

На правах рукописи

Чернов Василий Геннадьевич

Разработка распределенной системы
сбора данных и анализ формы
импульса событий
на установке «Троицк ню-масс»

01.04.01—«Приборы и методы экспериментальной физики»

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Москва — 2019

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте ядерных исследований Российской академии наук.

Научный руководитель: *Нозик Александр Аркадьевич*, кандидат физико-математических наук, Федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (национальный исследовательский университет)», лаборатория методов ядерно-физических экспериментов, старший научный сотрудник.

Официальные оппоненты:

1. *Крянев Александр Витальевич*, доктор физико-математических наук, Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский ядерный университет «МИФИ», кафедра прикладной математики, профессор.

2. *Белов Владимир Александрович*, кандидат физико-математических наук, Федеральное государственное бюджетное учреждение «Институт теоретической и экспериментальной физики имени А.И.Алиханова Национального исследовательского центра «Курчатовский институт», лаборатория сильных взаимодействий, старший научный сотрудник.

Ведущая организация:

Объединенный институт ядерных исследований, Дубна.

Защита состоится 23 апреля 2020 г. в 12.30 час. на заседании диссертационного совета Д002.119.01 при Федеральном государственном бюджетном учреждении науки Институте ядерных исследований Российской академии наук по адресу: 117312, Москва, проспект 60-летия Октября, 7А.

С диссертацией можно ознакомиться в библиотеке и на сайте ИЯИ РАН по адресу: <http://www.inr.ru/rus/referat/dis-zasch.html#che>.

Отзывы на автореферат в двух экземплярах, заверенные печатью учреждения, просьба направлять по адресу: 117312, Москва, проспект 60-летия Октября, 7А, ученому секретарю диссертационного совета Д002.119.01.

Автореферат разослан _____.

Ученый секретарь диссертационного совета Д 002.119.01,

канд. физ.-мат. наук

Демидов Сергей Владимирович

Общая характеристика работы

Актуальность темы

На сегодняшний день нейтринные исследования относятся к одной из самых активных и перспективных областей физики элементарных частиц. Экспериментальное открытие нейтринных осцилляций и смешиваний^{1,2} свидетельствует о выходе физики нейтрино за рамки Стандартной модели. Изучение свойств этой группы фундаментальных частиц открывает возможность прямого исследования явлений новой физики.

Экспериментально полученные ограничения на параметры нейтрино позволяют предположить существование еще одного вида частиц - стерильных (обладающих правой хиральностью) нейтрино. Большинство расширений Стандартной модели предполагают существование одного или нескольких видов такой частицы с модельно-зависимой массой. Аргументы в пользу гипотезы существования стерильных нейтрино:

- существование стерильного нейтрино естественным образом объясняет наличие массы активных нейтрино;
- для всех остальных фермионов были обнаружены как левосторонние, так и правосторонние частицы.

Эксперимент «Тройцк ню-масс» с 2012 года присоединился к поискам стерильного нейтрино [1]. Установка, несмотря на изначальное проектирование под измерение массы активного нейтрино, может быть использована и для поиска стерильных нейтрино с массой в диапазоне единиц кэВ. Эта область масс представляет интерес, т. к. стерильное нейтрино с массой порядка нескольких кэВ является одним из кандидатов на роль частиц темной материи³. Поиск стерильных нейтрино требует модификации установки. Необходимые улучшения описаны в [2]. Данная работа касается одного из основных необходимых улучшений - повышения чувствительности системы считывания сигнала.

Целью данной работы является совершенствование системы сбора данных установки «Тройцк ню-масс» и разработка алгоритмов разделения наложений при непрерывной оцифровке.

Для достижения поставленной цели необходимо было решить следующие **задачи**:

1. Провести общую модернизацию системы сбора данных. Разработать модульную архитектуру, позволяющую проводить изолированную разработку и отладку отдельных подсистем, в

¹ W. Walter, C. The Super-Kamiokande Experiment / C. W. Walter. 2008. Mar.

²The Sudbury Neutrino Observatory / A. Bellerive [et al.] // Nuclear Physics B. 2016. Apr. Vol. 908.

³A White Paper on keV Sterile Neutrino Dark Matter / R. Adhikari [et al.] // ArXiv e-prints. 2016. Feb. Vol. 2017.

- частности считывания сигнала. Адаптировать существующий алгоритм управления установкой под новую архитектуру. Добавить инструменты контроля качества набора в реальном времени. Необходимость разработки распределенной архитектуры вызвана масштабом системы сбора данных. Работа по уменьшению мертвого времени требует тестирования и отладки методов считывания сигнала. В силу сложности системы единственный способ эффективной разработки программного обеспечения - разбиение его на независимые подсистемы и разработка каждого модуля по отдельности. Дополнительными преимуществами модульности являются: возможность быстрой замены и одновременного использования нескольких вариантов считывающей электроники, возможность стандартизации интерфейса взаимодействия между подсистемами и использование единого транспортного протокола.
2. Провести исследование доступных вариантов альтернативного считывания сигнала. Оценить возможность использования в качестве замены исходной подсистемы считывания для установки «Троицк ню-масс». Встроить наиболее подходящий вариант в систему сбора данных. Реализовать для системы новый модуль, осуществляющий автоматизированный контроль считывания. Протестировать созданную подсистему и убедиться в ее корректной работе.
 3. Разработать алгоритмы обработки считанных данных, позволяющие работать с сигналами с высокой скоростью счета.

