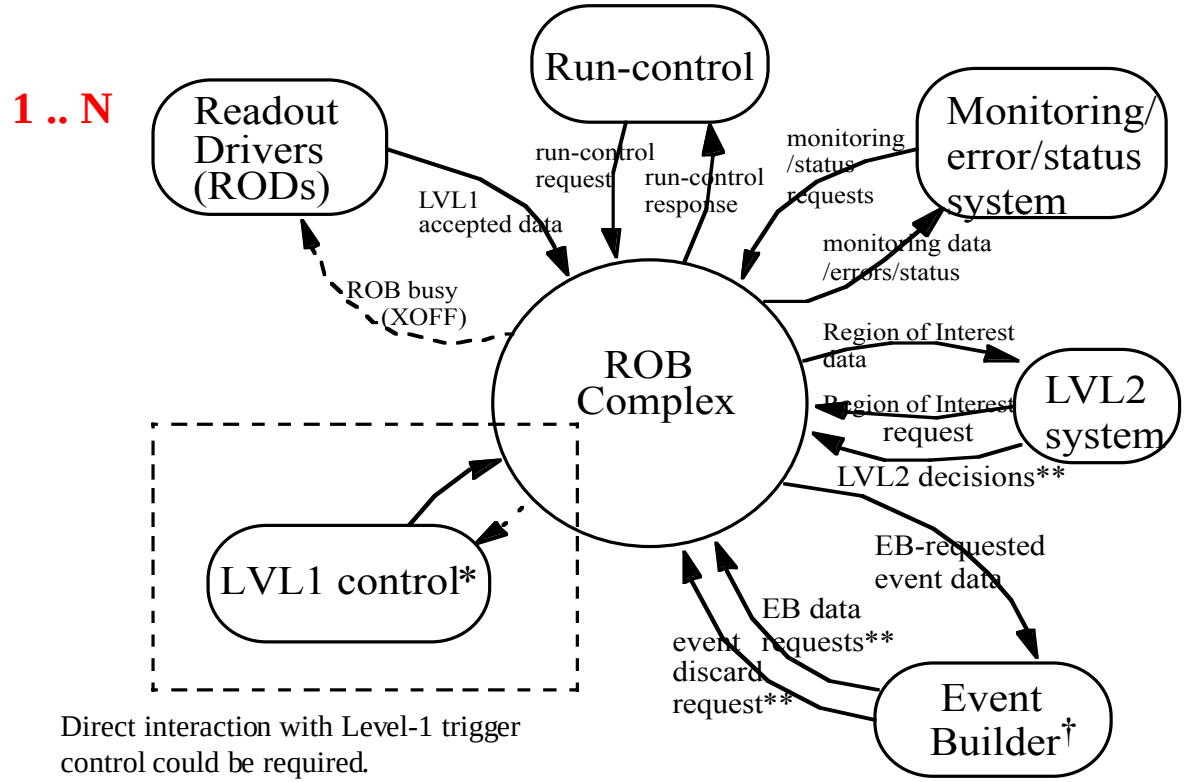
Report of the Pilot Project ROB Complex group

CERN, Mannheim, NIKHEF, Saclay, UCL/RHUL, Weizmann



ROB Complex is studied in the Pilot Project because

- it is a critical component for LVL2
- the most critical design issues are due to Level-2

But : the ROB is also a key DAQ component

The studies have led to the development of a lot of expertise within the Level-2 community

Project goals

1. Documentation of the options -> input for TP

- conclusions as far as possible
- explanation of the issues remaining to be resolved

2. Input for the ROB User Requirements Document (URD)

assumption revision of the ROB URD -> complete the Level-2 input for this URD

3. Prototyping

- building up knowledge
- provision of an existence proof
- for use in testbeds

Workplan items

1. provide input for URD, includes requirements from modelling
2. design studies, including performance measurements -> demonstration of feasibility
3. production and support of buffer prototypes
4. APIs relevant to the ROB complex
5. paper studies of scenarios

Output

- *Master Working Document*
- documents describing hard- and software designs, performance measurements and scenarios
- prototype hardware for use in testbeds

Input for URD : part of table of contents MWD

Input for URD

Modelling -> message and data rates and buffer sizes, benefits of grouping ROBs

See presentations on modelling : here LVL1 rate = 40 kHz or LVL1 rate = 75 kHz (all rates scaled up from LVL1 rate = 40 kHz, except for B-physics trigger and event building).

