Future DAQ note

The active ROB complex

R.K.Bock, D.Francis, J.Vermeulen, S.Wheeler
Preliminary draft,
29 Oct 1999

Abstract

We present results from a modeling exercise allowing to evaluate the changes brought about by grouping readout buffers in the traffic between individual readout buffers (ROBins) and level-2 processors. The assumption is made that ROBins are grouped into ‘active ROB complexes’ (AROBCs). An AROBC consists of a commercial off-the-shelf local network of multiple processors, sharing memory and access to a number of ROBins. Typically, today, this might be implemented by an SMP system with up to 8 processors with of the order of 20 PCI slots, or by adding substantial computing capacity inside a ROB crate as pursued by the project ‘DAQ prototype-1’. The grouping of ROBins we assume is adapted to individual detectors, and we set a limit of 16 ROBins per AROBC. We present the effect of this grouping on the traffic on the general communication network, and the computing requirements arising for AROBCs. The communication internal to the individual AROBCs will be considered at a later stage.

Introduction

An active ROB (AROBC) complex can alleviate potential problems, in particular in required transmission possibilities, in various ways:

- an AROBC receives grouped ROBin requests, reducing the number and frequency of requests for the general network and the steering processors;
- an AROBC selects RoI-related data from inside ROBins, reducing bandwidth requirements on transmission;
- an AROBC performs preprocessing tasks depending on the detector, reducing again the volumes transmitted, and alleviating computing requirements for feature extractors;
- for some detectors and depending on topological constraints, an AROBC may execute the full feature extraction for a region of interest (RoI).

In more detail, the processing possibilities in an AROBC arise in the following way. The processing of data in the level-2 trigger is usually split into the steps of feature extraction and decision making. In the (dominant) mode of region-of-interest guided analysis this is a particularly advantageous factorization of work: the first step runs on data of a single (sub-) detector and transforms raw data into a physics-related and shorter set of parameters (features), the second operates on multiple sets of features and looks for their compatibility with one or several of the physics hypotheses described in menu tables.

Raw data, however, are not presented to the trigger system in a grouping and format optimally adapted to feature extraction; they are constrained by front-end electronics and transmission up to the readout buffer, whereas the processing algorithms expect them arranged in handy arrays or objects. Summarily, the operations preparing data for feature extraction are named ‘preprocessing’. Typically, preprocessing does not transform raw data in an essential way, although for some detectors parts of the data present in the readout buffer may be considered unnecessary for level-2 processing, and therefore omitted, particularly if this alleviates algorithm and transmission requirements.

Most preprocessing also consists of operations that can be executed locally, e.g. inside a readout buffer - with the obvious exception of collecting all data for a region of interest from multiple readout buffers. While it is possible to consider preprocessing as first step of feature extraction, it is also reasonable to imagine architectures that have the required intelligent nodes upstream of feature extraction processors, e.g. implementations that ascribe preprocessing to groups of readout buffers (‘ROB complex’). Note that the execution of the preprocessing steps in a general-purpose processor can take non-negligible computer time, and the gain by executing these steps in parallel and locally, before a general transmission network is used, can result in critical savings in computer and transmission capacity.

Finally, it is conceivable that the aggregation of data in an AROBC is such that an RoI is fully available in most or many cases. If so, the feature extraction step, CP time intensive in some detectors, could conceivably also be executed in the ROB Complex (hence in parallel for different RoIs and detectors). This possibility is also explored, in particular the probability that data are complete enough and the computing requirements for this step.

Input

Preprocessing is very much a detector-dependent operation. Below, we describe summarily the preprocessing operations for the major detectors (see also LHCC 98-15, chapter 7), the time spent in a general-purpose computer for executing these operations, and the effect they have on subsequent operations. We use a grouping of detector data into readout buffers (ROBins) as described in ATL-COM-DAQ-99-014 (‘Detector and Read-Out Specification, and Buffer-RoI Relations, for Level-2 Studies’, P.Clarke et al.). We do not address the problem, common to all detectors, of identifying the readout buffers that contain information for a given region of interest, assuming this is mediated by some very fast lookup tables. . Benchmarking results have been taken from various measurements on 300 or 400 MHz processors, with no corrections (hence the computing needs are somewhat overestimated when assuming the availability of faster processors).

For the RoI access frequencies we use the physics model with sequential algorithms as used in modeling, taken from J.Vermeulen’s ‘Modelling Parameters’ (available only as presentation, note in preparation). They are based on menus resulting in approximately 40 kHz overall LVL1 trigger rate in both high and low luminosity mode (hence there is no contingency in frequencies). We also assume constant-size RoI-s, although they are not assumed the same for all detectors, and for electron, jet and muon triggers.