Научная новизна:

1. С учетом современных подходов к проектированию сложных программных комплексов^{4,5} разработана распределенная система сбора данных установки «Троицк ню-масс». В основе архитектуры лежит распределение обработки по независимым модулям, каждый из которых работает независимо и контролирует отдельную подсистему обработки. Модули имеют единообразный программный интерфейс и взаимодействуют через стандартный стек TCP/IP. Аппаратно модернизация основана на использовании встраиваемых ПК-контроллеров ССРС7⁶ производства ОИЯИ (г. Дубна) вместо обычных контроллеров КАМАК, подключающихся к ПК.

⁴Microservices: Yesterday, Today, and Tomorrow / N. Dragoni [et al.] // Present and Ulterior Software Engineering. Springer International Publishing, 2017. P. 195–216.

⁵*Al-Debagy, O. A Comparative Review of Microservices and Monolithic Architectures / O. Al-Debagy, P. Martinek. 2019. eprint: [arXiv:1905.07997](https://arxiv.org/abs/1905.07997).*

⁶AFI Electronics: ССРС7. URL: <https://afi.jinr.ru/CCPC7>.

2. В рамках проектирования системы сбора данных разработан оригинальный формат передачи и хранения данных, оптимизированный под задачи экспериментальной физики⁷.
3. Разработан оригинальный алгоритм разделения наложенных сигналов, основанный на форме событий. С помощью алгоритма получено эффективное мертвое время порядка 0.9 мкс при длине сигнала 6 мкс и одного канала оцифровки 320 нс. Помимо качества разделения, существенной особенностью алгоритма является отсутствие привязки к конкретной форме сигнала.

Практическая значимость. Распределенные системы контроля и сбора данных представляют существенный интерес для физических экспериментов. Успешный опыт использования этой концепции в эксперименте «Троицк ню-масс» будет применен при разработке новых систем. Распределенная система хранения данных и обмена сообщениями также представляет интерес как для экспериментальной физики, так и для промышленного использования.

Использование разработанных алгоритмов по разделению наложенных сигналов в рамках эксперимента «Троицк ню-масс» позволило добиться уменьшения эффективного мертвого времени в семь раз и, таким образом, обеспечить возможность работы на скорости счета вплоть до 50-60 кГц на канал, что позволит проводить измерения бета-спектра в широком диапазоне и с большой статистикой. Существенным является то, что алгоритмы не привязаны к конкретной форме сигнала и могут быть адаптированы для использования в других экспериментах.

Методология и методы исследования. При проектировании системы сбора данных проводилось исследование архитектуры комплексных систем с открытым исходным кодом. Проводился анализ подходов в существующих проектах, определялись их преимущества и возможность применения в разрабатываемой системе сбора. Для реализации компонент системы проводился поиск готового и поддерживаемого инструментария с открытым исходным кодом. В целом критериями выбора того или иного подхода и фреймворка были: его популярность, простота использования и степень распространенности в проектах.

При разработке методов разделения наложенных сигналов был проведен анализ существующих решений, изложенных в статьях и препринтах. Для оценки качества работы алгоритмов был разработан генератор событий, симулирующий реальный сигнал детектора. Адекватность работы генератора оценивалась с помощью сравнения статистических параметров с реальными данными. Техника определения метрик качества генерации данных и алгоритмы обработки описаны в соответствующих главах работы.

⁷DataForge Streaming envelope format documentation. 2019. URL: <http://npm.mipt.ru/dataforge/docs.html#envelopes>.

Основные положения, выносимые на защиту:

1. Проведена общая модернизация системы сбора данных установки «Троицк ню-масс». Выполнен переход от монолитной архитектуры исходной системы к микросервисной архитектуре. Основной код обработки был переписан на C++/Qt. Модернизированная система разбита на три независимых модуля, отвечающих за: считывание сигнала с детектора, управление напряжением спектрометра и проведение набора по сценарию. Модули имеют единообразный программный интерфейс и взаимодействуют через стек TCP/IP. В основе аппаратной модернизации лежит использование ПК-контроллеров крейтов КАМАК ССРС7 для управления подсистемами установки. Каждый модуль представляет собой программный сервер, который устанавливается в ССРС7 и осуществляет контроль аппаратуры через магистраль крейта, RS-232, RS-485 интерфейсы и другие специфические соединения.
2. Осуществлен переход на запись и обработку непрерывных кадров сигнала. Для модернизированной системы сбора разработан модуль считывания с помощью АЦП Лан10-12РС1. В процессе записи плата последовательно сбрасывает кадры максимальной длины по программному триггеру. Каждый последующий кадр сбрасывается сразу после сохранения предыдущего. Таким образом, сохраняется непрерывная оцифровка сигнала с пропусками на время сбросов. Данный подход позволяет решить проблему аппаратного мертвого времени, возникающую при стандартном наборе по триггеру, и проводить более сложную обработку сигнала в офлайн-режиме.
3. В рамках сотрудничества с группой TRISTAN⁸ из Института физики имени Макса Планка в Мюнхене, к системе сбора данных «Троицк ню-масс» был подключен прототип детектора TRISTAN, разрабатываемого для KATRIN⁹. Управление детектором осуществлялось через устройство обработки сигнала DANTE производства XGLab, Италия. Для него разработан эмулятор сигналов, работающий по предоставленной спецификации. При помощи эмулятора создан программный пакет, обеспечивающий взаимодействие с DANTE, управление считыванием, а также хранение событий вместе со специфическими для этой системы метаданными. При помощи этой системы проведены два сеанса измерений на

⁸A novel detector system for KATRIN to search for keV-scale sterile neutrinos / S. Mertens [et al.] // Journal of Physics G: Nuclear and Particle Physics. 2019. May. Vol. 46, no. 6. P. 065203.