Event building rate : 2 kHz
Maximum fragment size : 1800 Bytes -> max. LVl1 freq for 160 MByte/s LVL1 link = 89 kHz

Requirements 40 / 75 kHz LVL1 rate:

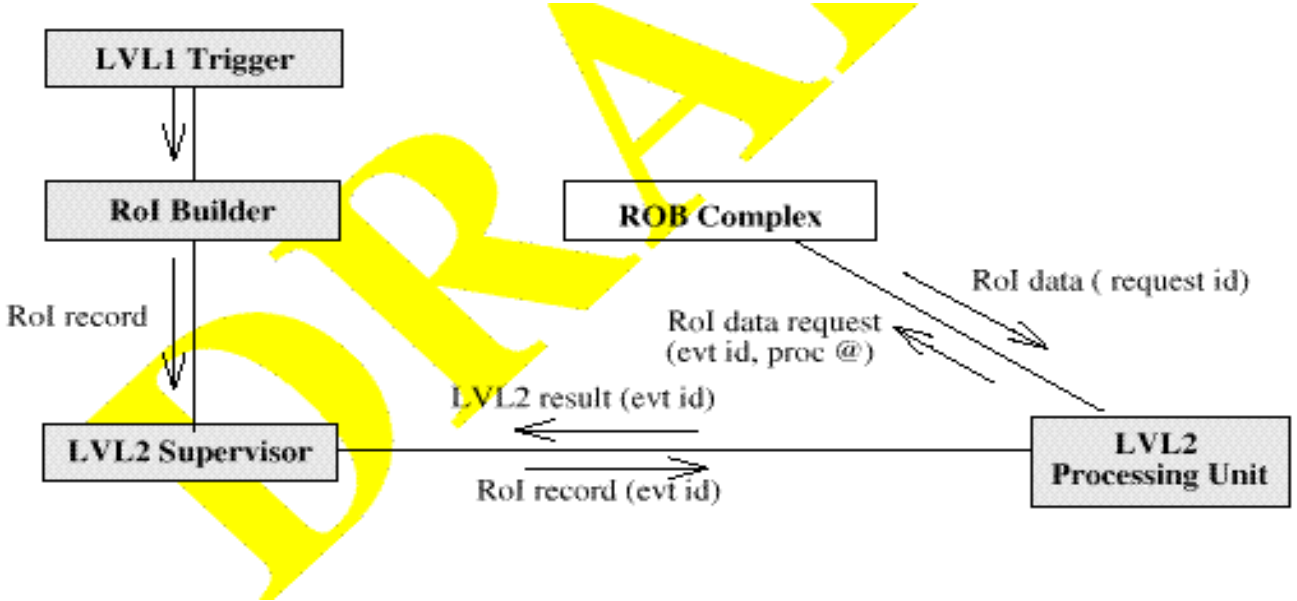
	ROIRs from LVL2	ROIRs from LVL1
Maximum input bandwidth (MByte/s)	72 / 135	72 / 135
No pre-proc. : maximum data vol out (MByte/s)	9.7 / 12.8	51.5 / 54.6
Pre-proc : maximum data vol out (MByte/s)	6.5 / 9.0	31.2 / 33.1
Pixels 112 Bytes : max. RoIR rate (kHz)	11.9 / 13.7	28.3 / 31.9
SCT 282 Bytes : max. RoIR rate (kHz)	11.1 / 12.1	26.4 / 28.4
TRT 332 / 782 Bytes : max. RoIR rate (kHz)	10.4 / 10.8	25.0 / 25.8
Cal.meters 1056 / 1832 Bytes : max. RoIR rate (kHz)	2.7 / 5.0	26.1 / 27.8

NB : data vol out = volume of data for LVL2 and EB

Existing URD + new insights ->
input for setting up a model in UML by M. Huet

See presentation M. Huet

Example diagram :



Overview of design studies

Hardware parts distinguished :

- ROBIn :
  - interface to ROL
  - event buffer
  - hardware and/or software management controlling :
    - + storage of event fragments
    - + location of event fragments requested
    - + clearing of event buffer
  - receives requests for data and delete
  - supplies the data requested or information on the location of the data
  - could also take care of pre-processing
- ROBController :
  - receives requests for data and delete messages and sends these to one or more ROBIns,
  - receives or retrieves event data from ROBIns,
  - does partial event building and and passes the result to the requester
  - could also take care of pre-processing
- ROBOut :
  - receives requests and outputs requested data,
  - could be considered to be part of the ROBController

