

High performance computing activities in hadron spectroscopy at BESIII

Beijiang Liu (for BESIII collaboration)

Institute of High Energy Physics, CAS, Beijing, China

E-mail: liubj@ihep.ac.cn

Abstract. BESIII experiment is operated in the τ -charm threshold energy region. It has collected the world's largest data samples of J/ψ , $\psi(3686)$, $\psi(3770)$ and $\psi(4040)$ decays. These data are being used to make a variety of interesting and unique studies of light hadron spectroscopy, precision charmonium physics and high-statistics measurements of D meson decays. As one of the experiments at the high luminosity frontier, data processing at BESIII is computationally very expensive for large data sets. In this presentation, we report two recent progresses in using high performance computing: a Tag-based preselection for data reduction and GPUPWA, a PWA framework harnessing the GPU parallel computing.

1. Introduction

The fundamental theory of the strong interactions (Quantum Chromodynamics, QCD) predicts that there should exist new forms of matter, such as glueballs (pure-gluon objects) and hybrids (qqbar states with explicit gluon) [1]. Experimental search of these predictions and subsequent investigation of their properties would provide validation of and valuable input to the quantitative understanding of QCD. The experimental study of hadron spectroscopy is to systematically map out all the resonances with the determination of their properties like mass, width, spin-parity as well as partial decays widths. The key to success lies in high levels of precision during the measurement and high statistics in the recorded data set complemented with sophisticated analysis methods.

BESIII (Beijing Spectrometer) [2] is a new state-of-the-art 4π detector at the upgraded BEPCII (Beijing Electron and Positron Collider) [3] that operated in the τ -charm threshold energy region between (about) 2 and 4.6 GeV with a designed luminosity of $10 \times 10^{33} \text{cm}^2 \text{s}^{-1}$. The BESIII experimental program addresses issues in light hadron physics, charmonium spectroscopy and decays, D and Ds meson decays, and numerous topics in QCD and τ physics [4]. To date, BESIII has accumulated data samples corresponding to 1.3 billion J/ψ decays, 0.6 billion $\psi(3686)$ decays and 2.9 fb at the peak of the $\psi(3770)$ resonance, which decays to $D\bar{D}$ meson pairs nearly 100% of the time. These are all world's-largest data samples at these c.m. energies and the J/ψ sample is the first ever to be collected in a high quality detector like BESIII.

Generally, a typical physics analysis can be categorized into three phases as shown in Table 1.

In this presentation, we report two recent progresses of using high performance computing in BESIII physics analysis: a Tag-based pre-selection for data reduction and the GPUPWA, a PWA framework for high statistics partial wave analyses harnessing the GPU parallel computing.

Table 1. Summary of the systematic errors from the event selection.

Phase	Purpose	Task	Bottle neck
Algorithmic Analysis	Reduction: event selection, record useful information	Processing reconstruction data	I/O, CPU
Interactive Analysis	Final selection, plotting, interactive studies	Processing Ntuple/Tree	I/O
Statistical Analysis	Extract physics	Fit	CPU

2. TAG-based preselection

The typical branching fractions for interesting physics channels are of the order of $O(10^{-3})$ [4, 5]. The traditional event-wise accessing of BOSS (Bes Offline Software System) [6] is not effective for the selective accessing with the rate of $O(10^{-3})$. Event tags (TAG) are event-level metadata which support efficient identification and selection of events of interest to a given analysis. TAG content includes, for example, run number, high level trigger decisions, event properties (like visible energy), and particle data (photons, charged tracks, electrons, muons, pions, kaons, protons, *etc*). To facilitate queries for event selection, TAG data can be stored in ROOT files or a relational database. The current TAG size is approximately 0.1 kilobyte per event. The computing cost of pre-selection with TAGs is much less than the pre-selection with the full reconstruction events. TAGs allow jobs to process only those reconstruction events that satisfy a given predicate. In the current implementation, the TAG data for every event is stored in the same reconstruction file as an individual ROOT tree. C++ modules are written using the TAG attributes for event selection. No payload data is retrieved for unselected events and data files containing only unselected events are not accessed. To improve the I/O performance of the event selection jobs, Hadoop and HDFS [7] could be employed for job management and distributed storage. The analysis algorithms are able to access the TAG data and reconstruction data from the local disks.