⁹KATRIN Collaboration. KATRIN design report 2004 : tech. rep. / KATRIN Collaboration, KATRIN Collaboration ; Forschungszentrum, Karlsruhe. 2005. 245 p. 51.54.01; LK 01.

установке с использованием детектора TRISTAN, по результатам которых была подготовлена совместные статьи [3] [4].

4. Разработаны алгоритмы выделения параметров событий (амплитуды и положения по времени) из непрерывных кадров с фиксированной длиной, полученных с платы Лан10-12РСІ. Для данных «Троицк ню-масс» при выделении параметров событий удалось добиться эффективного мертвого времени порядка 0.9 мкс для средней длины сигнала 6 мкс при шаге оцифровки 320 нс. Мертвое время при использовании исходной аппаратной обработки событий составляло около 7 мкс. Учет формы импульса при обработке позволил исключить систематическую ошибку восстановления амплитуд, вызванную наложением на хвосты предыдущих событий. Проведено обобщение алгоритмов для использования с произвольной формой импульса. Производительность алгоритма оптимизирована до уровня, когда он может быть использован в режиме реального времени. Исходный код с инструкциями по воспроизведению результатов выложен в открытый репозиторий.

Достоверность полученных результатов:

1. Правильность выбора архитектуры системы сбора данных подтверждается широким использованием выбранных подходов, как в коммерческих проектах, так и в проектах с открытым исходным кодом. Действенность используемых инструментов подтверждается их популярностью среди разработчиков программного обеспечения. Эффективность распределенной системы сбора данных была продемонстрирована при горячей замене отдельных подсистем и отдельных компьютеров. Такую замену практически невозможно осуществить в традиционной монолитной архитектуре.
2. Эффективность алгоритмов разделения наложений наглядно продемонстрирована в диссертации как при помощи всестороннего тестирования на сгенерированных данных, так и при помощи амплитудного и временного анализа результатов обработки реальных данных.

Апробация работы. Результаты диссертации докладывались на VII и VIII межинститутских молодежных конференциях «Физика элементарных частиц и космология», международной конференции «The XXI International Scientific Conference of Young Scientists and Specialists» и ряде научных семинаров в ИЯИ и КИТ, Карлсруэ.

Личный вклад

Разработка и тестирование всех программных компонентов, а также подготовка к публикации материалов по теме алгоритмов разделения наложенных событий проводились непосредственно автором. Также автор

принял активное участие в шести сеансах работы «Троицк ню-масс», по итогам которых были получены результаты мирового уровня:

- наилучшее ограничение на примесь стерильного нейтрино с массой в районе 1 кэВ [1],
- исследование и набор физических данных с помощью детектора нового поколения TRISTAN [3].

Содержание диссертации и основные положения, выносимые на защиту, отражают персональный вклад автора в работу.

Публикации. Основные результаты по теме диссертации изложены в шести печатных изданиях, рекомендованных ВАК.

Содержание работы

Во введении обосновывается актуальность исследований, проводимых в рамках данной диссертационной работы, формулируется цель, ставятся задачи работы, излагается научная новизна и практическая значимость представляемой работы.

В первой главе приведен обзор текущих экспериментов по поиску стерильных нейтрино. Кратко описываются их основные типы и метод работы.

Во второй главе приводится описание установки «Троицк ню-масс». Описываются ее основные компоненты: беззаконный газовый тритиевый источник, электростатический спектрометр, криогенная система и электронная пушка. Приводится описание методики набора данных с алгоритмической и аппаратной сторон.

Типичный сценарий сеанса набора выглядит следующим образом:

1. Установка выводится на рабочий режим. С помощью криогенной системы охлаждаются сверхпроводящие соленоиды спектрометра, производится термостабилизация объема источника. С помощью дифференциальной системы откачки в объеме спектрометра создается необходимый уровень вакуума.
2. Через систему ввода и очистки трития в объем источника напускается газообразный тритий.
3. Проводится набор интегрального спектра бета-распада трития. Для этого проводится последовательный набор точек спектра. Точка набирается следующим образом:
 - а) С помощью высоковольтной системы на электрод спектрометра подается запирающее напряжение.
 - б) В течение фиксированного отрезка времени производится запись зарегистрированных детектором событий.

Электрон в спектрометре, в случае если его энергия превышает выставленное запирающее напряжение, проходит потенциальный

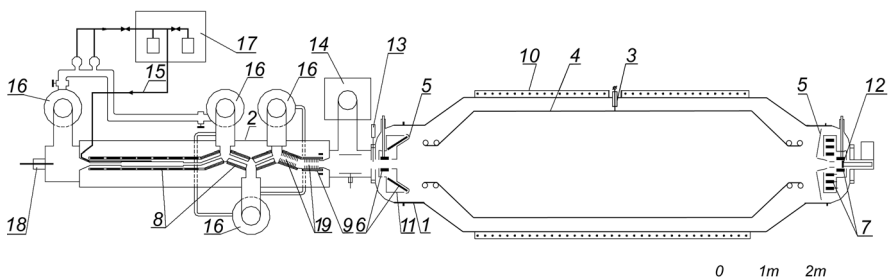


Рис. 1 — Схема установки: 1 - вакуумный объем спектрометра; 2 - вакуумный объем источника; 3 - высоковольтный ввод; 4 - электрод спектрометра; 5 - заземляющие электроды; 6-9 - сверхпроводящие катушки; 10 - индукционные обмотки для дополнительной коррекции; 11 - охлаждающий кожух с жидким азотом; 12 - детектор, охлажденный до температуры жидкого азота; 13 - аварийный шибер; 14 - магниторазрядный насос; 15 - тритиевый контур; 16 - ртутные диффузионные насосы; 17 - система ввода и очистки трития; 18 - электронная пушка; 19 - аргоновая ловушка.