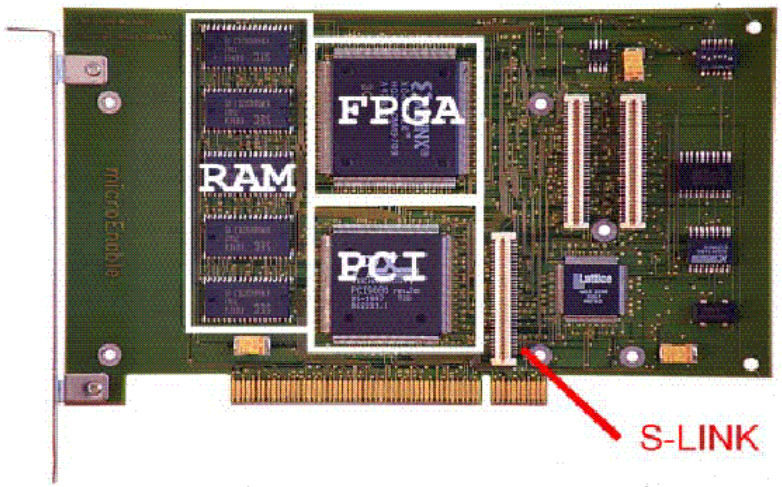
Parallel design studies / prototyping

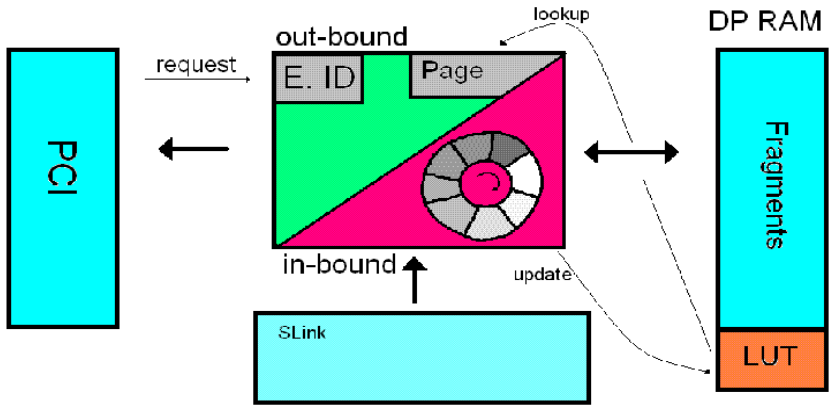
Probably not ideal

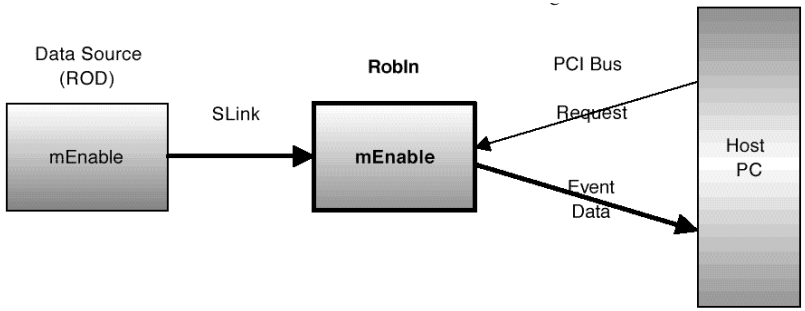
But :

- different parts of the design space were explored
- expertise was developed
- and there are results to base further work on

Mannheim : FPGA based ROBIn with paged cyclic memory implemented with µEnable, ROBController = PC. Moving to ATLANTIS hardware in CompactPCI environment







