

РЕОРГАНИЗАЦИЯ ПРОГРАММНОГО КОМПЛЕКСА SONIX+ ДЛЯ РАБОТЫ С ДАННЫМИ В ВИДЕ СПИСКА СОБЫТИЙ

А. С. Кирилов¹, И. А. Морковников,

С. М. Мурашкевич, Т. Б. Петухова, Л. А. Трунтова

Объединенный институт ядерных исследований, Дубна

Применение на измерительных установках реактора ИБР-2 новых DAQ-контроллеров, в которых отсутствует аппаратное гистограммирование данных, в ближайшие годы примет все более массовый характер. При этом происходит замена первичной формы представления данных в виде гистограмм в пользу накопления в виде списка отдельных событий, что соответствует общемировой тенденции. Это повлечет за собой определенную перестройку комплекса Sonix+. При этом, с одной стороны, гистограммы следует сохранить для оперативной визуализации и для программ обработки, которые специфичны для каждой установки. С другой стороны, снимается аппаратное ограничение по дискретизации при построении этих гистограмм, что, как ожидается, на практике приведет к существенному увеличению размеров гистограмм, в том числе превышающему лимит доступного виртуального адресного пространства для 32-разрядных приложений. В работе предложен подход к разрешению этой проблемы.

In the coming years, the application of new DAQ controllers in which there is no hardware data histogramming will become more and more widespread at the IBR-2 instruments. At the same time, the primary data presentation in the histogram form is replaced in favor of accumulation in the form of a list of individual events, which corresponds to the global trend. This will entail a certain restructuring of the Sonix+ complex. Thus, on the one hand, histograms should be kept for operative visualization, as well as for data reduction programs, which are specific for each instrument. On the other hand, the hardware sampling limitation in the construction of these histograms is removed, which is expected to lead to a significant increase in histogram size, including exceeding the limit of available virtual address space for 32-bit applications in practice. In this paper, an approach to solve this problem is proposed.

PACS: 29.50.+v; 28.41.-i

ВВЕДЕНИЕ

В последние годы в нейтронных центрах, применяющих времяпролетную методику измерений, растет интерес к работе с экспериментальными данными в форме списка событий (list-mode, event-mode) [1]. Ранее традиционной формой предоставления результатов измерения пользователям были массивы со спектральным распределением

¹E-mail: akirilov@jinr.ru

интенсивностей (гистограммы), причем сам процесс гистограммирования в основном выполнялся аппаратно с заранее заданными параметрами по временному и пространственному разрешению. При этом происходило определенное статистическое усреднение исходных данных, но результат был предельно компактен. Напротив, исходные данные в формате event-mode содержат максимально детализированную информацию об измерении, которую можно сортировать и гистограммировать с произвольной выборкой и разрешением.

Среди возможных, на наш взгляд, особый интерес представляет подход, принятый в коллаборации ESS (Лунд, Швеция), ISIS (Дидкот, Великобритания), PSI (Виллиген, Швейцария) [2–4]. С использованием системы EPICS [5] для управления устройствами и Mantid [6] для визуализации и первичной обработки данных предусматривает передачу данных DAQ в форме событий и параметров окружения образца и запись их в формате Nexus [7]. С учетом того, что EPICS является распределенной системой, а также отказа от аппаратного вето в процессе накопления, ответственность за корректную, без потерь, передачу данных в системе возрастает. Для решения той задачи предложено организовать конвейер передачи данных на основе брокера Apache kafka [8], адаптируя соответствующие компоненты в EPICS и Mantid. При этом для оперативной визуализации и перевода данных в гистограммную форму используются готовые средства Mantid. Аналогичный подход был позднее принят в PSI [9, 10] с заменой Apache kafka на Redpanda [11].

В последнее время на установках ИБР-2 в ЛНФ ОИЯИ также началось применение DAQ-контроллеров с выходными данными в виде списка событий [12, 13], причем замена старых контроллеров с аппаратным гистограммированием на новые варианты в ближайшие годы будет продолжена. Это должно получить соответствующую поддержку в программном комплексе Sonix+ [14], используемом для управления измерениями в лаборатории. Рассмотрим изменения в комплексе, которые позволяют организовать работу с данными в форме списков событий с учетом особенностей организации измерений в ЛНФ. Следует отметить, что на ряде установок ИБР-2 (ФДВР, ФСД, РТД) запись данных в форме списка событий применялась достаточно давно, однако поддержка со стороны Sonix+ ограничивалась только записью данных.