барьер и попадает в установленный за ним детектор. Электроны с более низкими энергиями отражаются от потенциального барьера и не доходят до детектора. Проведя серию наборов событий при разных значениях запирающего напряжения, мы получим и интегральный спектр энергий электронов бета-распада трития.

В третьей главе описана архитектура системы сбора данных. Глава разделена на две подглавы. В первой из них описывается разработанный формат хранения и передачи данных DataForge Envelope. Излагаются требования, которые учитывались при разработке, правила формирования пакета и преимущества формата. Проводится сравнение с существующими похожими форматами.

На рисунке 3 схематично изображена структура пакета:

1. Бинарный заголовок фиксированной длины, содержащий служебную информацию, такую как размер пакета и его внутренних частей.
2. Опциональный блок метаданных в человекочитаемом формате.
3. Опциональный блок бинарных данных.

Преимущества данного формата:

1. Структура формата оптимизирована под особенности экспериментальных данных. Типичная единица данных при наборе - комбинация параметров набора (напряжение, время набора, температуры и т. д.) и самих набираемых данных. Объединение обеих частей в одном пакете выглядит естественно и затрудняет потерю части данных.

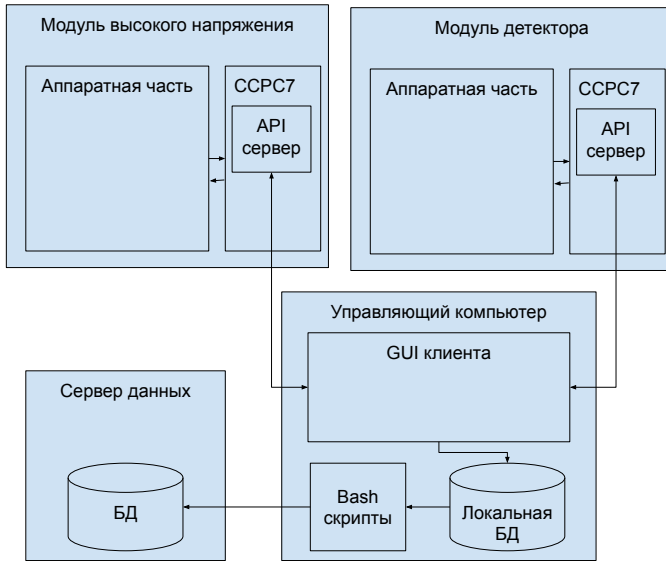


Рис. 2 — Распределенная система сбора данных.

2. Человекочитаемый формат метаданных позволяет оценить содержание пакета без использования дополнительных средств.
3. Встроенная система разделения пакетов позволяет создавать вложенные структуры данных и работать с потоком.
4. Отдельный блок метаданных может быть прочитан без парсинга всего пакета, что позволяет проводить быструю обработку, поиск и фильтрацию сообщений.

Во второй подглаве описывается разработанная автором распределенная система сбора данных установки «Троицк ню-масс». Разбирается работа каждого из модулей системы - описывается взаимодействие с аппаратной частью и логика работы и интерфейс управления. В основе архитектуры лежит микросервисный подход, при котором система разделяется на несколько максимально простых независимых модулей. В нашем случае система построена по следующим правилам:

1. Каждый модуль инкапсулирует в себе низкоуровневое аппаратное управление.
2. Взаимодействие происходит путем обмена сообщениями формата DataForge Envelope по стеку TCP/IP.
3. Интерфейс взаимодействия модуля включает только высокоуровневые команды, полностью отвязанные от аппаратной части.

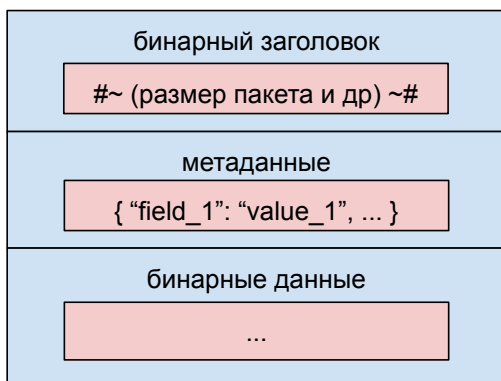


Рис. 3 — Формат сообщения DataForge Envelope.

Разработанная система схематично изображена на рисунке 2. Она имеет два микросервиса: высоковольтной стойки и детектора. Управляющая программа имеет графический интерфейс и с помощью высокоуровневых команд («выставить напряжение» для микросервиса высоковольтной стойки и «записать события за промежуток времени» для микросервиса детектора) контролирует процесс набора. Благодаря API каждый модуль системы сбора становится слабо связанным с другими модулями и может быть использован самостоятельно. Это значительно облегчает разработку, обслуживание и замену элементов системы. Отсутствие привязки API к конкретной аппаратной части позволяет использовать сервисы в виртуальном режиме, симулирующем реальную работу.

Использование программно-независимого формата DataForge Envelope допускает возможность написания компонент системы на разных языках программирования.

В четвертой главе представлены два примера расширения системы сбора данных: модуль АЦП Лан10-12PCI и модуль прототипа детектора TRISTAN из Института физики имени Макса Планка в Мюнхене.

Схема подключения для Лан10-12PCI представлена на рисунке 4. Создан новый микросервис, а сама плата располагается в отдельном компьютере и через разветвитель получает такой же сигнал, как и оригинальная система.