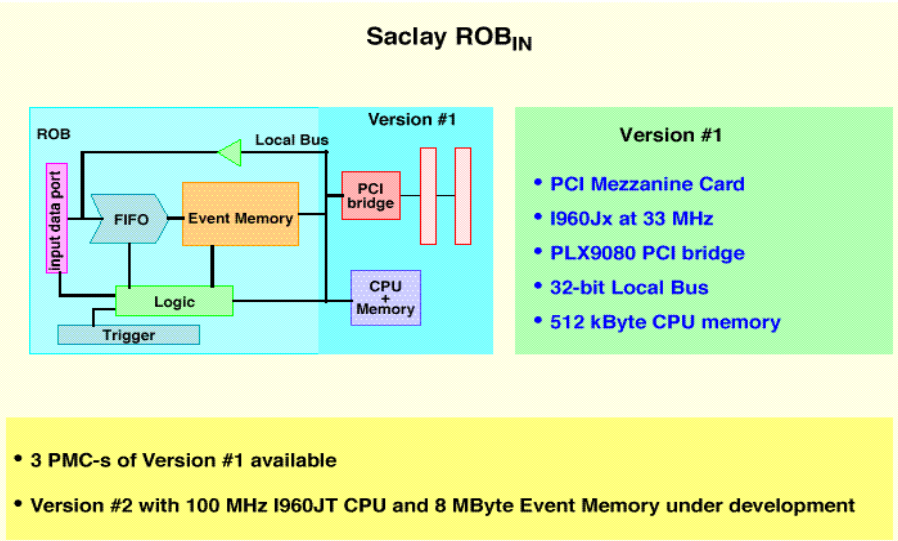
CERN/Mannheim : Active ROB Complex (AROBC) studies

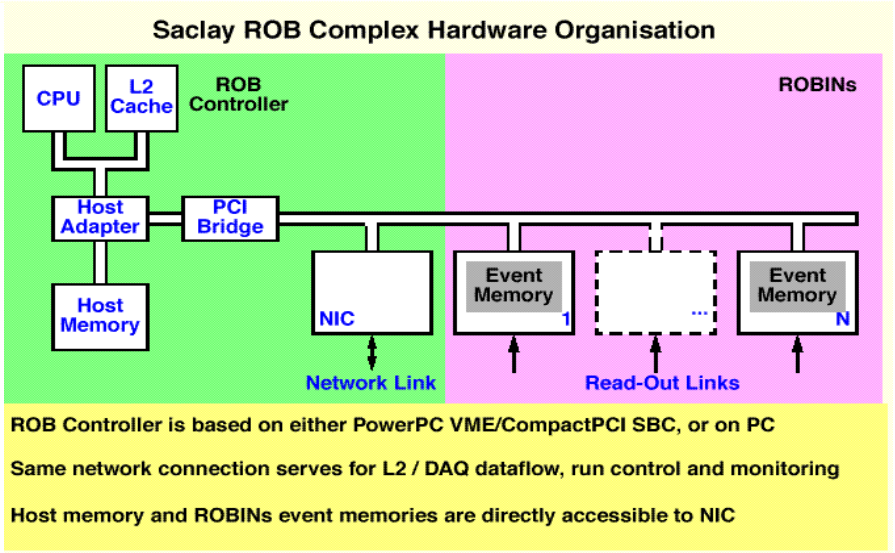
Powerful (multi-processor) ROBController, many ROBIns per ROBController

- grouped ROBin requests
- select RoI-related data from inside ROBins,
  -> reduction of bandwidth requirements on transmission
- perform preprocessing tasks depending on the detector
  -> reducing again the volumes transmitted
  -> alleviation of computing requirements for feature extractors
- may execute the feature extraction for a full RoI for some detectors and depending on topological constraints,

PC server based on a *four-processor board* (Intel SC450NX) purchased in October :
4 x 550 Mhz Pentium Xeon III , 2 PCI buses of 32 bits/32 MHz, *7 free slots*,
512 MB shared memory, 19 GB disk.
Standard commercial Micro-Enable boards used as ROBIns (Slinks are available but were not used).

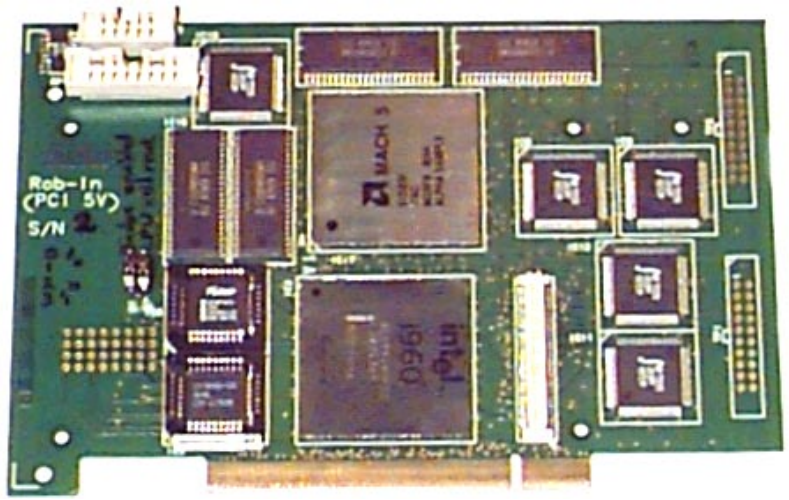
Saclay : partial i960 based ROBIn implementation with interfacing to PCI via PLX chip, full implementation before end of 1999, ROBController = CompactPCI or VME board or PC





