs to be applied to the measured latencies : these have to be increased with 12.0 µs. This number follows from the measurements and can be understood by considering that the time stamp at the start of outputting a new steering request will be made some time after the time stamp made on arrival of a decision. The correction of 12.0 µs to the measured latency is required to remove a systematic deviation between measured values and values for the latency calculated from the rate. The value of the correction has been chosen such that the average of the deviations between the corrected measured and calculated values of the latency for the configurations with a single ROB emulator is 0. If the correction is determined from the results discussed in this section it would have to be 14.5 µs, the value of 12.0 µs is a better choice if all errors are taken into account (see Figure 12 and Figure 13 in Section 3.7). The correction is probably required due to incorrect values for the processing times and / or by an inaccurate model of the different actions by the supervisor.

Event size (Bytes)	Number of ROB emulators	Measured rate (Hz)	Calculated rate (Hz)	Error in calculated rate (%)	Error in inverse rate (µs)	Measured latency + 12.0 (µs)	Calculated latency (Hz)	Error in calculated latency (%)	Error in calculated latency (µs)
64	2	4977	4958.6	-0.4	0.7	801	806.7	0.7	5.7
	3	4203	4196.6	-0.2	0.4	948	953.1	0.5	5.1
	4	3650	3637.6	-0.3	0.9	1091	1099.6	0.8	8.6
	5	3190	3210.0	0.6	-2.0	1249	1246.1	-0.2	-2.9
	6	2854	2872.4	0.6	-2.2	1399	1392.6	-0.5	-6.4
	7	2609	2599.0	-0.4	1.5	1531	1539.1	0.5	8.1
	8	2390	2373.1	-0.7	3.0	1672	1685.5	0.8	13.5
512	2	3583	3564.9	-0.5	1.4	1112	1122.1	0.9	10.1
	3	2788	2804.6	0.6	-2.1	1430	1426.2	-0.3	-3.8
	4	2307	2311.6	0.2	-0.9	1729	1730.4	0.1	1.4
	5	1952	1966.0	0.7	-3.7	2043	2034.6	-0.4	-8.4
	6	1699	1710.3	0.7	-3.9	2351	2338.8	-0.5	-12.2
	7	1511	1513.5	0.2	-1.1	2645	2642.9	-0.1	-2.1
	8	1360	1357.3	-0.2	1.5	2937	2947.1	0.3	10.1
1024	2	2684	2698.1	0.5	-2.0	1486	1482.5	-0.2	-3.5
	3	2028	2033.6	0.3	-1.4	1967	1966.9	0.0	-0.1
	4	1631	1631.8	0.0	-0.3	2444	2451.3	0.3	7.3
	5	1356	1362.5	0.5	-3.5	2947	2935.7	-0.4	-11.3
	6	1166	1169.6	0.3	-2.6	3424	3420.1	-0.1	-3.9
	7	1025	1024.5	-0.1	0.5	3897	3904.5	0.2	7.5
	8	915	911.4	-0.4	4.3	4360	4388.9	0.7	28.9

Table 16. Measured and calculated rates and latencies for the configuration with one supervisor, one steering processor, 2 - 8 ROB emulators and a maximum of 4 outstanding events.

3.3 Configuration with one supervisor, two steering processors, 2 - 8 ROB emulators and 8 outstanding events

For the configuration with one supervisor, two steering processors and 2 - 8 ROB emulators the rate was calculated by determining the rates supported by a ROB emulator, a supervisor and a steering processor, all running at 100 % utilization, and next taking the minimum of that rate. The latency was found from the calculated rate by multiplying its inverse with the maximum number of outstanding events, see the discussion in Section 3.2. A comparison between measured and calculated values is presented in Table 17. The maximum error is 2.2 %.