Новый микросервис работает как сниффер (программа, которая прозрачно встраивается в сетевой поток и перехватывает трафик) между стандартным детекторным блоком и управляющей программой. При получении команды от управляющей программы сервер передает ее в оригинальную систему без изменений. Ответы на команды, приходящие от оригинальной системы, также прозрачно транслируются обратно в управ-

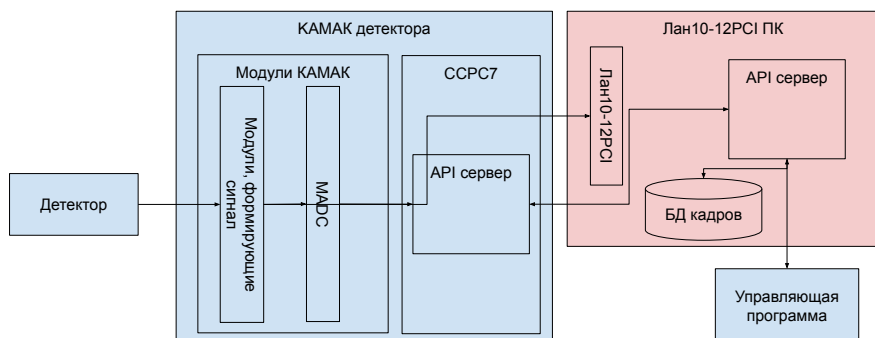


Рис. 4 — Подключение платы Лан10-12РСІ. Розовым цветом обозначены новые модули системы.

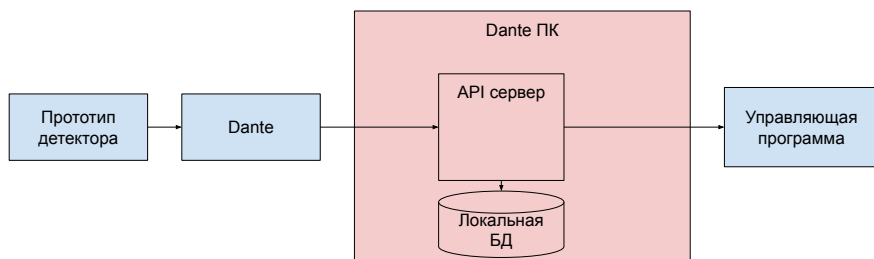


Рис. 5 — Интеграция DANTE в систему сбора данных.

ляющую программу. В случае если через сервер проходит команда на набор событий, помимо трансляции команды программа также запустит набор кадров через плату Лан10-12РСІ с таким же временем набора. После завершения набора обеими системами в управляющую программу будет передано ответное сообщение, аналогичное ответу на набор точки от оригинальной системы.

Пример Лан10-12РСІ показывает гибкость архитектуры системы сбора данных. Интеграция платы не потребовала изменения других компонент системы. Для работы в дублирующем режиме в управляющей программе необходимо только изменить IP-адрес и порт сервиса детектора. Также просто можно отключить новую систему.

В случае с прототипом детектора TRISTAN, управление детектором осуществлялось через устройство обработки сигнала DANTE производства XGLab, Италия. Схема подключения модуля к системе сбора представлена на рисунке 5. В этом случае новый микросервис полностью заменяет исходный.

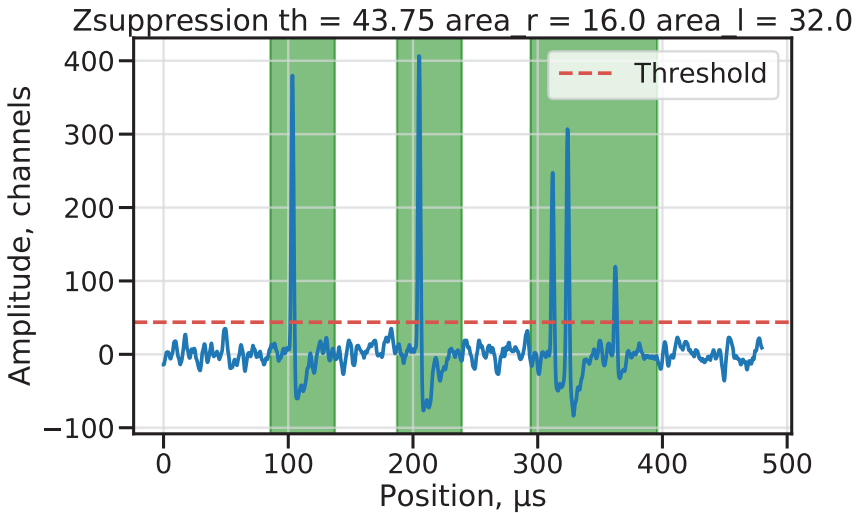


Рис. 6 — Алгоритм zero-suppression.

Пример TRISTAN показывает универсальность архитектуры разработанной системы сбора данных. Тщательное тестирование программного обеспечения DANTE выявило некоторые проблемы. При подключении через стандартный драйвер устройства результат набора устройством содержал ошибки. DANTE работал корректно только в случае управления через стандартную графическую программу, существовавшую в единственном экземпляре в одном из привезенных вместе с ней ноутбуков. Однако детектор был успешно встроен в систему сбора данных с помощью микросервиса-кликера, имитирующего человеческий ввод через клавиатуру и мышь.