АНАЛИЗ ТЕКУЩЕЙ СИТУАЦИИ

Ситуация с программным обеспечением измерительных установок на реакторе ИБР-2 ЛНФ и в упомянутых выше центрах по ряду моментов отличается существенно.

Если в ESS и PSI принято концептуальное решение о переходе к форме представления данных в виде списка событий в качестве замены гистограмм и под эту концепцию подбираются соответствующие аппаратные и программные средства, то в ЛНФ ОИЯИ ориентация на гистограммы остается прежней. Это определяет входной формат и отчасти функционал программ обработки измеренных данных, которые разработаны и сопровождаются самими пользовательскими группами.

Кроме того, в отличие от EPICS Sonix+ не является распределенной системой. За редким исключением для управления экспериментом используется единственный управляющий компьютер, к которому подключены все DAQ-контроллеры. Поэтому проблем с возможной потерей при сборе данных DAQ нет.

С учетом вышеизложенного до формулирования пользователями новых требований необходимости в радикальных изменениях структуры Sonix+ в настоящее время нет. Для адаптации комплекса к работе с новым типом данных достаточно:

- перевести DAQ-компоненты Sonix+ и средства визуализации в 64-разрядное виртуальное адресное пространство для работы с массивами более 2 ГБ;
- разработать модули для гистограммирования данных для всех типов контроллеров;
- доработать форматы записи данных для гистограмм размерности больше 4 ГБ, а также для записи данных в виде списка событий.

Обсудим вопрос необходимости расширения адресного пространства компонент. Sonix+ реализован как набор 32-разрядных модулей (которые далее для краткости и наглядности будем обозначать $\times 32$), и, следовательно, все его компоненты располагают виртуальным адресным пространством 2 ГБ для программных кодов и данных. Пока гистограммирование выполнялось аппаратно, этого было достаточно, так как максимальный размер запрашиваемых гистограмм для данных с двухкоординатных позиционно-чувствительных детекторов (2D ПЧД) на установках реактора ИБР-2 был существенно меньше 2 ГБ (таблица).

Ситуация изменилась при попытках использования диджитайзера N6730 фирмы CAEN [15, 16]. Получив возможность задавать произвольное разрешение, пользователи столкнулись с ошибками при попытках визуализации гистограмм из-за превышения лимита виртуального адресного пространства программ комплекса.

Рассмотрим два графика (рис. 1, 2), построенных с разным разрешением программой SpectraViewer [17] на основе данных измерения прямого пучка, выполненного на 2D ПЧД-детекторе на установке «Грейнс» в сентябре 2021 г. На них изображена центральная часть карты плотности (X , ToF), полученная при суммировании данных по оси Y (X , Y — пространственные каналы регистрации, ToF — временные каналы регистрации, заполняемые по методу времени пролета). Одномерные графики представляют результат суммирования данных по осям ToF и Y (справа вверх)

**Размерность гистограмм для данных
с двухкоординатных позиционно-чувствительных детекторов (2D ПЧД),
использованных на установках реактора ИБР-2 в мае 2021 г.**

Установка	Тип DAQ-контроллера	Ширина канала, мкс	ToF	X	Y	Длина гистограммы, МБ
Грейнс	DeLiDAQ-2	128	1024	128	256	128
ФДВР	DeLiDAQ-2	64	1024	128	128	64
Рефлекс	DeLiDAQ-1	192	668	160	80	32,6
Ремур	DeLiDAQ-2	256	512	256	128	64
РТД	DeLiDAQ-2	64	204	32	64	16
	DeLiDAQ-1	64	2048	16	32	4
ЮМО	—	128	512	256	256	128