Event size (Bytes)	Number of ROB emulators	Measured rate (Hz)	Calculated rate (Hz)	Error in calculated rate (%)	Error in inverse rate (µs)	Measured latency + 12.0 (µs)	Calculated latency (Hz)	Error in calculated latency (%)	Error in calculated latency (µs)
64	2	8206	8203.4 S	0.0	0.0	972	975.2 S	0.3	3.2
	3	8156	8203.4 S	0.6	-0.7	979	975.2 S	-0.4	-3.8
	4		7275.2				1099.6		
	5	6399	6420.0	0.3	-0.5	1248	1246.1	-0.2	-1.9
	6	5772	5744.7	-0.5	0.8	1384	1392.6	0.6	8.6
	7	5297	5198.0	-1.9	3.6	1508	1539.1	2.1	31.1
	8	4845	4746.3	-2.0	4.3	1649	1685.5	2.2	36.5
512	2	7060	7019.6	-0.6	0.8	1131	1139.7	0.8	8.7
	3	5498	5609.2	2.0	-3.6	1454	1426.2	-1.9	-27.8
	4		4623.2				1730.4		
	5	3901	3932.0	0.8	-2.0	2048	2034.6	-0.7	-13.4
	6	3396	3420.6	0.7	-2.1	2354	2338.8	-0.6	-15.2
	7	3012	3027.0	0.5	-1.6	2654	2642.9	-0.4	-11.1
	8	2706	2714.5	0.3	-1.2	2949	2947.1	-0.1	-1.9
1024	2	4966	4937.9 R	-0.6	1.1	1609	1620.1 R	0.7	11.1
	3	3990	4067.3	1.9	-4.8	2002	1966.9	-1.8	-35.1
	4		3263.6				2451.3		
	5	2718	2725.1	0.3	-1.0	2941	2935.7	-0.2	-5.3
	6	2326	2339.1	0.6	-2.4	3437	3420.1	-0.5	-16.9
	7	2044	2048.9	0.2	-1.2	3911	3904.5	-0.2	-6.5
	8	1818	1822.8	0.3	-1.4	4393	4388.9	-0.1	-4.1

Table 17. Measured and calculated rates and latencies for the configuration with one supervisor, two steering processors, 2 - 8 ROB emulators and a maximum of 8 outstanding events. “S” and “R” indicate that the supervisor, respectively the ROB emulator, was determining the rate.

3.4 Configuration with 2 supervisors, 4 steering processors, 2 - 8 ROB emulators and 8 outstanding events

For the configuration with two supervisors, 4 steering processors and 2 - 8 ROB emulators the rate was calculated in the same way as in the previous section. The latency was again found from the calculated rate by multiplying its inverse with the total maximum number of outstanding events (i.e. 16), see the discussion in Section 3.2. A comparison between measured and calculated values is presented in Table 18. The largest errors are now 2 - 4 % in the latency, found for 64 Byte event fragments. Errors in the rate are at maximum 1.5 %.

Event size (Bytes)	Number of ROB emulators	Measured rate (Hz)	Calculated rate (Hz)	Error in calculated rate (%)	Error in inverse rate (µs)	Measured latency + 12.0 (µs)	Calculated latency (Hz)	Error in calculated latency (%)	Error in calculated latency (µs)
64	4	11159	11122.6 R	-0.3	0.3	1432	1438.5 R	0.5	6.5
	5	11070	11122.6 R	0.5	-0.4	1445	1438.5 R	-0.4	-6.5
	6	11066	11122.6 R	0.5	-0.5	1384	1438.5 R	3.9	54.5
	7	10554	10396.0	-1.5	1.4	1508	1539.1	2.1	31.1
	8	9483	9492.6	0.1	-0.1	1649	1685.5	2.2	36.5
512	4	6987	7019.6 R	0.5	-0.7	2289	2279.3 R	-0.4	-9.7
	5	6963	7019.6 R	0.8	-1.2	2296	2279.3 R	-0.7	-16.7
	6	6773	6841.3	1.0	-1.5	2364	2338.8	-1.1	-25.2
	7	6003	6053.9	0.8	-1.4	2665	2642.9	-0.8	-22.1
	8	5386	5429.1	0.8	-1.5	2968	2947.1	-0.7	-20.9
1024	4	4945	4937.9 R	-0.1	0.3	3235	3240.2 R	0.2	5.2
	5	4912	4937.9 R	0.5	-1.1	3257	3240.2 R	-0.5	-16.8
	6	4610	4678.2	1.5	-3.2	3470	3420.1	-1.4	-49.9
	7	4067	4097.8	0.8	-1.9	3934	3904.5	-0.8	-29.5
	8	3634	3645.6	0.3	-0.9	4402	4388.9	-0.3	-13.1

Table 18. Measured and calculated rates and latencies for the configuration with 2 supervisors, 4 steering processors, 2- 8 ROB emulators and a maximum of 8 outstanding events per supervisor."R" indicate that the ROB emulators were determining the rate.