В пятой главе представлены разработанные автором алгоритмы выделения параметров событий из непрерывного сигнала и разделения наложений. Обосновывается необходимость разработки алгоритмов разделения событий. Приводятся характеристики платы Лан10-12PCI и метод записи кадров непрерывного сигнала. Описываются принцип работы и результаты валидации обычного и обобщенного генераторов сигналов. Описывается методика измерения эффективности алгоритмов. Исследуются пять разработанных алгоритмов выделения.

Пример кадра непрерывного сигнала изображен на рисунке 6 (используются только выделенные области, остальные части кадра обрезаются предобработкой). Формы событий извлечены из данных путем группировки событий по значению амплитуды и последующего усреднения форм по группам. Пример усреднения по группам показан на рисунке 7.

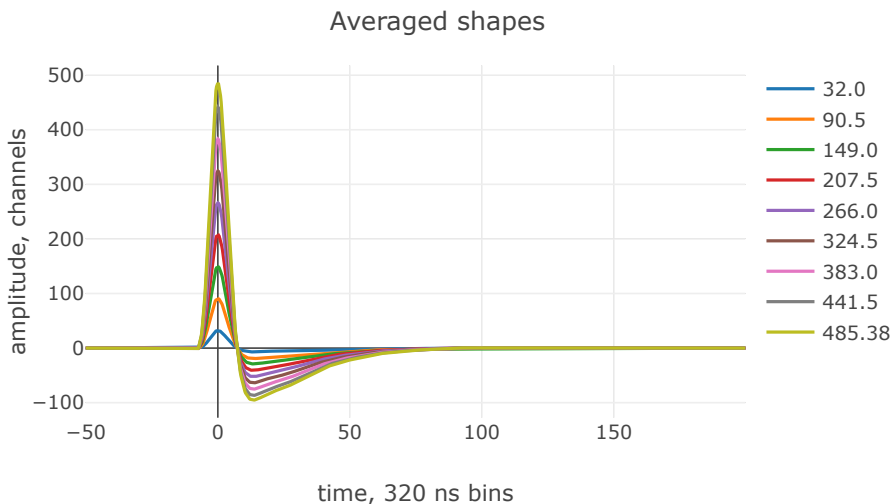


Рис. 7 — Усредненные формы событий.

Для проверки эффективности алгоритмов выделения используются размеченные кадры, созданные генератором. Принцип работы генератора:

1. Для генерации события используется интерполяция по сетке усредненных форм.
2. Шумовой сигнал собирается из фрагментов реального сигнала, вырезанного из начала кадров.

Для определения соответствий между исходными реконструированными событиями используются следующие правила (изображены на рисунке 8):

1. Реконструированное событие соответствует исходному, если:
 - а) текущее реконструированное событие - ближайшее к исходному по сравнению с остальными реконструированными событиями;
 - б) разница между положениями пика не превышает 10 бинов (т. е. 3.2 мкс).
2. Исходное событие считается наложенным, если:
 - а) разница между положениями пика исходного и реконструированного событий не превышает 10 бинов;
 - б) соответствующее реконструированное событие имеет другое, более близкое исходное событие.
3. Исходное событие считается нераспознанным, если для него не существует реконструированного события с положением пика, отличным от исходного менее чем на 10 бинов.

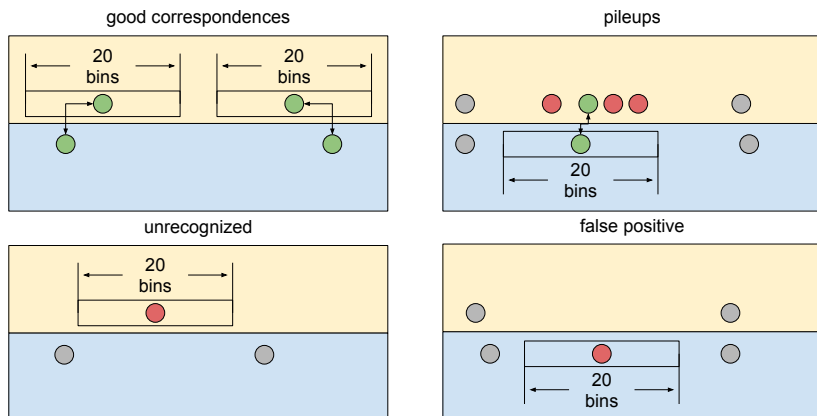


Рис. 8 — Правила определения соответствий между событиями. Обозначения: желтый прямоугольник содержит исходные события, голубой - восстановленные, зелеными кружками отмечены взаимно соответствующие события, красным - нераспознанные\ложнораспознанные\наложенные события в зависимости от примера, серыми - не относящиеся к примеру события, изображенные для наглядности.

4. Реконструированное событие считается ложнораспознанным, если для него не существует исходного события с положением пика, отличным от реконструированного менее чем на 10 бинов.

Наилучший результат дает алгоритм выделения с фитированием двумя разными способами. При средней длине события 6 мкс и размере одного канала оцифровки 320 нс алгоритм имеет эффективное мертвое время 0.9 мкс. Схема работы:

1. Запись текущего состояния в стек.
2. Поиск самого левого локального максимума выше порога (максимум определяется как самый большой бин в плавающем окне размером пять бинов).
3. Извлечение параметров события с помощью параболического фитирования по пяти точкам в окрестности локального максимума.
4. Сохранение полученных параметров.
5. Вычитание из кадра реконструированной по параметрам формы события.
6. Проверка очередности последних двух событий. Переход к п. 1, если очередность соблюдена или расстояние между событиями больше восьми бинов. В противном случае:

- а) Откат состояния до этапа, на котором выделенное событие было левее текущего. Одновременно ищутся самое левое и самое правое события, отмененные откатом.
- б) Выделение двойного наложения фитированием. Функция - сумма двух форм событий, область интерполяции - промежуток между крайними отмененными событиями.
- в) Запись извлеченной пары событий.
- г) Вычитание извлеченных событий из кадра.
- д) Переход к п. 1.