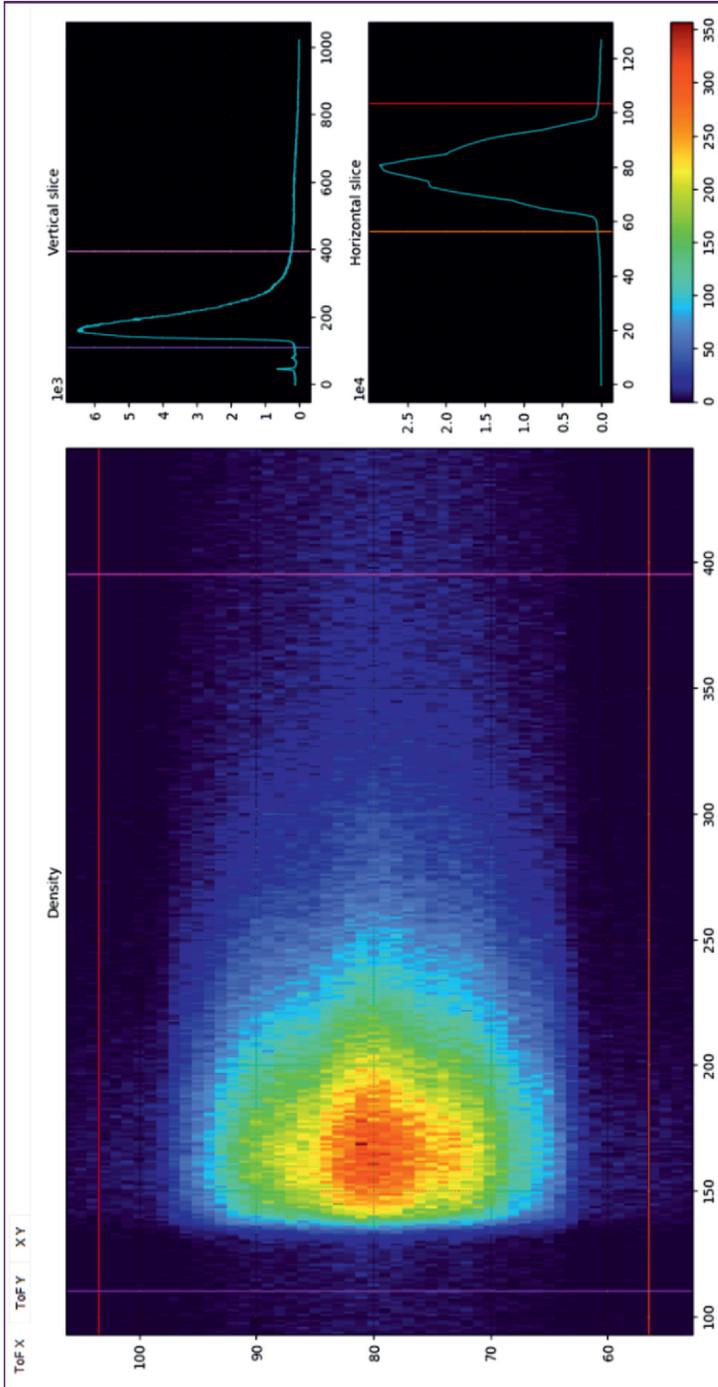


Рис. 1. Графики прямого пучка установки «Грейнс» (10-й канал), построенные с разрешением ToF = 1024, X = 128, Y = 256. Размер гистограммы 128 МБ. Все поверхности и кривые выглядят гладкими