3. GPUPWA

Extracting resonance properties from experimental data is however far from straightforward; resonances tend to be broad and plentiful, leading to intricate interference patterns. In such an environment, simple fitting of mass spectra is usually not sufficient and a partial wave analysis (PWA) is required to disentangle interference effects and to extract resonance properties. In the cases discussed here, the full kinematic information is used and fitted to a model of the amplitude in a partial wave decomposition. The partial wave amplitude is constructed with an angular part and a dynamical part. The model parameters are determined by an unbinned likelihood fit to the data, while the event-wise efficiency correction is included. In a typical PWA (we use the radiative decay $J/\psi \rightarrow \gamma\eta\eta$ [12] as an example), the quasi two-body decay amplitudes (isobar model) in the sequential decay process $J/\psi \rightarrow \gamma X, X \rightarrow \eta\eta$, are constructed using covariant tensor amplitudes described in Ref. [8]. The basis of the likelihood fitting is that a hypothesized probability density function (PDF) would produce the data set under consideration. The probability to observe the event characterized by the measurement ξ is :

$$P(\xi) = \frac{\omega(\xi)\epsilon(\xi)}{\int d\xi\omega(\xi)\epsilon(\xi)}, \quad (1)$$

where $\epsilon(\xi)$ is the detection efficiency and $\omega(\xi) \equiv \frac{d\sigma}{d\Phi}$ is the differential cross section, and $d\Phi$ is the standard element of phase space.

$$\frac{d\sigma}{d\Phi} = \left| \sum_j^{N_{wave}} \Lambda_j A_j(\xi) \right|^2, \quad (2)$$

where A_j is the partial wave amplitude with coupling strength determined by a complex coefficient Λ_j . The normalization integral is performed numerically by Monte Carlo techniques. The likelihood for a particular model is

$$\mathcal{L} = \prod_{i=1}^{N_{data}} P(\xi_i). \quad (3)$$

A series of likelihood fits are performed for parameter estimation and model evaluation. In the log likelihood calculation, the likelihood value of background events are given negative weights, and are removed from data since the log likelihood value of data is the sum of signal and background.

Most computing problems in particle physics are trivially parallel, as data consists of independent events. The traditional approach to parallelism is thus to treat different subsets of events on different cores/machines/in different locations and then have a relatively lightweight piece of code perform a synthesis. For PWA, this synthesis (namely the likelihood sum) has to be performed very frequently (for every fit iteration) and thus network latency quickly dominates for distributed architectures. As the calculations per event are relatively simple and exactly the same for all events, it is desirable to have a very large number of (simple) cores available in a single machine.

GPUPWA [9] has been developed as the working framework of BESIII, harnessing GPU parallel acceleration. GPUPWA is now developed with *OpenCL* [10] as described in [11]. The performance is improved by O(100) of magnitude comparing to FORTRAN implementation used in BES II (without multi-threading).

The framework now provides facilities for amplitude calculation, minimization and plotting and is widely used for analyses (e.g. [12, 13]) at BES III. It continues to be developed and is available at [14]. A cluster of 35 computing nodes with 140 GPUs is established to serve the PWA at BESIII.

4. Summary

Now five years from our first collisions, BESIII has established a broad and successful program in charm physics. Recently, in 2012, even larger samples have been accumulated at the J/ψ and $\psi(3686)$; total samples are now about 1.2 billion and 0.35 billion decays, respectively. Furthermore, our 2013 dataset includes more data near 4260 MeV, and also a large sample at the $Y(4360)$. With the excellent performance of the accelerator and detector, more interesting results are expected. High performance computing techniques will play a more and more important role in the physics analysis of high statistics experimental data.

5. Acknowledgements

This presentation is partly supported by Youth Innovation Promotion Association, CAS.