Блок-схема алгоритма изображена на рисунке 9. Обработка использует стек состояний, хранящий историю предыдущих этапов и возможность отката на предыдущие состояния. При откате назад выделенные на текущем этапе события удаляются из массива результатов, а их формы возвращаются в кадр. Т. е. при нарушении очередности алгоритм возвращает состояния до последнего правильного и меняет способ выделения на фитирование одновременно двух событий по полной форме.

В **заключении** приведены основные результаты работы:

1. Проведена общая модернизация системы сбора данных установки «Троицк нью-масс». Разработана модульная архитектура, позволяющая проводить изолированную разработку и отладку отдельных подсистем. Исходный алгоритм управления переписан на C++/Qt и перенесен на новую архитектуру. Разработаны инструменты визуального контроля качества набора в реальном времени. Разработан и реализован оригинальный формат передачи и хранения данных DataForge Envelope, оптимизированный под задачи экспериментальной физики¹⁰. Написаны реализации протокола на C++(Qt), Python, Java. Преимущества формата:
 - Единица данных - пакет, содержащий метаданные и бинарные данные. Метаданные имеют текстовый формат и используются для хранения параметров набора и настроек. Объединение метаданных и бинарных данных в один пакет исключает потерю части набранных данных при манипуляциях с ними.
 - DataForge Envelope может использоваться как для хранения, так и для передачи данных. В качестве хранилища используется файловое дерево, каждый пакет записывается в виде файла, таким образом для просмотра данных требуется только текстовый редактор. По служебным

¹⁰DataForge Streaming envelope format documentation. 2019. URL: <http://npm.mipt.ru/dataforge/docs.html#envelopes>.

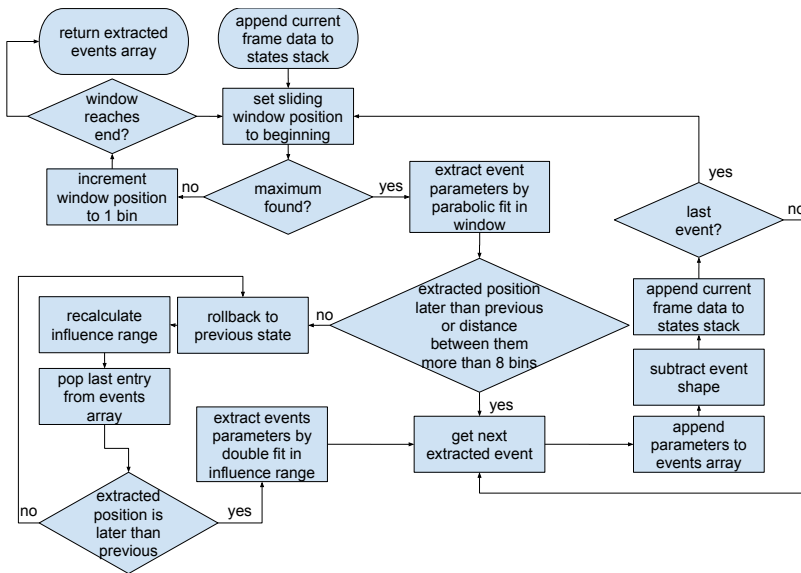


Рис. 9 — Блок-схема алгоритма выделения параметров событий.

параметрам пакета можно определить его размер, что позволяет создавать вложенные сообщения и использовать формат для работы в потоке стека TCP/IP.

- Возможно быстрое чтение и предобработка сохраненных данных только по метаданным.

Реализована распределенная система сбора данных. Алгоритм управления установкой разбит на три независимых модуля, отвечающих за: считывание сигнала с детектора, управление напряжением спектрометра и проведение набора по сценарию. Модули имеют единообразный программный интерфейс и взаимодействуют через стек TCP/IP. В основе аппаратной модернизации лежит использование ПК-контроллера крейтов КАМАК ССРС7 для управления подсистемами установки. Каждый модуль представляет собой программный сервер, который устанавливается в ССРС7 и осуществляет контроль аппаратуры через магистраль крейта, RS-232, RS-485 интерфейсы и другие специфические соединения. Для упрощения разработки реализована работа модулей в виртуальном режиме, в котором не требуется подключение к реальной аппаратной части. Виртуальный режим упрощает разработку и позволяет проводить неполный запуск системы (отключенные модули заменяются виртуальными). Проведено успешное тестирование, в результате которого подтверждено полное соответствие функционалу исходной системы.