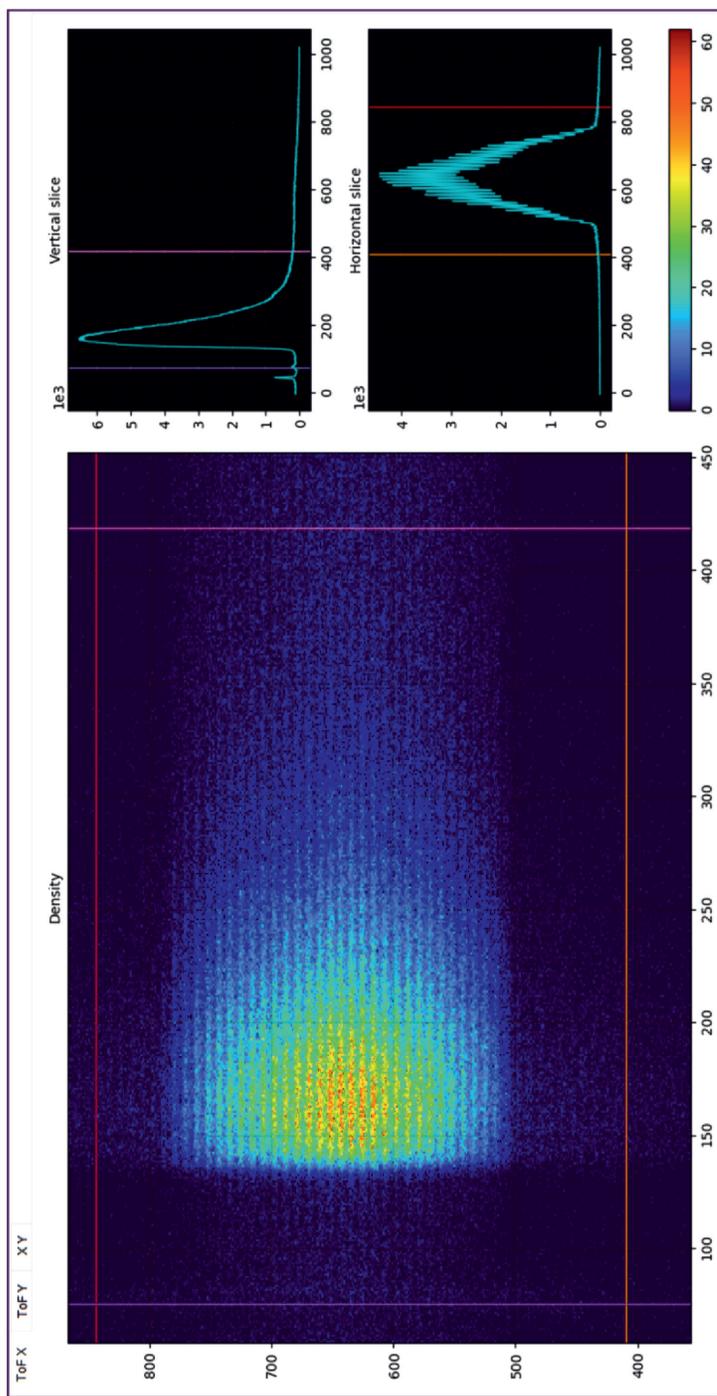


Рис. 2. Графики для того же измерения, построенные с разрешением $\text{ToF} = 1024$, $X = 1024$, $Y = 1000$. Размер гистограммы 4 000 МБ. Явственно видна структура регистрационных элементов детектора

и аналогично — результат суммирования по осям ToF и X (справа внизу). Первая гистограмма размером 128 МБ построена с помощью 32-разрядной версии Python (далее Python32), вторая — размером 4000 МБ с помощью 64-разрядной версии Python (далее Python64). На первом графике все кривые и поверхности гладкие, на втором явно видна структура регистрационных элементов детектора.

С учетом данных таблицы разрешение на рис.2 является избыточным, однако консультации с пользователями подтвердили интерес к гистограммам с подобными разрешениями. Кроме того, на установках ИБР-2 в течение 3–5 лет планируется подключение новых детекторов с большей площадью, для которых более высокое разрешение будет востребованным.

СТРУКТУРА DAQ Sonix+ И «УЗКИЕ МЕСТА» В НЕЙ

Последовательность обработки данных DAQ в Sonix+ приведена на рис. 3. Данные от контроллеров запрашиваются драйверами и далее передаются на следующий уровень. При необходимости драйверы формируют файлы со списком событий. В конце экспозиции сервер S_expo запрашивает данные в виде гистограмм от всех драйверов и формирует общий файл в формате Sonix+, который далее в скрипте трансформируется в набор файлов конкретного спектрометра. Сервер S_spectra выполняет аналогичную работу по запросу программы оперативной визуализации SpectraViewer.

Все модули нижнего уровня (драйверы и серверы) написаны на языке C++ и оформлены как приложения $\times 32$.

Программа визуализации данных SpectraViewer и остальной пользовательский интерфейс (GUI) написаны на языке Python и запускаются через его 32-разрядный интерпретатор.

