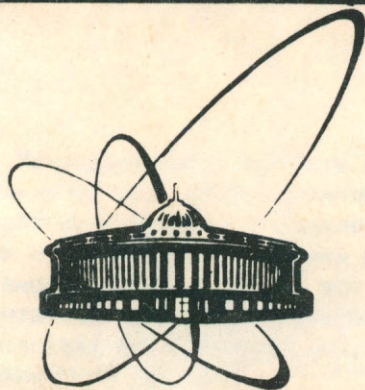


89-293



ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА

P10-89-293 †

Н.Н.Говорун, А.П.Воробьев*, В.Ю.Глебов *,
И.А.Голутвин, С.Г.Каданцев, В.М.Котов,
З.В.Крумштейн, Г.В.Мицельмахер, В.А.Сергеев *,
В.П.Шириков

ВАРИАНТ ТРИГГЕРА ВЫСОКОГО УРОВНЯ УКД УНК
НА ОСНОВЕ СТАНДАРТА МАЛТИБАС II.

Направлено в Оргкомитет VII Всесоюзного симпозиума
"Модульные информационно-вычислительные системы",
Новосибирск, 11-13 мая 1989 г.

* Институт физики высоких энергий, Серпухов

1989

Исследования в области физики высоких энергий на Ускорительно-накопительном комплексе /УНК/ ИФВЭ требуют создания аппаратно-программных комплексов, способных принимать с детекторов и обрабатывать потоки информации в $10^{10} \pm 10^{12}$ бит/с/1/. Невозможно зафиксировать всю информацию, получаемую в экспериментах во время их проведения, поэтому анализ и отбор данных, представляющих физический интерес, необходимо производить в реальном времени.

Система сбора и обработки данных /ССОД/ современных экспериментов должна выполнять функции триггирования /отбора/ событий, осуществляемые последовательно в несколько этапов /уровней/. Общая схема прохождения информации и распределение функций аппаратуры и программ ССОД, приведенные на рис. 1, предложены рабочей группой по проекту SSC/2/, а суммарная вычислительная мощность, необходимая для обработки данных в реальном времени, оценивается в несколько сотен ЭВМ типа VAX 11/780.

Следует отметить отличительную особенность организации обработки данных с детекторов коллайдеров на энергиях в несколько ТэВ - это применение принципа конвейерной обработки данных (pipeline), что позволяет минимизировать мертвое время системы. В проекте H¹/3/ для триггера 1-го уровня принята длина pipeline /Тр = 2112 нс/, равная 22 периодам пересечения банчей /Т = 96 нс/, а время решения триггера 1-го уровня составляет примерно $0,8 \pm 1,2$ мкс.

Триггер 2-го уровня анализирует события, прошедшие отбор на 1-м уровне. Частота событий на его входе на порядок меньше, и время его решения может быть значительно больше /Т₂ = 10 мкс/, поэтому на этом уровне могут применяться DSP (DIGITAL SIGNAL PROCESSORS), представляющие собой RISC-процессоры с предельно ограниченным составом команд.

Передавать всю информацию дальше, даже для отобранных событий, бессмысленно. В среднем на одно событие только 5% ячеек калориметра дают информацию при больших значениях поперечных импульсов. Поэтому применение DSP в составе триггера 2-го уровня позволяет также производить эффективное "сжатие" информации - преобразование ее в компактную форму перед процессом записи данных (READIN).

Сбрасыванием триггера 2-го уровня заканчивается процесс подавления фоновых и отбор полезных событий, осуществляемый,

Исследования в области физики высоких энергий на Ускорительно-накопительном комплексе /УНК/ ИФВЭ требуют создания аппаратно-программных комплексов, способных принимать с детекторов и обрабатывать потоки информации в $10^{10} \div 10^{12}$ бит/с/1/. Невозможно зафиксировать всю информацию, получаемую в экспериментах во время их проведения, поэтому анализ и отбор данных, представляющих физический интерес, необходимо производить в реальном времени.

Система сбора и обработки данных /ССОД/ современных экспериментов должна выполнять функции триггирования /отбора/ событий, осуществляемые последовательно в несколько этапов /уровней/. Общая схема прохождения информации и распределение функций аппаратуры и программ ССОД, приведенные на рис. 1, предложены рабочей группой по проекту SSC/2/, а суммарная вычислительная мощность, необходимая для обработки данных в реальном времени, оценивается в несколько сотен ЭВМ типа VAX 11/780.

Следует отметить отличительную особенность организации обработки данных с детекторов коллайдеров на энергиях в несколько ТэВ - это применение принципа конвейерной обработки данных (pipeline), что позволяет минимизировать мертвое время системы. В проекте H¹/3/ для триггера 1-го уровня принята длина pipeline /Тр = 2112 нс/, равная 22 периодам пересечения банчей /Т = 96 нс/, а время решения триггера 1-го уровня составляет примерно $0,8 \div 1,2$ мкс.