Full CPLD based implementation of ROBIn hardware under way with paged memory management, memory = SDRAM

UCL/RHUL : i960 + CPLD based ROBIn with SRAM and paged memory management, PCI and PMC implementations, ROBController = PC or VME board (DAQ-1 setup)



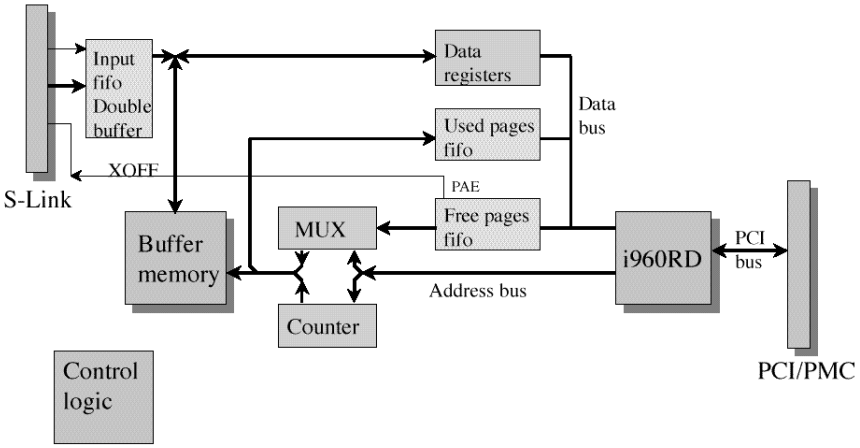
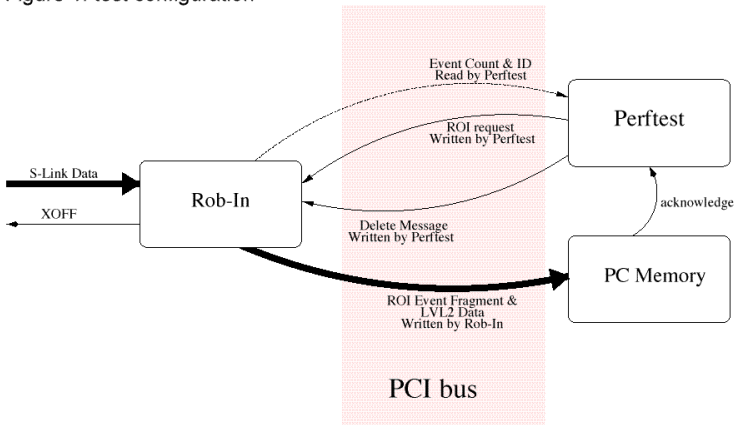
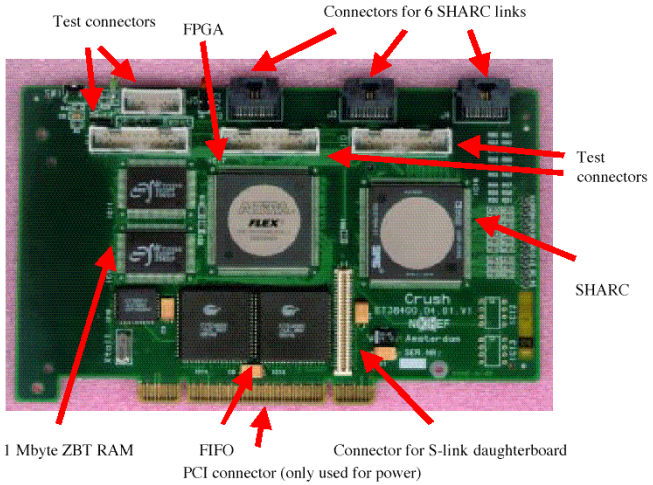