3.5 Configuration with one supervisor, one 300 MHz steering processor and one ROB emulator

For the configuration with one supervisor, one 300 MHz steering processors and 2 - 8 ROB emulators the rate was calculated assuming 100 % utilization of the steering processor. The latency was found from the *measured* rate by multiplying its inverse with the total maximum number of outstanding events, see the discussion in Section 3.2. A comparison between measured and calculated values is presented in Table 19. For the lowest values of the number of outstanding events the steering processor as well as the ROB emulator and the supervisor are not fully utilized, so the rate cannot be predicted, as is seen from the large errors. The errors are again small (below 1 %) for the highest values of the number of outstanding events. The latencies calculated from the measured rate are in very good agreement with the latencies measured, if the measured values are increased with 12.0 µs.

Workers	Event size (Bytes)	Number of outstanding events	Measured rate (Hz)	Calculated rate for 100 % utilization of steering processor (Hz)	Deviation of calculated from measured rate (%)	Deviation of calculated from measured inverse rate (µs)	Measured latency + 12.0 (µs)	Number of outstanding events * inverse rate (µs)	Deviation of calculated from measured latency (%)	Deviation of calculated from measured latency (µs)
2	64	2	4557	5203	14.2	-27.2	440	438.9	-0.3	-1.1
		3	4861	5203	7.0	-13.5	619	617.2	-0.3	-1.8
		4	5254	5203	-1.0	1.9	762	761.3	-0.1	-0.7
		5	5231	5203	-0.5	1.0	957	955.8	-0.1	-1.2
	1024	2	2260	3281	45.2	-137.7	886	885.0	-0.1	-1.0
		3	2990	3281	9.7	-29.6	1004	1003.3	-0.1	-0.7
		4	3139	3281	4.5	-13.8	1273	1274.3	0.1	1.3
3	64	3	4871	5203	6.8	-13.1	617	615.9	-0.2	-1.1
		4	4855	5203	7.2	-13.8	824	823.9	0.0	-0.1
		5	4877	5203	6.7	-12.8	1026	1025.2	-0.1	-0.8
		6	5232	5203	-0.6	1.1	1147	1146.8	0.0	-0.2
		7	5219	5203	-0.3	0.6	1342	1341.3	-0.1	-0.7
	1024	3	2729	3281	20.2	-61.6	1100	1099.3	-0.1	-0.7
		4	3126	3281	4.9	-15.1	1278	1279.6	0.1	1.6
		5	3127	3281	4.9	-15.0	1593	1599.0	0.4	6.0
		6	3270	3281	0.3	-1.0	1833	1834.9	0.1	1.9

Table 19. Measured and - for 100 % utilization of the steering processor - calculated rates and latencies or the configuration with one supervisor, one steering processor (300 MHz), one ROB emulator, 2 or 3 threads and a variable number of outstanding events per supervisor.

3.6 Configuration with 1 supervisors, 1 400 MHz steering processor and 1 ROB emulator

For the configuration with one supervisor, one 400 MHz steering processor and 2 - 8 ROB emulators the rate was calculated, as in the previous section, assuming 100 % utilization of the steering processor. The latency was found from the *measured* rate by multiplying its inverse with the total maximum number of outstanding events, see the discussion in Section 3.2. A comparison between measured and calculated values is presented in Table 19. For the lowest values of the number of outstanding events the steering processor as well as the ROB emulator and the supervisor are not fully utilized, so the rate cannot be predicted, as is seen from the large errors. The errors are again small (below 1 %) for the highest values of the number of outstanding events. There is one exception : the result for 3 threads, 7 outstanding events and 1024 Byte event fragments. The rate measured is 4630 Hz, while the rate for 6 and 8 outstanding events is 4365 and 4360 Hz respectively. It seems likely that the experimental value has been incorrectly typed and that it should be 4360 Hz. The latencies calculated from the measured rate are in very good agreement with the latencies measured, if the measured values are increased with 12.0 µs.