Триггер 2-го уровня анализирует события, прошедшие отбор на 1-м уровне. Частота событий на его входе на порядок меньше, и время его решения может быть значительно больше /Т₂ = 10 мкс/, поэтому на этом уровне могут применяться DSP (DIGITAL SIGNAL PROCESSORS), представляющие собой RISC-процессоры с предельно ограниченным составом команд.

Передавать всю информацию дальше, даже для отобранных событий, бессмысленно. В среднем на одно событие только 5% ячеек калориметра дают информацию при больших значениях поперечных импульсов. Поэтому применение DSP в составе триггера 2-го уровня позволяет также производить эффективное "сжатие" информации - преобразование ее в компактную форму перед процессом записи данных (READIN).

Сбрасыванием триггера 2-го уровня заканчивается процесс подавления фоновых и отбор полезных событий, осуществляемый,

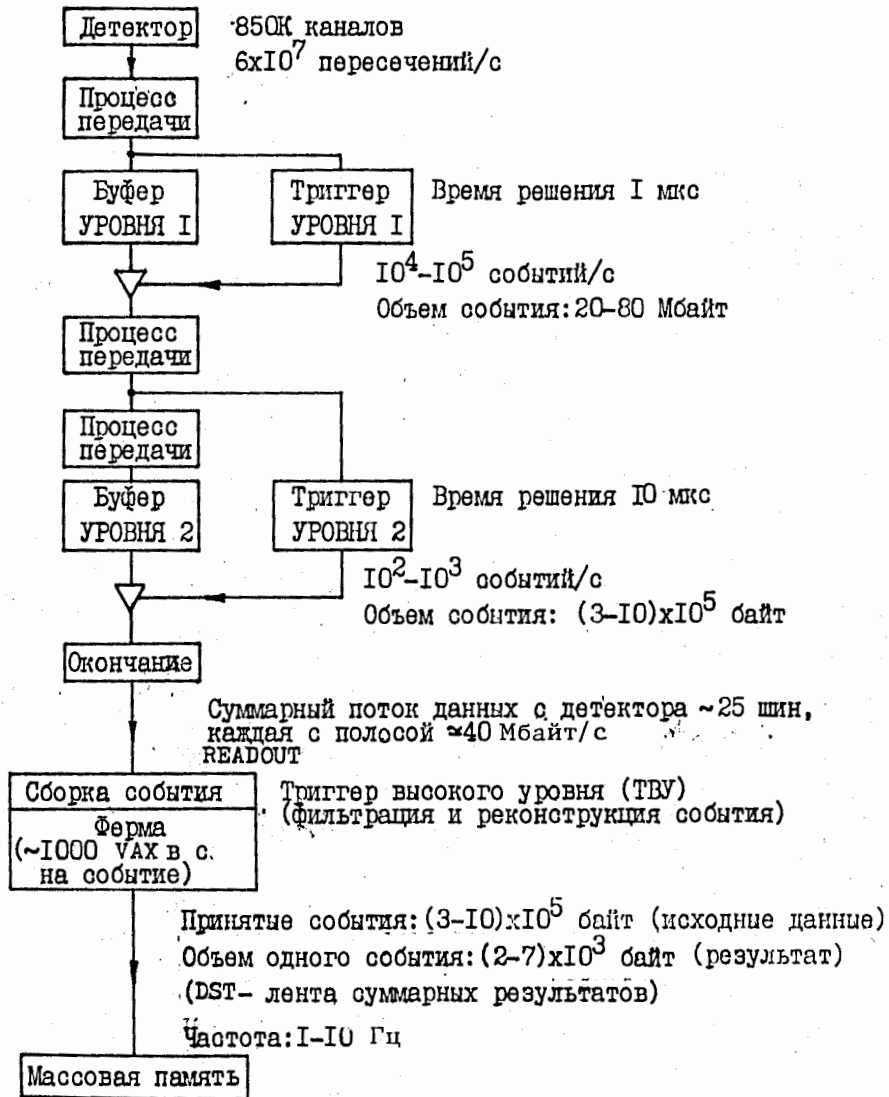


Рис. 1.
Общая модель потока данных.

в основном, аппаратными методами. Алгоритмы отбора здесь тесным образом связаны с особенностями детекторов, регистрирующей электроникой и физическими задачами, решаемыми на данном детекторе. Именно триггер 2-го уровня, выделяя полезное собы-

тие, начинает его оцифровку и производит запись данных (READIN) в память FRONT-END электроники отдельных детекторов.

Дальнейший анализ можно проводить только при наличии полных данных о событии, собранных из буферов отдельных детекторов. Алгоритмы обработки на этом уровне весьма сложные и требуют применения быстрых спецпроцессоров и мощных универсальных процессоров. Это - функции триггера высокого уровня, проблеме создания которого посвящена эта работа.