В комплексе можно отметить три «узких места» при работе с «большими» массивами:

- ограниченность размера внутренних буферов прежде всего в серверах S_expo и S_spectra;

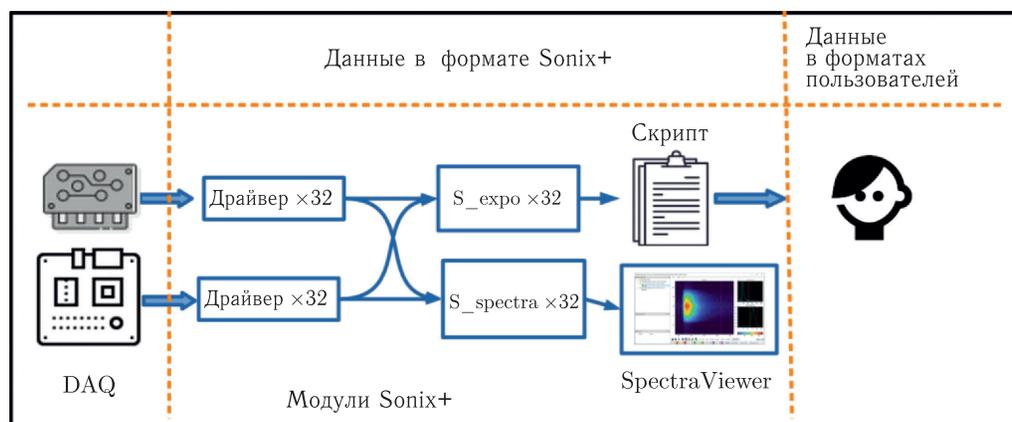


Рис. 3. Структура конвейера обработки данных в Sonix+

- аналогичная ситуация с программой SpectraViewer;
- недостаточность выбранной размерности длины гистограммы в 4 Б в формате Sonix+.

ВОЗМОЖНЫЕ ПОДХОДЫ И ОСНОВНАЯ ИДЕЯ РЕШЕНИЯ

В связи с тем, что на управляющих компьютерах установлена 64-разрядная версия системы, а поддержка подсистемы WoW64 (Windows-on-Windows 64-bit), позволяющей запускать 32-разрядные приложения, не лимитирована, ограничение размеров массивов данных является фактически единственным, которое необходимо устранить.

Лобовое решение проблемы по полному переводу комплекса в $\times 64$ сопряжено с трудностями перевода ядра комплекса на C++ [18] и не выглядит рациональным. Тем более, что значительная часть компонентов комплекса, написанная на языке Python (GUI, визуализация, системные скрипты Sonix+), может работать под Python64 и такого перевода не требует.

Вместе с тем очевидно, что, поскольку при работе с файлами событий гистограммирование выполняется программно после измерения, можно решить эту проблему, выделив перевод данных в формат гистограмм, а также их визуализацию в $\times 64$. При этом управляющая часть комплекса на C++ с небольшой модификацией сохраняется в $\times 32$.

Основной сложностью в подобном подходе является разработка нового средства взаимодействия между модулями, написанными на C++, и модулями, написанными на Python, поскольку коммуникационная библиотека динамической компоновки, которая обеспечивала эту связь, работает только в $\times 32$ и не импортируется в Python64.

Реорганизация комплекса выглядит следующим образом. Детекторы (DAQ-контроллеры) в режиме event-mode выделяются в отдельный класс. Данные от этих детекторов записываются в особом формате (рис. 4).

Тип контроллера	Имя детектора	Список параметров для гистограммирования	Данные
-----------------	---------------	--	--------

Рис. 4. Формат записи данных в event-mode

Здесь и далее цветом выделены поля, которых не было ранее. Список параметров для гистограммирования зависит от типа контроллера и включает все необходимые параметры помимо разрешения.

Для гистограммирования данных каждого контроллера разрабатывается отдельная программа-конвертер как приложение $\times 64$. Размерность гистограмм и имена входного и выходного файлов задаются через строку параметров конвертера. Конвертеры могут вызываться как автоматически из скрипта в процессе проведения эксперимента, так и вручную из специализированного GUI по запросу пользователя.

Каждая из конвертированных гистограмм сохраняется в отдельном файле по формату на рис. 5.

Необходимые изменения для работы с новым классом детекторов вносятся в программы драйверов, серверов и SpectraViewer. При этом отработка данных с остальных детекторов (контроллеров) сохраняется неизменной.

Признак формата	Имя детектора	Число детекторов	Размерность (ToF, X, Y)	Длина (8 Б)	Гистограмма
-----------------	---------------	------------------	-------------------------	-------------	-------------

Рис. 5. Формат записи конвертированной гистограммы

Для связи ядра Sonix+ и модулей на языке Python была адаптирована служба обмена сообщениями PyChannel [19], что явилось самым важным изменением в структуре комплекса при его реорганизации.