ROB grouping and processing in detail

The readout buffers have been grouped into a total of 134 active ROB complexes: Pixels 12, SCT 12, TRT 24, Ecal 56, Hcal 10, MuPrec 16, and MuTr 4. Details are given in the appendix (in terms of files which are the input to a preprocessing program for modeling). This results in a much reduced number of access lines for the general network (a total of 1530 ROBins before grouping).

The AROBC for the pixel and SCT precision trackers

The readout of the wafer-based detectors presents several specific problems. Wafers are constant-size detector elements that cannot be arranged to cover specific areas in eta/phi, and the information is sparse compared to the number of channels; the two facts together result in comparatively large and irregular areas collected in each ROBin. We assume the readout grouping in layers (S.Wheeler, ATL-DAQ-99-006, ‘Optimisation of ROB mapping for SCT and Pixel detectors’). Furthermore, tracks leave multiple signals in each layer, so that a clustering procedure is necessary. The steps of preprocessing are then the following:

- selection of wafers, inside a readout buffer, that are part of the currently considered region of interest;
- clustering of adjacent strips or pixels into hits corresponding to a single track;
- conversion of local coordinates (wafer ID and strip or pixel number) to global coordinates (eta/phi/radius or similar); this includes combining the normal and stereo wafers for the SCT (present algorithms all are based on space points only, but this may not remain true forever).

Benchmarking of these algorithm parts has given the following numbers per ROBin (from J.Baines/R.Dankers): selection of wafers: 0.02 msec; clustering: 0.025 / 0.17 msec (SCT / Pix); coordinate transformations and space points: 0.10 / 0.14 msec (SCT / Pix). The accumulated effect is a data volume reduction by a factor ~4, most of it due to wafer selection. Feature extraction algorithms take 10 msec per RoI, the RoI size is assumed 0.2 x 0.2.

The overall analysis frequency of RoIs in the precision trackers is 12.7/30.6 kHz for high/low luminosity. The AROBC access rates are given in tabular form, below. Details about the grouping of ROBins into ROB complexes are found in the appendix.

The AROBC for the transition radiation tracker

The readout of the TRT has been arranged such that the relation between RoIs and ROBins is optimal, and comparatively little non-RoI information is transmitted without selecting data inside ROBins. A subdivision in eta of this detector in the barrel does not exist (except left/right halves), and is not practical (and would not help the data volumes much) in the endcap; selection of a suitable phi-sector thus is sufficient to obtain access to all digitizings relevant for a region of interest.

Preprocessing for the TRT consists mostly of format conversion (zero suppression), and has been benchmarked to take some 0.17 msec per ROBin (from M.Sessler). The probability of wires being hit by some track (‘occupancy’) can go to very high values, and the preliminary decisions taken for the TRT data format are giving high priority to the high-luminosity (= high-occupancy) case; this means that, at least for low luminosity bulk analysis of the TRT (‘full scan’) a more efficient zero suppression is possible. The effect of preprocessing is a more suitable data format (algorithm time for feature extraction reduced by ~30%) and a luminosity-dependent reduction of data volume, which at high L is not more than 1.3, but at low L, we assume a compression factor of 5. Feature extraction has been benchmarked at approximately 1 msec per RoI; the size of the RoI is assumed to be 0.2 x 0.2.

The overall analysis frequency of RoIs in the TRT is 12.7/30.6 kHz for high/low luminosity. The AROBC access rates are given in tabular form, below. Details about the grouping of ROBins into ROB complexes are found in the appendix.

The AROBC for the calorimeters

The readout grouping for calorimeter ROBins is still under discussion; we follow here ATL-COM-DAQ-99-014 quite literally.

Preprocessing for the calorimeter consists of selecting the RoI-related subsets of data in each ROBin and has not been benchmarked for the present assumptions; we assume 0.03 msec per ROBin, from earlier work. We estimate a reduction in data volume by a factor of two (recent work by S.Gonzalez), but the arranging of data for the algorithms typically takes more computing time than feature extraction, although it can only be done partially inside an AROBC. For jets, preprocessing could consist of presenting the data in a much coarser and unified granularity, and would then substantially reduce the volume of data to be transmitted (we make no assumptions). Feature extraction is assumed to take approximately 0.1 msec per RoI (e.m. calorimeter and hadronic each). The RoI size for electron and muon triggers is assumed to be 0.3 x 0.3 for ROB access requirements, but algorithms are thinkable which select (in finer granularity) a smaller RoI (like 0.2 x 0.2) for transmission. For jets, the RoI size is 0.8 x 0.8. Preprocessing for jets can be guessed to take 0.1 msec (summing of cells).