Figure 1: test configuration

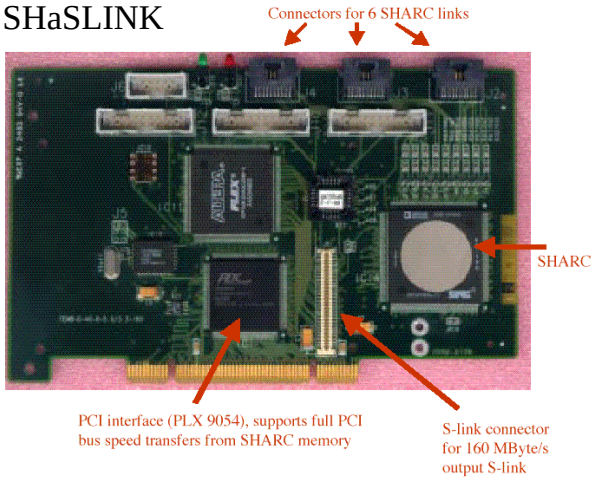


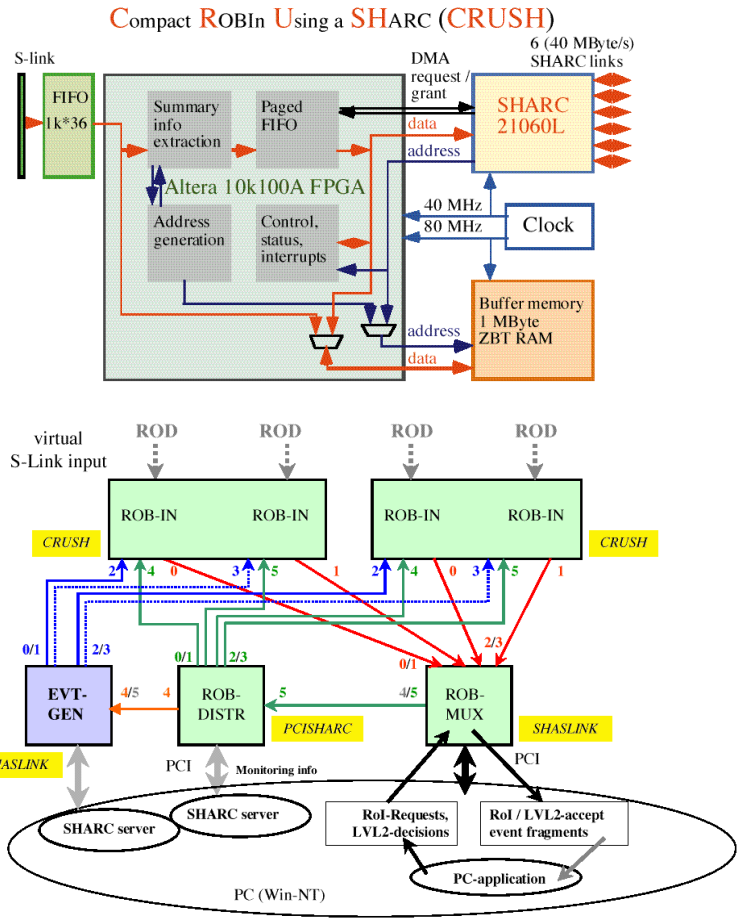
NIKHEF : SHARC + FPGA based ROBIn, ZBTRAM, cyclic buffer memory with software fragmentation control, SHARC based ROBController, connects to PCI bus or to S-link

CRUSH



SHaSLINK





ROB Complex test set-up

ROBIn performance calculated from measurement results @ 40 kHz LVL1 rate, accept rate = 2 kHz

		UK	Mannheim	NIKHEF
Maximum input bandwidth (MByte/s)		132 - output BW	80	160
1832 Bytes : maximum data vol out (MByte/s)		60	13.9	31
1056 Bytes : maximum data vol out (MByte/s)		60	8.5	27
Pixels	112 Bytes : max. RoIR rate (kHz)	23	~ 7.4	68 **)
SCT	282 Bytes : max. RoIR rate (kHz)	20	6.9	51 **)
TRT	332 / 782 Bytes : max. RoIR rate (kHz)	19 / 14.4	6.5 or 2*) / 6.2	48 **) / 30
Cal.meters	1056 / 1832 Bytes : max. RoIR rate (kHz)	~14 / ~8	6.0 / 5.6	24 / 15