Workers	Event size (Bytes)	Number of outstanding events	Measured rate (Hz)	Calculated rate for 100 % utilization of steering processor (Hz)	Deviation of calculated from measured rate (%)	Deviation of calculated from measured inverse rate (µs)	Measured latency + 12.0 (µs)	Number of outstanding events * inverse rate (µs)	Deviation of calculated from measured latency (%)	Deviation of calculated from measured latency (µs)
2	64	2	5113	6890	34.8	-50.4	393	391.2	-0.5	-1.8
		3	6328	6890	8.9	-12.9	476	474.1	-0.4	-1.9
		4	6855	6890	0.5	-0.7	585	583.5	-0.3	-1.5
		5	6890	6890	0.0	0.0	727	725.7	-0.2	-1.3
		6	6885	6890	0.1	-0.1	872	871.5	-0.1	-0.5
	1024	3	3394	4355	28.3	-65.0	885	883.9	-0.1	-1.1
		4	3522	4355	23.7	-54.3	1134	1135.7	0.2	1.7
		5	3537	4355	23.1	-53.1	1413	1413.6	0.0	0.6
		6	3530	4355	23.4	-53.7	1700	1699.7	0.0	-0.3
		7	3530	4355	23.4	-53.7	1981	1983.0	0.1	2.0
3	64	4	6418	6890	7.4	-10.7	623	623.2	0.0	0.2
		5	6402	6890	7.6	-11.1	782	781.0	-0.1	-1.0
		6	6896	6890	-0.1	0.1	871	870.1	-0.1	-0.9
		7	6876	6890	0.2	-0.3	1019	1018.0	-0.1	-1.0
		8	6875	6890	0.2	-0.3	1165	1163.6	-0.1	-1.4
	1024	4	4046	4355	7.6	-17.5	988	988.6	0.1	0.6
		5	4198	4355	3.7	-8.6	1192	1191.0	-0.1	-1.0
		6	4365	4355	-0.2	0.5	1374	1374.6	0.0	0.6
		7	4630	4355	-5.9	13.6	1604	1511.9	-5.7	-92.1
		8	4360	4355	-0.1	0.3	1832	1834.9	0.2	2.9
4	64	5	6304	6890	9.3	-13.5	793	793.1	0.0	0.1
		6	6312	6890	9.2	-13.3	950	950.6	0.1	0.6
		7	6322	6890	9.0	-13.0	1108	1107.2	-0.1	-0.8
		8	6839	6890	0.7	-1.1	1171	1169.8	-0.1	-1.2
	1024	5	4152	4355	4.9	-11.2	1203	1204.2	0.1	1.2
		6	4168	4355	4.5	-10.3	1439	1439.5	0.0	0.5
		7	4169	4355	4.5	-10.3	1677	1679.1	0.1	2.1
		8	4342	4355	0.3	-0.7	1841	1842.5	0.1	1.5

Table 20. Measured and - for 100 % utilization of the steering processor - calculated rates and latencies or the configuration with one supervisor, one steering processor (300 MHz), one ROB emulator, 2, 3 or 4 threads and a variable number of outstanding events per supervisor.

3.7 Overview of deviations between calculated and measured values

In this section an overview of the errors in the calculated values of latencies and rates is presented.

Figure 11 contains a histogram of the relative errors in the rate as calculated for all configurations with two or more ROB emulators. For the configurations with one outstanding event the rates were calculated from the latency, for all other configurations from the processing time of component with the longest processing time running assuming 100 % utilization.