The overall analysis frequency of RoIs in the e.m. calorimeter is 42.9/31.3 kHz for high/low luminosity. The corresponding rates in the hadron calorimeter are 14.5/18.7 kHz. For the subset of jets, the rate is 2.5/3.4 kHz in both e.m. and hadronic calorimeters. The AROBC access rates are given in tabular form, below. Details about the grouping of ROBins into ROB complexes are found in the appendix.

The AROBC for the muon detectors

The muon detectors pose no major frequency or bandwidth problems, and a preprocessing task with any sizeable effect on computing or gain in bandwidth has not been defined. For these detectors, we give only the access rates and bandwidths. Details about the grouping of ROBins into ROB complexes are found in the appendix.

The effect of active ROB complexes on transmission

With the definitions above, we obtain the following tables. Note that the columns labeled (ROBin) are numbers for the totally ungrouped ROBins, given for comparison. The bandwidth given for results of feature extraction is derived from the access frequency, assuming a result packet of 256 bytes. Added traffic due to non-containment of RoIs is not included.

High Luminosity

		Access frequencies [kHz]				Required bandwidth [MB/s]			
Detector	#ROBin/ #AROBC	(ROBin)	average	upper limit	Lower limit	(ROBin)	all data	with preproc.	Ev.Buil. (added)
Pixels	84/12	0.7	2.8	4.9	1.1	0.57	4.1	1.0	11.5
SCT	92/12	0.4	1.8	3.1	0.7	0.75	5.8	1.4	24.7
TRT	256/24	0.3	0.9	1.1	0.8	0.26	2.8	2.2	21.7
Ecal	760/56	1.0	2.2	2.9	0.7	1.44	19.6	9.8	49.0
Hcal	98/10	0.9	2.2	3.3	0.4	1.53	15.1	7.6	35.4
MuPrec	92/16	0.2	0.9	1.0	0.8	0.12	1.5	-	19.7
MuTrig	48/ 4	0.3	3.3	3.3	3.3	0.13	1.7	-	9.7

Low Luminosity

		Access frequencies [kHz]				Average required bandwidth [MB/s]			
Detector	#ROBin/ #AROBC	(ROBin)	average	upper limit	lower limit	(ROBin)	all data	with preproc.	Ev.Buil. (added)
Pixels	84/12	11.0	14.4	17.9	11.8	1.2	9.1	2.3	1.6
SCT	92/12	10.6	12.8	14.9	11.0	3.0	23.4	5.8	4.3
TRT	256/24	10.3	11.4	11.7	11.2	8.1	86.5	17.3	16.7
Ecal	760/56	0.7	1.7	2.1	0.6	1.1	15.0	7.5	49.9
Hcal	98/10	1.2	2.9	4.2	0.7	2.0	19.7	9.8	36.0
MuPrec	92/16	0.3	1.8	2.0	1.6	0.2	3.0	-	20.0
MuTrig	48/ 4	0.6	6.5	6.5	6.5	0.3	3.4	-	9.9

The computing aspect of active ROB complexes

We estimate the computing capacity needed in the AROBCs for the different detectors for preprocessing. We also give the probability for a RoI to be contained in a single AROBC (excluding jet RoIs), and the computing capacity that would be needed for feature extraction (high/low luminosity). Computing capacity is given in terms of the number of processors (as used in benchmarking) required, on average, in a single AROBC. Preprocessing estimates use the per-ROBin processing time, and multiply with the number of ROBins used and with the frequency of requests (separately for the three different types of RoIs). Feature extraction is estimated by multiplying the AROBC access frequency by the processing time per RoI; this is a clear overestimation for RoIs spanning several AROBCs. We should underline that in any case, the numbers in this table are crude estimates, not backed up by up-to-date benchmarked algorithms, and ignoring both the requirements in AROBC-internal traffic and all communication-related computing. Also, no attempt is made to estimate the additional communication needs arising for the general network from the necessity to collect fragments for non-contained RoIs.

Detector	#active ROBins/ AROBC	average frequency (hi/lo L)	computing load, high L (prepr.)	computing load, low L (prepr.)	computing load (appr) (feat.ex.)	probab.of RoI contained
Pixels	84/12	2.8 / 14.4	2.4	12.3	28 / 144	0
SCT	92/12	1.8 / 12.8	0.4	3.2		0.4
TRT	256/24	0.9 / 11.4	0.3	3.4	1 / 11	0.5
Ecal	760/56	2.2 / 1.7	0.2	0.1	0.2 / 0.2	0.1
Hcal	98/10	2.2 / 2.9	0.1	0.1	0.2 / 0.3	0.7

Conclusions

This modeling exercise for the set of assumptions explained above has shown the following effects:

- a reduction in number of ports for the general network by a factor 11.4;
- a corresponding substantial reduction of the total number of messages exchanged over the network;
- the possibility to stay within technology limits for frequencies and bandwidth needed for level-2 traffic, assuming that 15 kHz and 20 MB/s are acceptable; this assumes that the computing capacity in the AROBC is (partly) used for compacting the data;
- the limit in bandwidth is, in fact, reached (and the above numbers exceeded) by the Event Builder traffic; we have assumed, for the purposes of this note, that data compaction for the EB is undesirable. The aggregation of data into larger packets (> 4 kB) may, of course, help to better utilize the available hardware;
- there seems to be no problem in providing enough computing capacity in an AROBC to execute preprocessing, for all detectors.
- the possibility of using AROBCs for feature extraction for at least some detectors seems fairly remote, given the low probabilities of containing RoI data in a single AROBC. The implication is that substantial communication with other AROBCs would be necessary, over a general network. This traffic would be less, however, than for feature extraction in a separate (‘steering’) processor. We have given some indication for the computing load for feature extraction in order to show that also the availability of sufficient computing capacity may be problematic (precision trackers).

Appendix : geometrical definition of AROBCs

The active ROB complex definition for pixels:

```
#Pixels  grouped  mapping (RKB 22/09/99)
Format EtaMinMaxPhiMinMax Detector Pixels

0     -2.6923365  -1.7005045  -0.0436335  3.0979595
1     -2.6923365  -1.7005045   3.0979595  6.2395525
2     -2.8934445   0.65457071 -0.1737665  3.3329845
3     -2.8934445   0.65457071  3.3168935  6.1255115
4     -2.0530605   0.29472457 -0.0775335  3.3151025
5     -2.0530605   0.294724576 3.3063685  6.2139915
6      0.22253605  2.0530605  -0.0775335  3.3151025
7      0.22253605  2.0530605   3.3063685  6.2139915
8      0.654570716 2.8934445  -0.1737665  3.3329845
9      0.654570716 2.8934445   3.3168935  6.1255115
10    1.7005045   2.6923365   -0.0436335  3.0979595
11    1.7005045   2.6923365    3.0979595  6.2395525
```

The active ROB complex definition used for the SCT:

```
#SCT   grouped  mapping (RKB 22/09/99)
Format EtaMinMaxPhiMinMax Detector SCT

0     -2.5678865  -2.0436645  -0.0785405  3.0811775
1     -2.5678865  -2.0436645  3.0630535   6.2227705
2     -2.4197275  -1.1798135  -0.0785405  3.0811775
3     -2.4197275  -1.1798135  3.0630535   6.2227705
4     -1.6406675  0.0130525   -0.29952380 3.2486485
5     -1.6406675  0.0130525   3.1937655   6.1201105
6     -0.0130525  1.6406675   -0.29952380 3.2486485
7     -0.0130525  1.6406675   3.1937655   6.1201105
8     1.1798135   2.4197275   -0.0785405  3.0811775
9     1.1798135   2.4197275   3.0630535   6.2227705
10    2.0436645   2.5678865   -0.0785405  3.0811775
11    2.0436645   2.5678865   3.0630535   6.2227705
```

The active ROB complex definition for the TRT:

```
#TRT ROB Map grouped mapping (RKB 22/09/99)
Format Normal Detector TRT
8 -2.4 -0.7
4 -1.1 0.0
4 0.0 1.1
8 0.7 2.4
```

The active ROB complex definition for the calorimeters:

```
L-Ar Cal ROB map Updated DAQ note 62, layers (modified RKB 22/09/99), ~12)
Format Normal Detector EmCal
#EmCal combined layers (presampler + front + middle + back)
2 -3.2 -2.4
2 -2.4 -2.2
2 -2.2 -2.0
2 -2.0 -1.8
2 -1.8 -1.6
2 -1.6 -1.4
8 -1.4 -0.8
8 -0.8  0.0
8  0.0  0.8
8  0.8  1.4
2  1.4  1.6
2  1.6  1.8
2  1.8  2.0
2  2.0  2.2
2  2.2  2.4
2  2.4  3.2

#HadCal ROB map Updated DAQ note 62, layers RKB 23/09/99
Format Normal Detector HadCal
1 -3.2 -2.5
1 -2.5 -1.6
1 -1.6 -1.5
4 -1.6 1.6
1 1.5 1.6
1 1.6 2.5
1 2.5 3.2
```

The active ROB complex definitions for muon precision and trigger chambers:

```
#Mu precision ROB map : modified RKB 23/09/99
Format Normal Detector MuPrec
4 -2.7 -1.05
4 -1.05 0.0
4  0.0  1.05
4  1.05 2.7

#Mu triggerdetector ROB map modified RKB 23/09/99
Format Normal Detector MuTrig
2 -2.5  0.0
2  0.0  2.5
```