*) if pre-processing included

**) 33 if PCI bus + ROB-MUX in the loop

Requirements :

		ROIRs from LVL2	ROIRs from LVL1
Maximum input bandwidth (MByte/s)		72	72
No pre-proc. : maximum data vol out (MByte/s)		9.7 (TRT)	51.5 (Cal.meter)
Pre-proc : maximum data vol out (MByte/s)		6.5 (Cal.meter)	31.2 (Cal.meter)
Pixels	112 Bytes : max. RoIR rate (kHz)	11.9	28.3
SCT	282 Bytes : max. RoIR rate (kHz)	11.1	26.4
TRT	332 / 782 Bytes : max. RoIR rate (kHz)	10.4	25.0
Cal.meters	1056 / 1832 Bytes : max. RoIR rate (kHz)	2.7	26.1

NB : data vol out = volume of data for LVL2 and EB

Mannheim : no fragmentation handling, no delete handling
NIKHEF : fragmentation detection included,
fragmentation handling reduces rate
UK, NIKHEF : grouping of 100 deletes

ROBIn performance calculated from measurement results @ 75 kHz LVL1 rate, accept rate = 2 kHz

		UK	Mannheim	NIKHEF
Maximum input bandwidth (MByte/s)		132 - output BW	80	160
1832 Bytes : maximum data vol out (MByte/s)		--	--	30
1056 Bytes : maximum data vol out (MByte/s)		--	--	26
Pixels	112 Bytes : max. RoIR rate (kHz)	14.3	8.5	65 **)
SCT	282 Bytes : max. RoIR rate (kHz)	12.4	~ 7.4	49 **)
TRT	332 / 782 Bytes : max. RoIR rate (kHz)	11.9 / 8.7	6.5 or 2 *) / 6.2	46 **) / 28
Cal.meters	1056 / 1832 Bytes : max. RoIR rate (kHz)	--	--	23 / 14

*) of pre-processing included

**) 33 if PCI bus + ROB-MUX in the loop

Requirements :

		ROIRs from LVL2	ROIRs from LVL1
Maximum input bandwidth (MByte/s)		135	135
No pre-proc. : maximum data vol out (MByte/s)		12.8	54.6
Pre-proc : maximum data vol out (MByte/s)		9.0	33.1
Pixels	112 Bytes : max. RoIR rate (kHz)	13.7	31.9
SCT	282 Bytes : max. RoIR rate (kHz)	12.1	28.4
TRT	332 / 782 Bytes : max. RoIR rate (kHz)	10.8	25.8
Cal.meters	1056 / 1832 Bytes : max. RoIR rate (kHz)	5.0	27.8

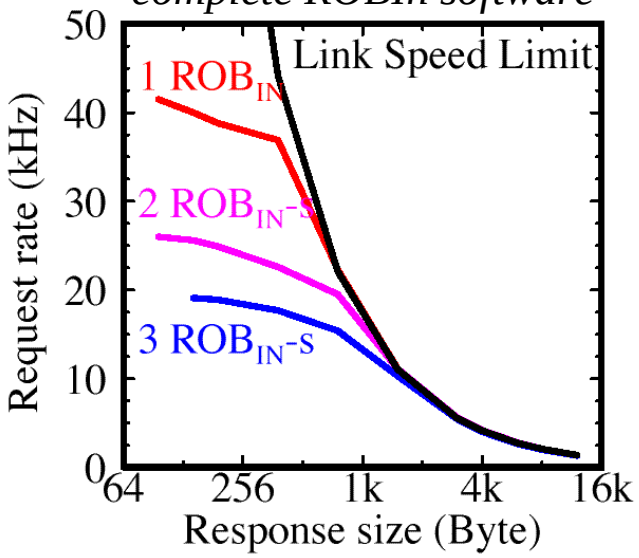
NB : data vol out = volume of data for LVL2 and EB

Mannheim : no fragmentation handling, no delete handling
NIKHEF : fragmentation detection included,
fragmentation handling reduces rate
UK, NIKHEF : grouping of 100 deletes

ROB Complex performance calculated from measurement results @ 40 kHz LVL1 rate, accept rate = 2 kHz, results for 1/2/3/4 ROBIns