ОПЕРАТИВНАЯ ВИЗУАЛИЗАЦИЯ

В Sonix+ для визуализации измеренных данных предназначена написанная на языке Python программа SpectraViewer [17]. Существенным изменением при ее адаптации к Python64 стало использование службы обмена сообщениями PyChannel для передачи сообщений от SpectraViewer модулю S_spectra и обратно. Кроме того, для работы с данными в виде списка событий в нее были добавлены вызовы соответствующих конвертеров. При этом схема оперативной визуализации не претерпела существенных изменений. Пользователь может посмотреть график текущего измерения в любой момент времени. По его запросу программа посылает запрос модулю S+spectra с именем интересующего детектора. По этом запросу S_spectra запрашивает данные от драйвера и формирует файл данных, который передается для визуализации. Если данные контроллера допускают гистограммирование «на лету», то гистограмма строится в драйвере непосредственно в процессе накопления и время выполнения запроса минимально. В остальных случаях гистограммирование выполняется в S_spectra с помощью соответствующего конвертера и время реакции несколько увеличивается. При длительности эксперимента от получаса до суток такой вариант оперативной визуализации для пользователей вполне приемлем.

GUI ДЛЯ РУЧНОГО КОНВЕРТИРОВАНИЯ

В процессе эксперимента конвертирование данных при формировании выходных данных в формате пользователя выполняется автоматически с заданными параметрами в скрипте. При этом данные сохраняются и в виде гистограмм, и в первичной форме в виде списка событий. Это дает пользователю возможность произвести гистограммирование с нужными параметрами самостоятельно после эксперимента. Для этого на языке Python разработана программа EventModeToHistogram, интерфейс которой приведен на рис. 6. Размерность гистограммы задается пользователем, а значения специфических параметров извлекаются из файла данных. После запуска программа проверяет корректность заданных параметров и в случае успеха выполняет гистограммирование.

Список конвертеров вместе с необходимым для каждого составом параметров вынесен в конфигурационный файл программы, чем обеспечивается ее универсальность.

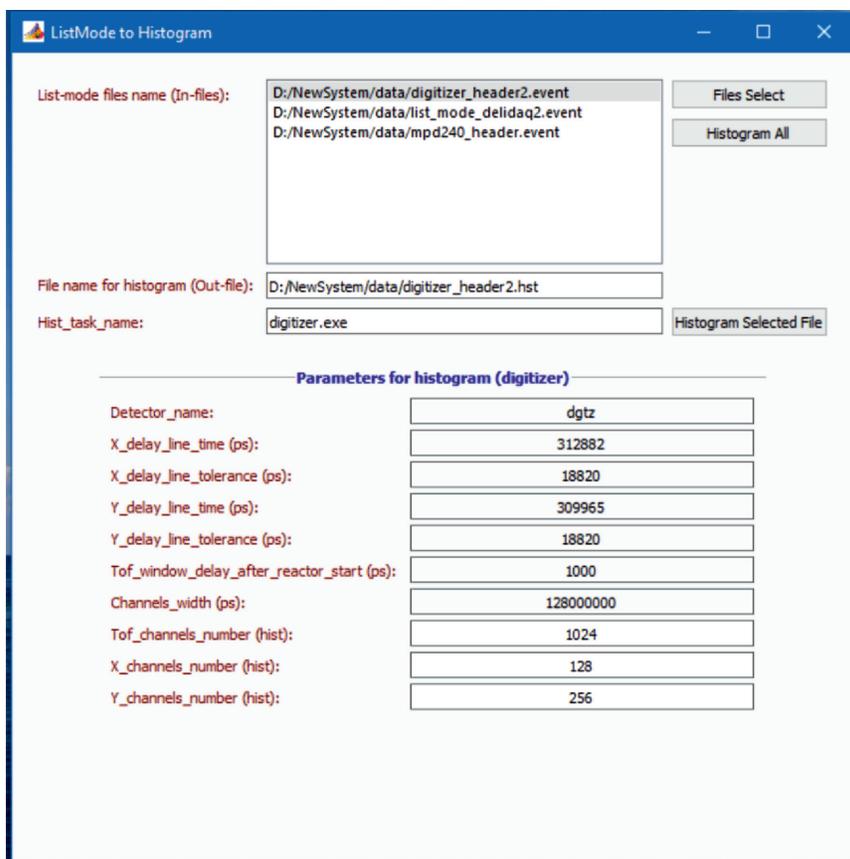


Рис. 6. Пользовательский интерфейс программы EventModeToHistogram для ручного гистограммирования