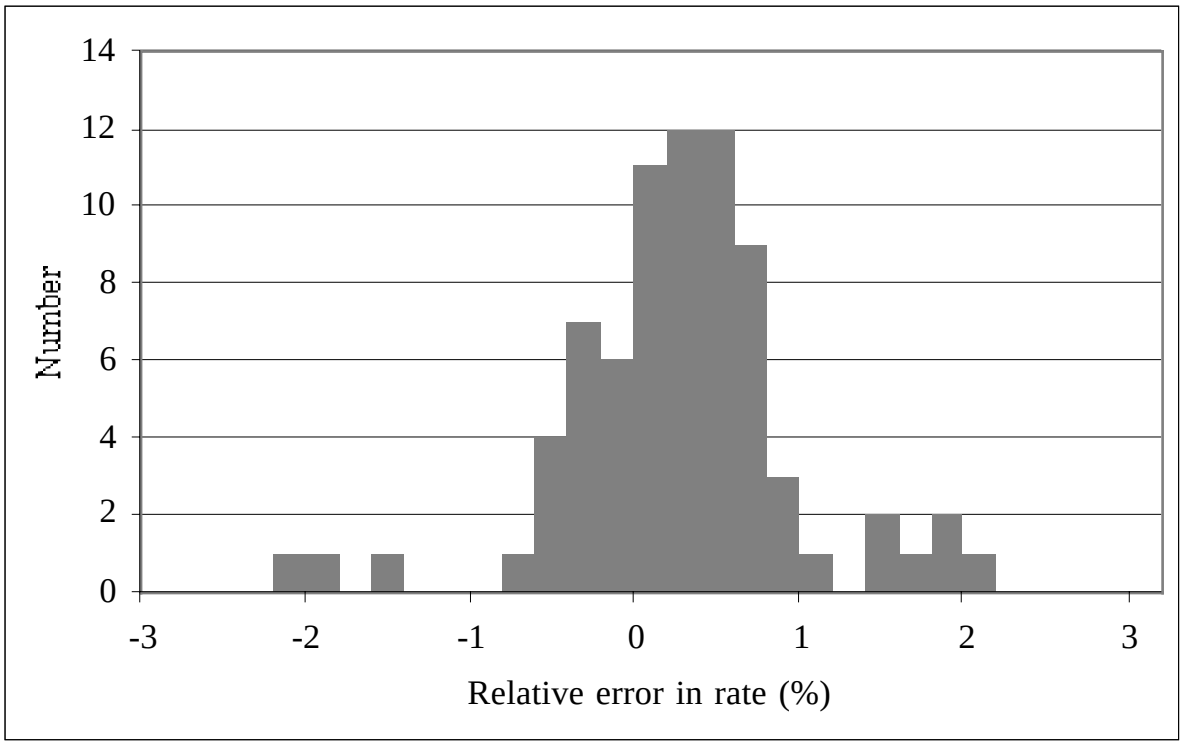

Figure 11. Relative error in the calculated rate for all configurations with 2 or more ROB emulators.

Figure 12 and Figure 13 contain histograms of the relative errors in the calculated latencies for all configurations with two or more ROB emulators and assuming a correction of 12.0 and 14.5 µs respectively (see Section 3.2). For the configurations with one outstanding event the latencies were calculated from the individual contributions to it, for all other configurations from the *calculated* rate (i.e. from the maximum number of outstanding events multiplied with the inverse calculated rate). The distribution is somewhat sharper for a correction of 12.0 µs than for a correction of 14.5 µs, the value obtained from the measurement results for the configuration with one supervisor, one steering processor, one ROB emulator and 4 outstanding events. With the value of 12.0 µs the narrow distribution shown in Figure 14 is obtained for the relative error for the configurations with one ROB emulator, where the latencies have been calculated from the *measured* rates (i.e. from the maximum number of outstanding events multiplied with the inverse measured rate).

