BRAHMS

RCF quarterly meeting

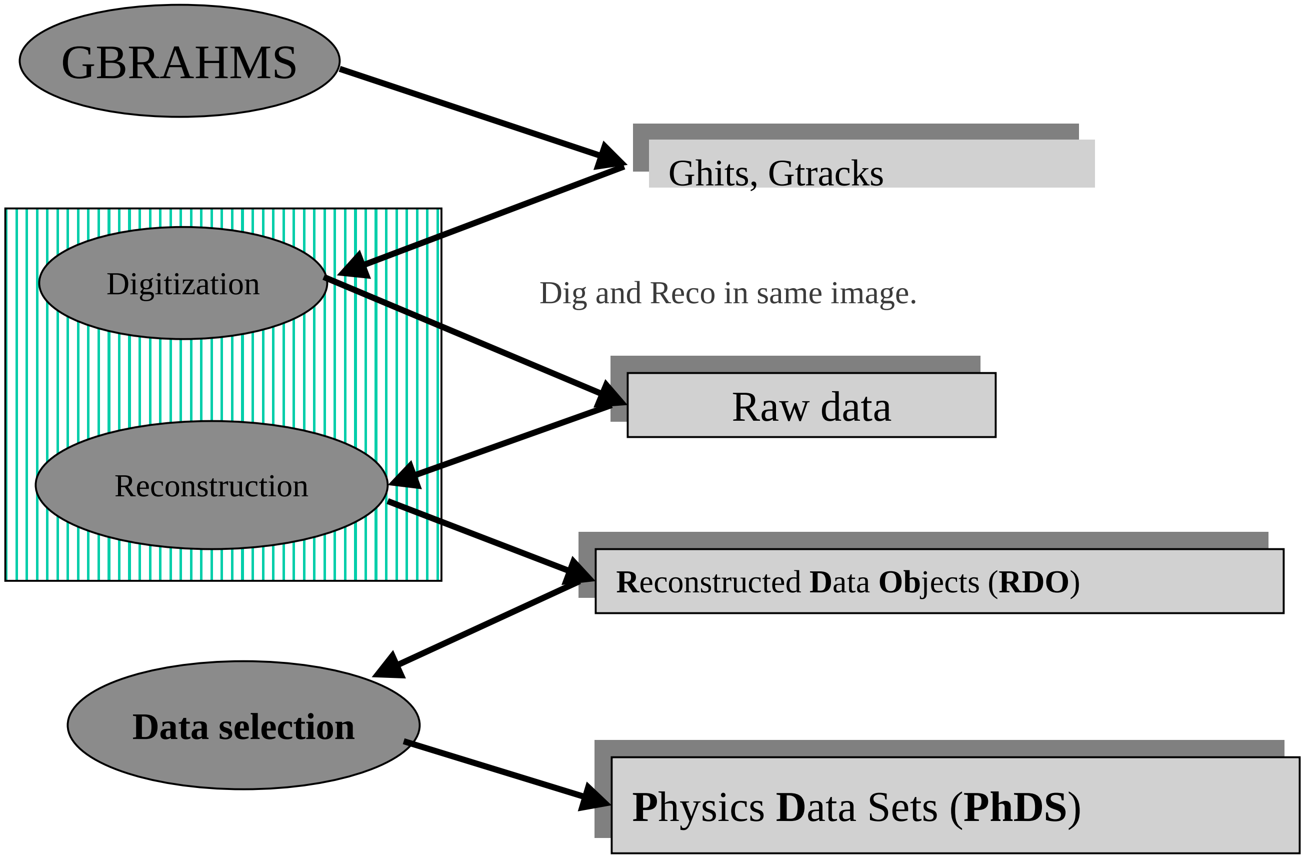
F.Videbaek

April 1, 1999

MDC-II

- Out activities was similar to those in from MDC-I. Coordinator Alv Kjetil Holme.
- Scripts have been written to access the HPSS, CRS system in a controlled fashion
  - Brahms add own layer of scripts , mostly written in PERL.
    - Sinking of *raw* data to HPSS (single or multiple jobs)
    - CRS job submission
    - Un-sinking of reconstructed data from HPSS to CAS disk farm.
    - Sinking of μDST to HPSS.
  - Various utilities to examine HPSS and create job summaries.
- Reconstruction code for Forward Spectrometer (BramReco)
- Event selection code (Bminor)
- Simple analysis code (BramAna)
- The reconstruction software worked quite well during the MDC-II in fact from our end no real bugs were found.
- The web pie charts provided by Dave Morrison was very useful.