ЗАКЛЮЧЕНИЕ

Применение DAQ-контроллеров с выходными данными в виде списка событий вместо массивов со спектральным распределением интенсивностей (гистограмм) и отсутствием функции аппаратного гистограммирования, начавшееся в последнее время на установках ИБР-2 в ЛНФ ОИЯИ, потребовало определенной реорганизации программного комплекса Sonix+. Вместе с тем общая ориентация на использование гистограмм для дальнейшей обработки в программах пользователей остается неизменной. В работе рассмотрено решение, которое позволяет адаптировать Sonix+ к новой ситуации с минимальными затратами.

Авторы выражают признательность В. И. Боднарчуку за поддержку работы и ценные замечания при подготовке данной рукописи. Мы также благодарны рецензенту, критические замечания которого позволили более четко выразить основную цель данной работы.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ (соглашение №075-10-2021-115 от 13 октября 2021 г. (внутренний номер 15.СИН.21.0021)).

СПИСОК ЛИТЕРАТУРЫ

1. *Peterson P. F., Campbell S. I., Reuter M. A., Taylor R. J., Zikovskiy J.* Event-Based Processing of Neutron Scattering Data. arXiv:1507.01838 [physics.ins-det].
2. *Gahl T., Hagen M., Hall-Wilton R., Kolya S., Koennecke M., Rescic M., Rod T. H., Sutton I., Trahern G., Kirstein O.* Hardware Aspects, Modularity and Integration of an Event Mode Data Acquisition and Instrument Control for the European Spallation Source (ESS) // Proc. of ICANS XXI, Mito, Japan, 2014. P. 497–502.
3. *Clarke M. J., Akeroyd F. A., Moore L. A., Arnold O., Draper N. J., Gigg M., Jones M. D., Richter T. S.* Live Visualization of Experiment Data at ISIS and the ESS // Proc. of the 16th Intern. Conf. on Accelerator and Large Exp. Control Systems (ICALEPCS2017), Barcelona, Spain, 2017. P. 431–434.
4. *Jones M. D., Akeroyd F. A., Arnold O., Clarke M. J., Draper N., Gigg M., Moore L., Richter T.* A New Design for Live Neutron Event Data Visualisation for ISIS and the ESS // J. Phys.: Conf. Ser. 2018. V. 1021. 012013.
5. <https://epics.anl.gov/>
6. <https://www.mantidproject.org/>
7. <https://www.nexusformat.org/>
8. <https://kafka.apache.org/>
9. *Brambilla M.* Data Streaming at SINQ. https://indico.psi.ch/event/12738/contributions/38879/attachments/22754/40115/SINQ_streaming_Michele.pdf
10. <https://www.youtube.com/watch?app=desktop&v=0bw07N-izza>
11. <https://redpanda.com>
12. *Levchanovskiy F. V., Murashkevich S. M.* // Proc. of the XXIV Intern. Symp. on Nucl. Electronics and Computing (NEC'2013), Varna, Bulgaria, Sept. 9–16, 2013. P. 176–179.
13. *Levchanovskiy F. V., Murashkevich S. M.* // Phys. Part. Nucl. Lett. 2016. V. 13, No. 5. P. 591–594.
14. <https://sonix.jinr.ru/wiki/doku.php?id=ru:index>
15. <https://www.caen.it/products/n6730/>
16. *Кирилов А. С., Мурашкевич С. М.* Адаптация программного комплекса Sonix+ для работы с DAQ-контроллерами DeLiDAQ-2 и диджитайзером N6730 фирмы CAEN. Препринт ОИЯИ P10-2022-5. Дубна, 2022.
17. *Кирилов А. С.* Новые версии программ юстировки и визуализации спектров для рефлектометров ИБР-2 // Письма в ЭЧАЯ. 2016. Т. 13, № 1(199). С. 208.
18. *Карпов А.* 7 шагов по переносу программы на 64-битную систему. <https://pvs-studio.ru/ru/blog/posts/cpp/a0042/>.
19. *Морковников И. А., Трунтова Л. А.* Служба обмена сообщениями PyChannel в программном комплексе Sonix+. Препринт ОИЯИ P10-2023-27. Дубна, 2023.

Получено 15 июня 2023 г.