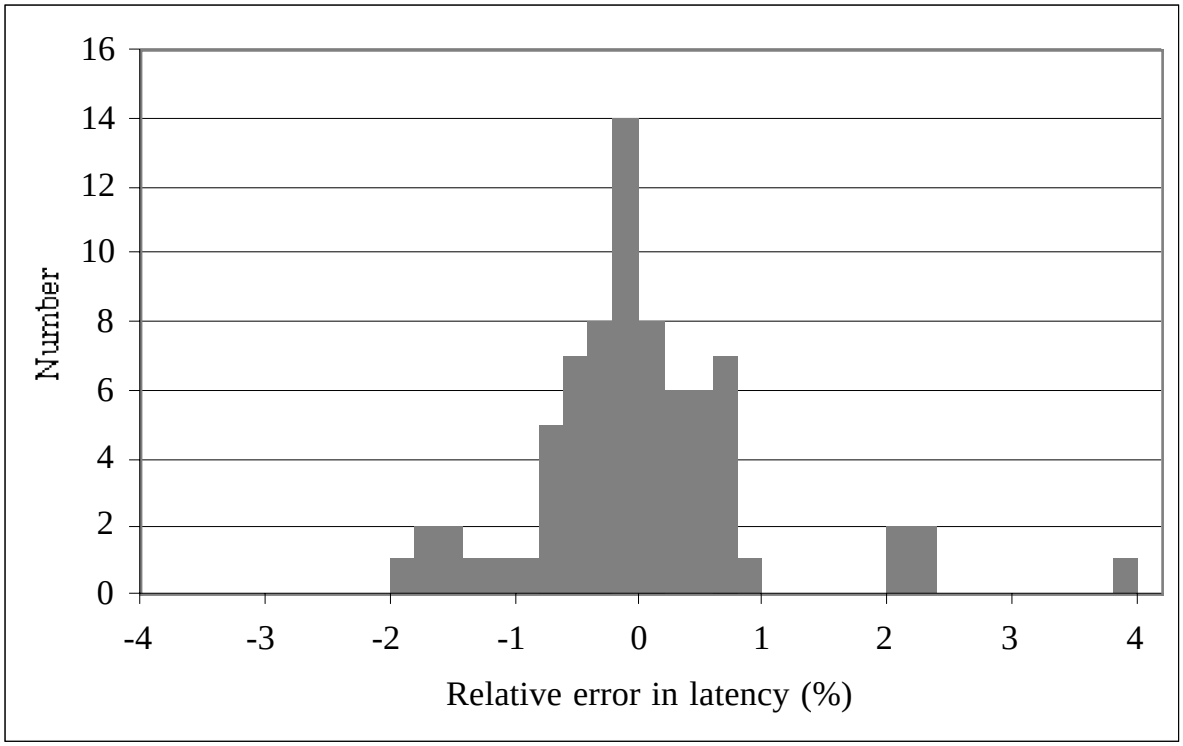

Figure 12. Relative error in the calculated latency, if the measured latency is increased with 12.0 µs, for all configurations with 2 or more ROB emulators.