Data Flow



GBRAHMS

Ghits, Gtracks

Digitization

Dig and Reco in same image.

Raw data

Reconstruction

Reconstructed **D**ata **Ob**jects (**RDO**)

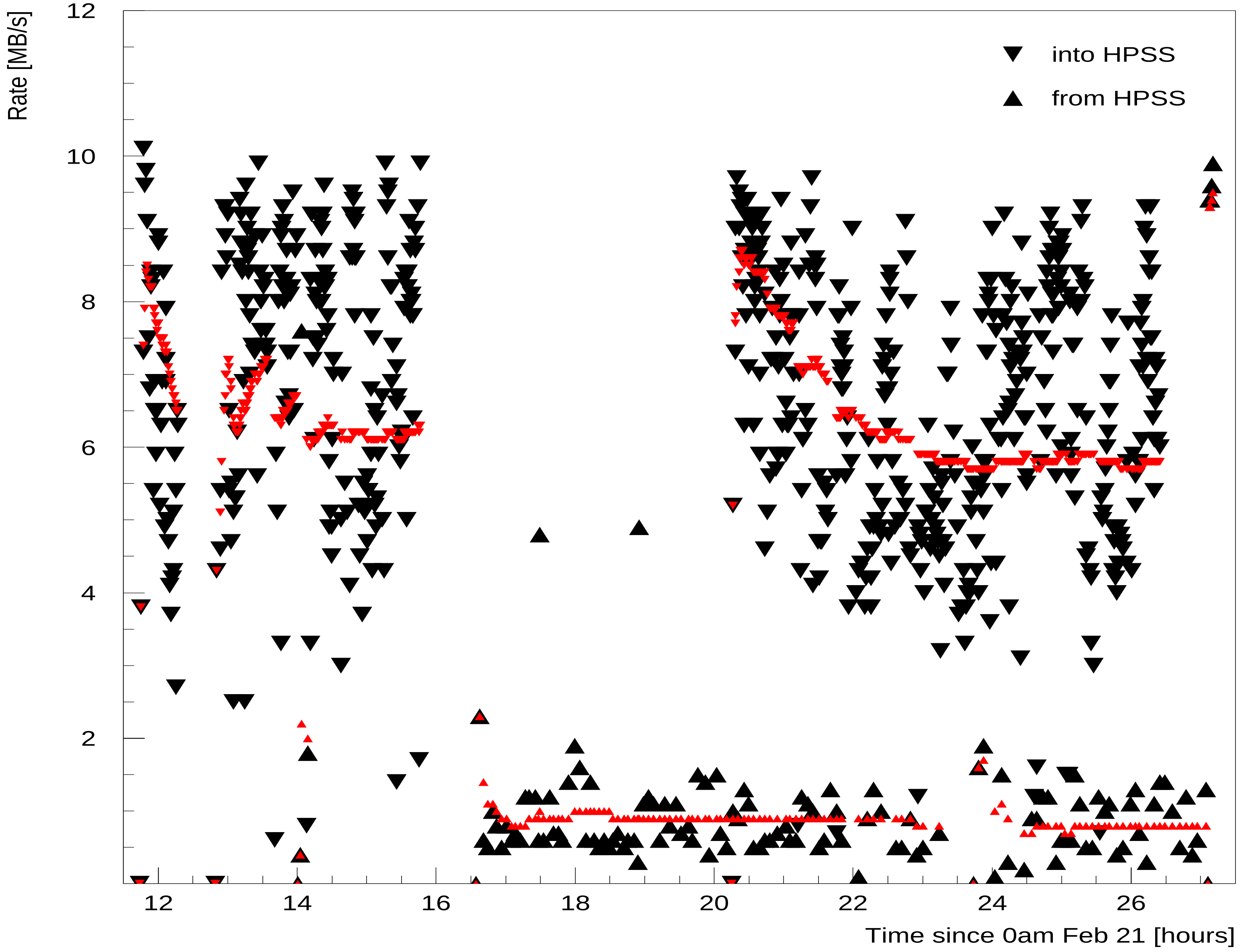
Data selection

Physics **D**ata Sets (**PhDS**)