2. Проведен переход на запись непрерывного сигнала. Для модернизированной системы сбора разработан модуль записи данных непрерывной оцифровки с помощью АЦП Лан10-12РСІ. В процессе записи плата последовательно сбрасывает кадры максимальной длины по программному триггеру. Каждый последующий кадр сбрасывается сразу после сохранения предыдущего. Таким образом, сохраняется непрерывная оцифровка сигнала с пропусками на время сбросов. Такой подход позволяет обойти проблему аппаратного мертвого времени, возникающую при стандартном наборе по триггеру, и проводить более сложную обработку сигнала в офлайн-режиме. Особенности модуля Лан10-12РСІ:
 - Модуль полностью совместим с программным интерфейсом системы и не требует модификации других модулей при интеграции/отключении.
 - Модуль прозрачно встраивается в систему (т. е. подключается между управляющим и детекторным модулями, транслируя их сообщения без изменений) и набирает данные в дублирующем режиме.
 - Для взаимодействия с АЦП Лан10-12РСІ используется написанное в рамках работы [5] CLI-приложение.
3. В рамках сотрудничества с TRISTAN на установке «Троицк ню-масс» проведено подключение прототипа детектора производства XGLab для эксперимента KATRIN к системе сбора данных установки. Разработан эмулятор сигналов, работающий по спецификации, предоставленной XGLab. При помощи эмулятора создан программный пакет, обеспечивающий взаимодействие с АЦП DANTE, управление считыванием, а также хранение событий вместе со специфическими для этой системы метаданными. При помощи этой системы проведены два сеанса измерений на установке с использованием детектора TRISTAN.
4. Разработаны алгоритмы выделения параметров событий (амплитуды и положения) из оцифрованного сигнала, снятого платой Лан10-12РСІ. Для данных «Троицк ню-масс» при выделении параметров событий удалось добиться рекордного эффективного мертвого времени порядка 0.9 мкс для средней длины «колокола» события 6 мкс и размера одного канала оцифровки 320 нс (мертвое время при аппаратной обработке событий составляло около 7 мкс). Также обработка формы импульса позволила исключить систематическую ошибку восстановления амплитуд, вызванную наложением на хвосты предыдущих событий. Проведено обобщение алгоритмов на произвольную форму импульса. Производительность алгоритма оптимизирована до уровня, когда он может быть использован в режиме реального времени. Исходный

код с инструкциями по воспроизведению результатов выложен в открытый репозиторий^{11,12}. Описание алгоритмов опубликовано в [6].

- Разработаны обобщенные генераторы шума и событий, симулирующие реальный сигнал АЦП. Генераторы извлекают параметры из кадров АЦП и не требуют ручной настройки. Разработаны тесты для оценки качества генерируемых данных.
- Разработан фреймворк для тестирования алгоритмов обработки, имеющий функционал: расчета процентного соотношения событий по типам (нераспознанные, ложно-распознанные и т. д.), построения распределения ошибок распознавания амплитуды и времени, графического отображения случайного примера кадра, вызывающего ошибку обработки.
- Разработан алгоритм выделения параметров событий, имеющий мертвое время 0.9 мкс. В основе его работы лежит последовательное вычитание из кадра распознанных событий от первого к последнему. Для выделения параметров используются два метода: параболическое фитирование пика для одиночных или хорошо разделенных событий и фитирование функцией суммы форм двух событий для наложенных событий. Условием для использования второго метода выступает нарушение очередности извлекаемого события.

Проделанная работа представляет ценность как в рамках эксперимента «Троицк ню-масс», так и для развития экспериментальной физики частиц в целом.

Распределенные системы контроля и сбора данных представляют существенный интерес для экспериментов следующего поколения. Успешный опыт использования этой концепции в эксперименте «Троицк ню-масс» будет применен при разработке новых систем. Распределенная система хранения данных и обмена сообщениями также представляет интерес как для экспериментальной физики, так и для промышленного использования.

Существенной особенностью разработанных алгоритмов по разделению наложений является то, что они не привязаны к конкретной форме сигнала и могут быть в перспективе адаптированы для использования в других экспериментах.

Все программные компоненты, которые реализованы в данной работе и имеют потенциал повторного использования, были подробно задокументированы и выложены в репозитории с открытым доступом.

¹¹signal-utils. 2019. URL: <https://github.com/kapot65/signal-utils>.

¹²lan10-processing. 2019. URL: <https://bitbucket.org/Kapot/lan10-processing>.

Публикации автора по теме диссертации

1. First measurements in search for keV sterile neutrino in tritium beta-decay in the Troitsk nu-mass experiment / J. N. Abdurashitov [et al.] // JETP Letters. — 2017. — June. — Vol. 105, no. 12. — P. 753–757.
2. The current status of "Troitsk nu-mass" experiment in search for sterile neutrino / D. N. Abdurashitov [et al.] // Journal of Instrumentation. — 2015. — Apr. — Vol. 10, no. 10. — T10005–T10005.
3. Measurements with a TRISTAN prototype detector system at the “Troitsk nu-mass” experiment in integral and differential mode / T. Brunst [et al.] // Journal of Instrumentation. — 2019. — Nov. — Vol. 14, no. 11. — P11013–P11013.
4. Silicon drift detector prototypes for the keV-scale sterile neutrino search with TRISTAN / K. Altenmüller [et al.] // Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. — 2018. — Dec. — Vol. 912. — P. 333–337.
5. *Abdurashitov, D. N.* The Long-Term Stability of a Fused-Silica Proportional Counter / D. N. Abdurashitov, V. G. Chernov // Instruments and Experimental Techniques. — 2019. — Jan. — Vol. 62. — P. 5–9.
6. *Chernov, V.* Shape-based event pileup separation in Troitsk nu-mass experiment / V. Chernov, A. Nozik // Journal of Instrumentation. — 2019. — Aug. — Vol. 14. — T08001–T08001.

Научное издание

Чернов Василий Геннадьевич

Разработка распределенной системы сбора данных и анализ формы импульса
событий на установке «Троицк ню-масс»

Автореферат

диссертации на соискание ученой степени
кандидата физико-математических наук на тему

Ф-т 60x84/16 Уч.-изд.л. 1,1 Зак. № 22458 Тираж 80 экз. Бесплатно

Печать цифровая

Федеральное государственное бюджетное учреждение науки
Институт ядерных исследований Российской академии наук

Издательский отдел

117312, Москва, проспект 60-летия Октября, 7а