References

- [1] Close F E 1988 *Rept. Prog. Phys.* **51** 833
- Godfrey S and Napolitano J 1999 *Rev. Mod. Phys.* **71** 1411
- Amsler C and Tornqvist N A 2004 *Phys. Rept.* **389** 61
- Klempf E and Zaitsev A 2007 *Phys. Rept.* **454**, 1
- Crede V and Meyer C A 2009 *Prog. Part. Nucl. Phys.* **63** 74

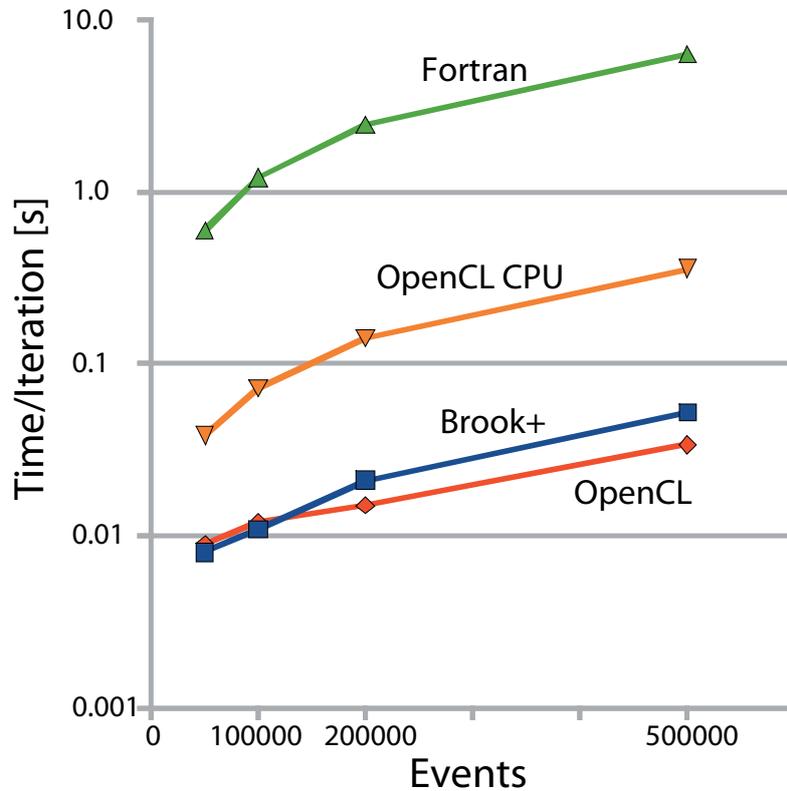


Figure 1. Comparison of the running time of one fit iteration of the benchmark $J/\psi \rightarrow \gamma K^+ K^-$ analysis in different implementations on the same machine. Both *OpenCL* datasets were obtained using the AMD implementation. The detailed information for the setup is described in Ref. [11].

[2] Ablikim M *et al* (BESIII Collaboration) 2010 *Nucl. Instrum. Meth. A* **614** 345
 [3] Bai J *Zet al* (BES Collaboration) 1994 *Nucl. Instrum. Meth. A* **344** 319
 Bai J *Zet al* (BES Collaboration) 2001 *Nucl. Instrum. Meth. A* **458** 627
 [4] Asner D M, Barnes T, Bian J M, Bigi I I, Brambilla N, Boyko I R, Bytev V and Chao K T *et al* 2009 *Int. J. Mod. Phys. A* **24**, S1
 [5] Beringer J *et al* (Particle Data Group) 2012 *Phys. Rev. D* **86** 010001
 [6] Li W 2006 *Proc.Int.Conf.Comput.High Energy and Nucl.Phys* 225
 [7] <http://hadoop.apache.org/>
 [8] Zou B S and Bugg D V 2003 *Eur. Phys. J. A* **16** 537
 [9] Berger N Liu B J and Wang J K 2010 *J. Phys. Conf. Ser.* **219**, 042031
 [10] <http://www.khronos.org/opencv/>
 [11] Berger N 2011 *AIP Conf. Proc.* **1374** 553
 [12] Ablikim M *et al* [BESIII Collaboration] 2013 *Phys. Rev. D.* **87** 092009
 [13] Ablikim M *et al* [BESIII Collaboration] 2012 *Phys. Rev. Lett.* **108** 112003
 [14] GPUPWA package, available at <http://sourceforge.net/projects/gpupwa/>