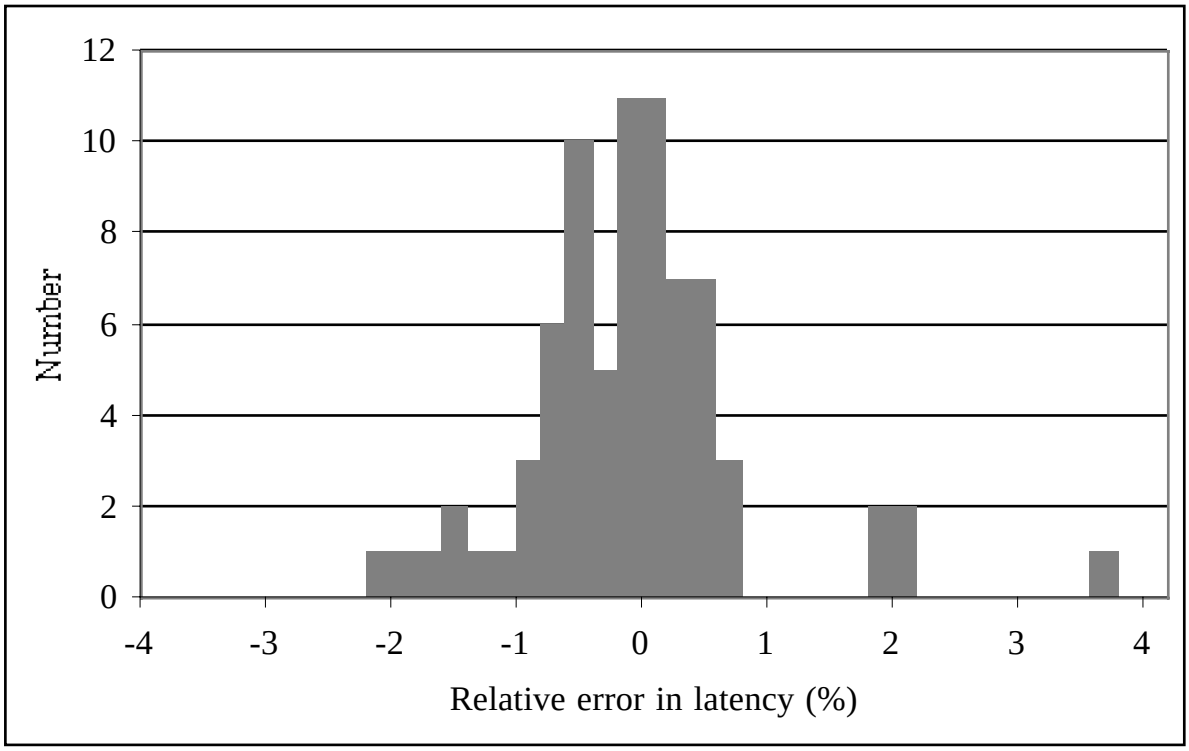

Figure 13. Relative error in the calculated latency, if the measured latency is increased with 14.5 µs, for all configurations with 2 or more ROB emulators.

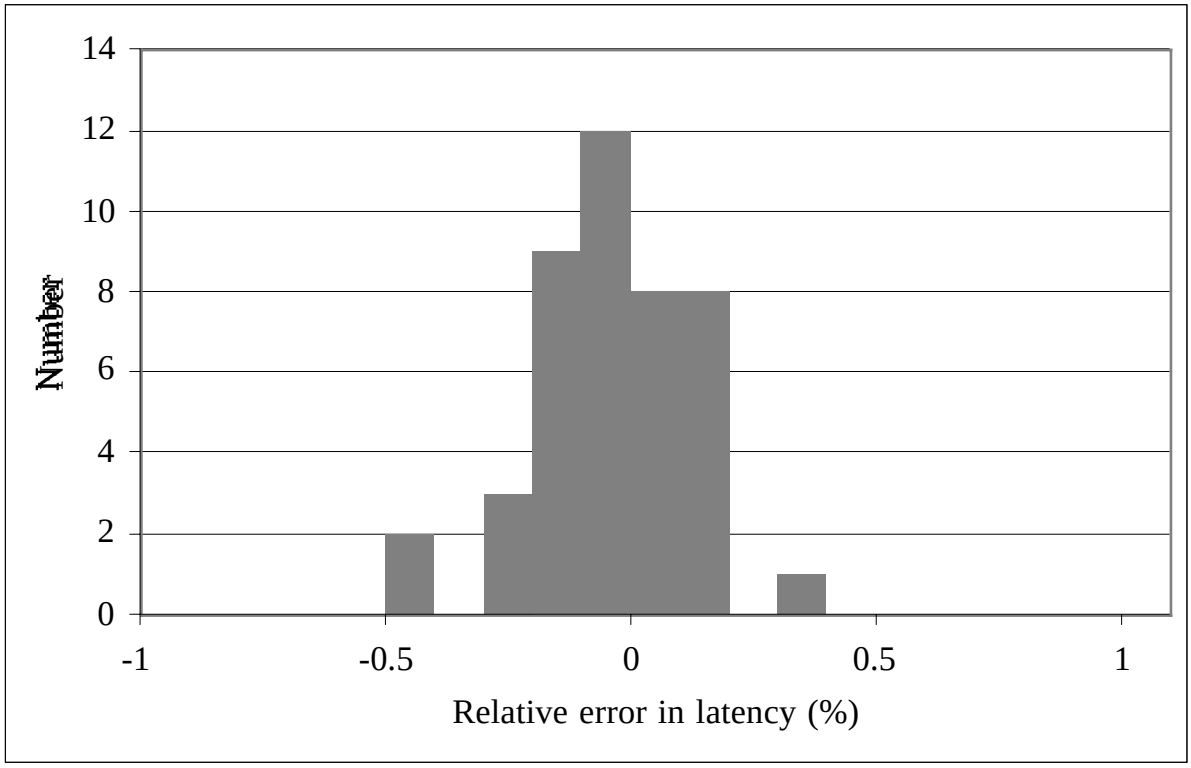

Figure 14. Relative error in the calculated latency, if the measured latency is increased with 12.0 µs, for the configurations with one supervisor, one steering processor (300 MHz or 400 MHz), one ROB emulator, 2, 3 or 4 threads and a variable number of outstanding events per supervisor.