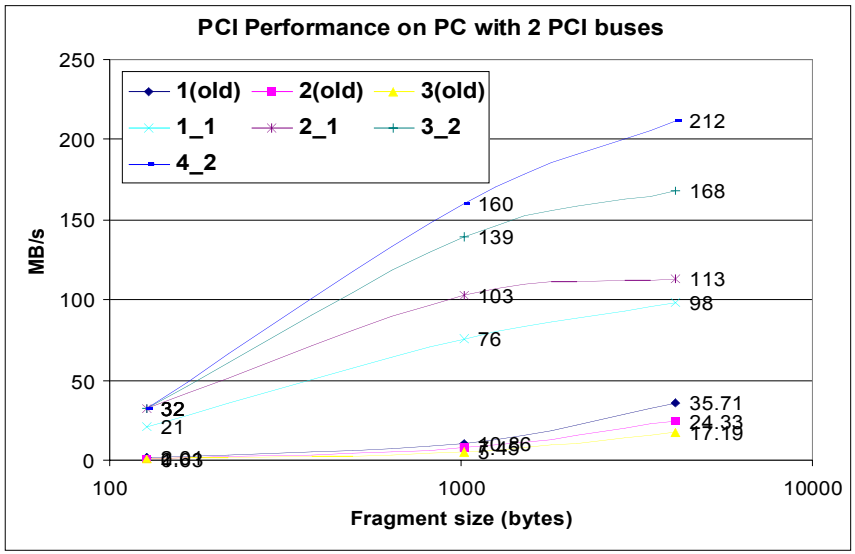
		Saclay	CERN/Mannheim	NIKHEF
Pixels	112 Bytes : max. RoIR rate (kHz)	26/23/17	175/152/98/71	33/25/19/15
SCT	282 Bytes : max. RoIR rate (kHz)	36/19/12	137/105/77/59	33/25/19/15
TRT	782 Bytes : max. RoIR rate (kHz)	20/10/6 *)	83/55/47/39	29/21/16/12
Cal.meters	1056 Bytes : max. RoIR rate (kHz)		68/43/38/33	22/18/13/10
Cal.meters	1832 Bytes : max. RoIR rate (kHz)		44/26/25/22	15/13/9/6

Saclay results for less complete ROBIn software



*) Limited by network bandwidth

CERN/Mannheim AROBC results



ROB complex group report, Beatenberg, JV, 7- 12-99

ROB Complex performance calculated from measurement results @ 75 kHz LVL1 rate, accept rate = 3.75 kHz, results for 1/2/3/4 ROBIns

		Saclay	CERN/Mannheim	NIKHEF
Pixels	112 Bytes : max. RoIR rate (kHz)	24/21/15	173/150/96/69	33/25/17/13
SCT	282 Bytes : max. RoIR rate (kHz)		135/103/75/57	33/23/17/13
TRT	782 Bytes : max. RoIR rate (kHz)		81/53/45/37	25/18/14/11
Cal.meters	1056 Bytes : max. RoIR rate (kHz)		66/41/36/30*)	20/15/11/8
Cal.meters	1832 Bytes : max. RoIR rate (kHz)		42/24/23/21*)	13/10/7/3.7

*) Input Bandwidth not sufficient with current micro-Enable ROBIn implementation

Output BW limiting factor

Scenarios

Three categories :

1. Each ROBIn has its own network interface

2a. ROBIns connect via PCI bus (Compact PCI, via PMC or PCI connector on VME card or PCI in (multi-CPU, multi-bus) PC) to ROB Controller.
DAQ-1 ROC falls into this category
2b. ROB Complex connects via PCI bus to CPU managing network interface

3. ROB Complex connects via simple point-to-point link to LVL2 farm processor.
One network interface per crate with O(100) ROBIns for requesting data

Which is the best ? Maybe no unique answer, in any case the answer is time dependent :

1. Interesting, not yet demonstrated
2. "Standard" solution with different possible choices, emphasis on PCI bus
3. Weak coupling between purpose built hardware and commercially available hardware may be an advantage

Conclusions

- Feasibility demonstrated
- Hardware available for prototype/testbed studies
- MWD, with input for URD, being finished
- In view of risk of hardware being obsolete in 2005 :
  avoid taking decisions now, follow up on different scenarios
  with second category as baseline option ?

- One group looking to ROBs very desirable

- Different parallel activities were fine for "brainstorming" phase
  (in view of the "brainstorming process" itself and in view of education),
  but now it is time to find a way to agree on the assignment of
  responsibilities within a single team