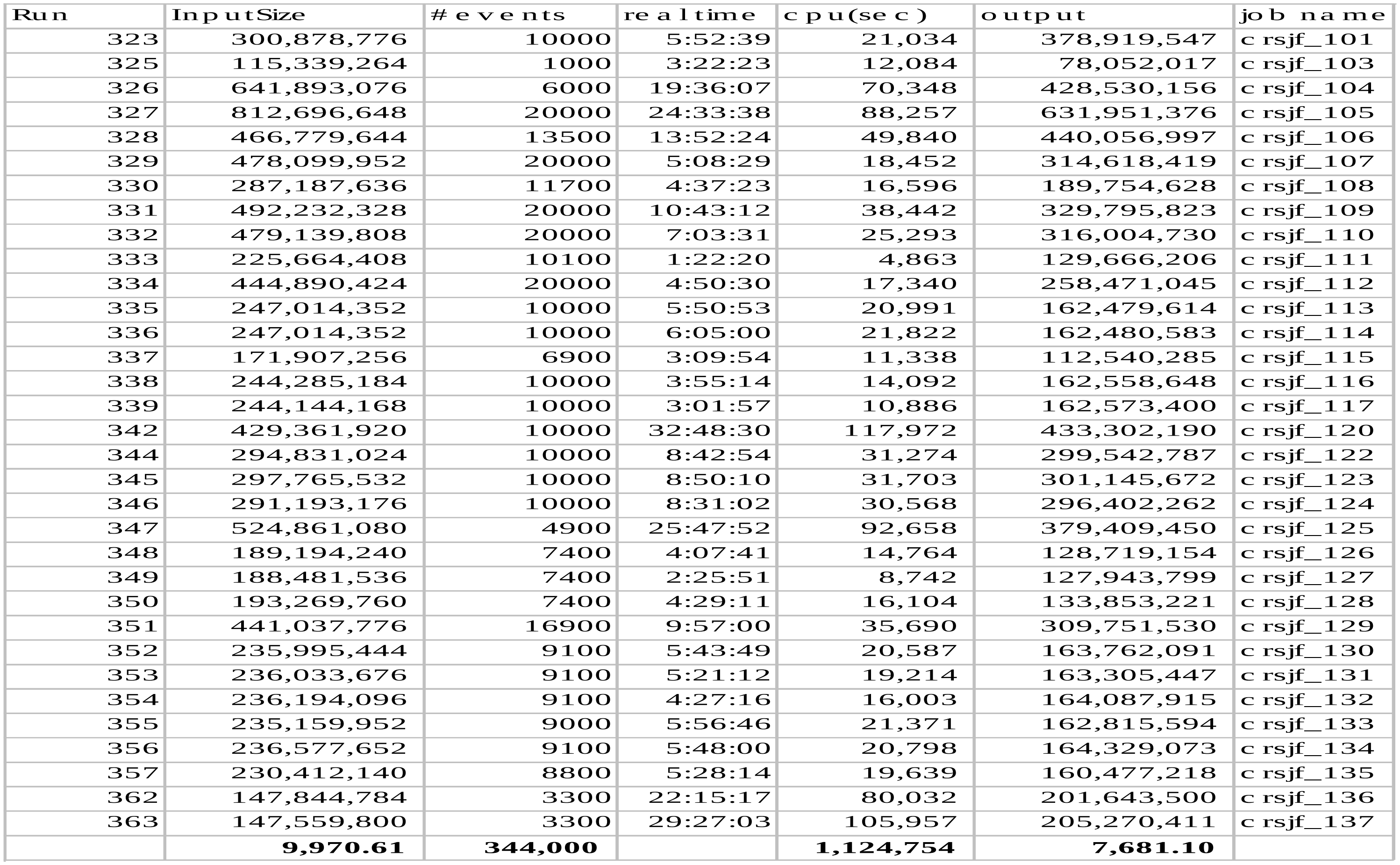
Data mining and Analysis

- The simple data mining and data analysis code were developed for MDC-I were essentially also used for MDC-II. They were judged to be quite typical.
- Additionally 'experiments' were made with more 'fixed' Tree structures that allows for much faster access to parts of a given events thus achieving the very high nominal access speeds for ROOT objects.
- Data mining, was done by selecting events with full spectrometer tracks. For central events at forward angles this reduced data volumes by about 1:3. With the present event model (sequential reading of hierarchical structure of root objects) the data mining is close to being dominated by CPU usage not i/o (for single task).
- One particular sequence of task entailed reading from different files, but executing the same code on 4*4 parallel jobs on the CAS nodes over a 3 hours sequence. An aggregated reading speed of 4Mb/sec was achieved on the Linux CA farms.
- The following slide show transfers (sinking) into HPSS and data mining reading during a active period towards the end of MDC-II. Note the saturation for writing into at about 6 Mb/s.