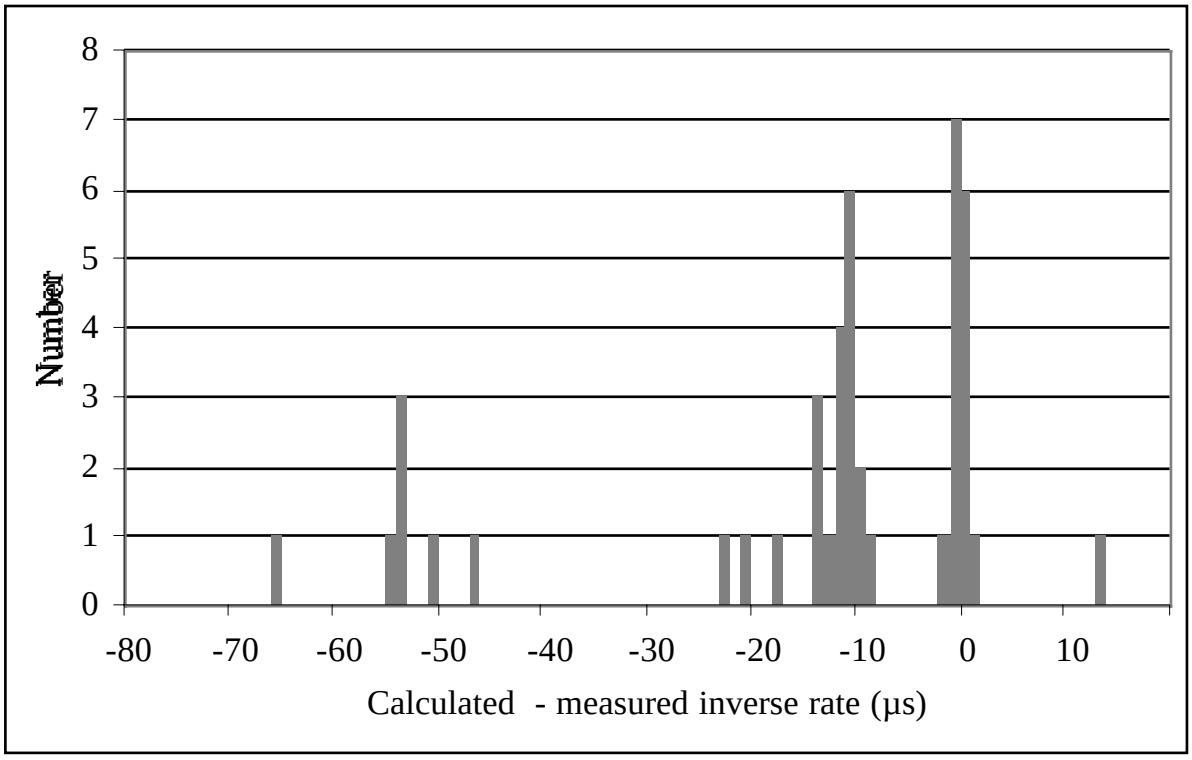

Figure 15. Differences between the inverse of the processing time per event for the steering processor and the measured inverse rate, for the configurations with one supervisor, one steering processor (300 MHz or 400 MHz), one ROB emulator, 2, 3 or 4 threads and a variable number of outstanding events per supervisor. All results for the 300 Mhz processor have been multiplied by 0.75.

In Figure 15 a histogram with the differences between the inverse of the processing time per event for the steering processor and the inverse rate, for the configurations with one ROB emulator, is presented. For a number of configurations the difference is small, for a number of other configurations there seems to be a systematic difference of somewhat larger than 10 µs. This probably can be explained from changes in the order of processing steps dealing with different events in the steering processor (TO BE CHECKED WITH COMPUTER MODELS).

4 Comparison with simulation results

5 Conclusion

6 References

[1] Testbed description
[2] Ptolemy testbed model
[3] Simdaq
[4] Source of specified message sizes
[5] Ptolemy processor star