Sinking from and extracting to rsun00



Sample statistics



Run	InputSize	#events	realtime	cpu(sec)	output	job name
323	300,878,776	10000	5:52:39	21,034	378,919,547	crsjf_101
325	115,339,264	1000	3:22:23	12,084	78,052,017	crsjf_103
326	641,893,076	6000	19:36:07	70,348	428,530,156	crsjf_104
327	812,696,648	20000	24:33:38	88,257	631,951,376	crsjf_105
328	466,779,644	13500	13:52:24	49,840	440,056,997	crsjf_106
329	478,099,952	20000	5:08:29	18,452	314,618,419	crsjf_107
330	287,187,636	11700	4:37:23	16,596	189,754,628	crsjf_108
331	492,232,328	20000	10:43:12	38,442	329,795,823	crsjf_109
332	479,139,808	20000	7:03:31	25,293	316,004,730	crsjf_110
333	225,664,408	10100	1:22:20	4,863	129,666,206	crsjf_111
334	444,890,424	20000	4:50:30	17,340	258,471,045	crsjf_112
335	247,014,352	10000	5:50:53	20,991	162,479,614	crsjf_113
336	247,014,352	10000	6:05:00	21,822	162,480,583	crsjf_114
337	171,907,256	6900	3:09:54	11,338	112,540,285	crsjf_115
338	244,285,184	10000	3:55:14	14,092	162,558,648	crsjf_116
339	244,144,168	10000	3:01:57	10,886	162,573,400	crsjf_117
342	429,361,920	10000	32:48:30	117,972	433,302,190	crsjf_120
344	294,831,024	10000	8:42:54	31,274	299,542,787	crsjf_122
345	297,765,532	10000	8:50:10	31,703	301,145,672	crsjf_123
346	291,193,176	10000	8:31:02	30,568	296,402,262	crsjf_124
347	524,861,080	4900	25:47:52	92,658	379,409,450	crsjf_125
348	189,194,240	7400	4:07:41	14,764	128,719,154	crsjf_126
349	188,481,536	7400	2:25:51	8,742	127,943,799	crsjf_127
350	193,269,760	7400	4:29:11	16,104	133,853,221	crsjf_128
351	441,037,776	16900	9:57:00	35,690	309,751,530	crsjf_129
352	235,995,444	9100	5:43:49	20,587	163,762,091	crsjf_130
353	236,033,676	9100	5:21:12	19,214	163,305,447	crsjf_131
354	236,194,096	9100	4:27:16	16,003	164,087,915	crsjf_132
355	235,159,952	9000	5:56:46	21,371	162,815,594	crsjf_133
356	236,577,652	9100	5:48:00	20,798	164,329,073	crsjf_134
357	230,412,140	8800	5:28:14	19,639	160,477,218	crsjf_135
362	147,844,784	3300	22:15:17	80,032	201,643,500	crsjf_136
363	147,559,800	3300	29:27:03	105,957	205,270,411	crsjf_137
	9,970.61	**344,000**		**1,124,754**	**7,681.10**	

MDC-II Conclusions

- MDC-II was a useful with a concentrated effort during a short time. The preparation and startup was hampered by availability of final scripts not until just-in-time and severe network connectivity during this period.
- Increased stability of HPSS systems is needed. More tools for automated handling of errors are needed. Better access to
- Brahms still needs to develop additional scripts to make submission, checking of in particular (calibration), reconstruction and data mining. It is necessary to examine how to incorporate calibration and reconstruction in the CRS environment.).
- We were able to handle an increased number of events (Min Bias) satisfactory in the reconstruction in the order of 340,000 events in spectrometer different configurations.
- Gained experience with (p)ftp access to RCF.
- MDC-II was useful in identifying areas of improvements
  - Data selection
  - simplification of DST data model that makes better use of root trees.
- We did not really get around to study the ‘competing’ access to HPSS (raw/reco/mining)

Misc.

- LSF farm software for batch generation is still reasonable. RCF is missing tools to keep it up running (went dead in a couple of instances, without being detected). An example of not having tools to monitor the state of all systems, and report to users.
- The division of the rlnx03 has worked well for us so far, but there will be a need to have a dedicated Linux development machine.
- BRAHMS BRAT progress (K.Hagel, A.Kjetil,JHLee,..)
  - DB Detetor Parameters API defined
  - ROOT based event display
  - Revised Object model for spectrometertracks
  - Improved OO model for TPC data, and gain in reconstruction (*4-5)
  - Review and revision of mDST data structures.