СХЕМА ТРИГГИРОВАНИЯ ВЫСОКОГО УРОВНЯ СИСТЕМЫ СБОРА И ОБРАБОТКИ ДАННЫХ

Рассматривается вариант функциональной схемы триггирования последнего /высокого/ уровня, именуемого далее - триггер высокого уровня /ТВУ/ системы сбора и обработки данных для универсального калориметрического детектора /УКД/ УНК /рис.2/.

Основная задача ТВУ заключается в считывании данных из буферов памяти FRONT-END процессоров отдельных детекторов установки, сбора полного события, восстановления его в пространстве детекторов, фильтрации для окончательного отбора полезных событий и записи их на носители массовой памяти для дальнейшего анализа. Этот процесс сопровождается мониторингом, калибровочными и тестовыми процедурами, обеспечивающими "дружественный" режим работы для экспериментаторов по контролю и управлению системой сбора и обработки данных на установке. Допустимая частота событий на выходе системы определяется объемом события и скоростью записи на носитель и составляет 1 ÷ 10 Гц. Связи ТВУ и его место в общей схеме ССОД показаны на рис. 3.

Частота поступления событий на вход ТВУ определяется величиной режекции триггеров предыдущих уровней и составляет не менее 100 Гц. Учитывая объем события, который может колебаться от 0,3 до 1 Мбайт, полосу пропускания на входе следует выбирать равной примерно 100 Мбайт/с. Эти оценки относятся к средней скорости поступления событий на вход ТВУ, т.е. предполагается, что глубина буферизации в FRONT-END электронике достаточна для выравнивания импульсных перегрузок по потокам еще до уровня READOUT /например, величина буферной памяти в контроллерах READOUT установки Н1/3/ составляет 10 ÷ 20 событий среднего объема/.

Так как длина кабелей для соединения электроники FRONT-END детекторов и ТВУ должна быть несколько десятков метров, то система ввода данных ТВУ должна иметь параллельный вход, рассчитанный на ввод 5 ÷ 10 отдельных шин с полосой пропускания

в 20 Мбайт/с каждая, с последующим мультиплексированием их для сбора полного события в отдельных локальных областях памяти буфера полных событий.

Процедура фильтрации событий в ТВУ определяет целесообразность выделения 2-х внутренних уровней в ТВУ и буферизации при обмене между ними для выравнивания скорости передачи данных внутри ТВУ. Предварительная фильтрация /триггер 3-го уровня/ осуществляется на спецпроцессорах, ориентированных на выполнение быстрых алгоритмов предварительной фильтрации с использованием триггерных данных. На этом же уровне осуществляется декодировка исходных данных в виде, удобном для полной фильтрации, так как ранее в процессе считывания на этапе READIN производится "сжатие" информации /выбрасываются "нулевые" значения, вычитаются пьедесталы и т.п./, и данные форматируются компактно. Введение предварительной фильтрации уменьшает требования на полосу пропускания и на количество процессоров, необходимых для полной фильтрации на 4-м уровне ТВУ. Разбиение ТВУ на два отдельных уровня присуще практически всем ССОД современных детекторов. Анализируя структуру как созданных, так и проектируемых систем сбора и обработки данных для экспериментов на коллайдерах, можно выделить наиболее удачные, на наш взгляд, архитектуры ССОД в H1 HERA/3/ ФРГ, CDF/4/ США, DO/5/, DELPHI ЦЕРН/6/, опыт создания которых учитывается при разработке данной схемы.

Требования широкой полосы пропускания на входе в ТВУ определили необходимость применения параллельного способа записи в буфер полных событий. Основным элементом системы записи предлагается, так же, как и в ССОД DO FERMILAB/5/, набор двухпортовых памятей, один порт каждой из которых выведен на переднюю панель модуля памяти и соединяется при помощи кабеля данных с процессором чтения данных из буфера событий одного из детекторов, осуществляющим процесс READOUT /рис. 2/. Чтобы триггер 3-го уровня не уменьшал полосу пропускания ТВУ по входу, предлагается ввести несколько секций буферов полных событий глубиной не менее 2-х событий, каждая из которых имеет собственный триггер /спецпроцессор/ 3-го уровня. В этом случае составитель события /координатор/, входящий в состав мастер-крейта ССОД /рис. 2/, выполняет функции супервизора ввода и выбирает свободный, готовый к работе триггер 3-го уровня и организует закачку полного события из кабелей данных в 2-портовую память именно этого триггера. Каждая секция из 2-х триггеров 3-го уровня имеет контроллер быстрой шины, который, выполняя функции host-процессора шины, производит передачу отобранных триггером 3-го уровня событий в один из процессоров фермы одноплатных универсальных ЭВМ, выполняю-

щих фильтрацию полного события, т.е. в триггер 4-го уровня ТВУ /рис. 2/. Количество секций триггеров 3-го уровня определяется по формуле /1/.

Структура получаемых в процессе READOUT данных, в виде отдельных событий, обработка /фильтрация/ которых может производиться независимо, создает возможности для применения параллельной обработки событий в отдельных процессорах, объединенных шиной магистрально-модульных систем и быстрой шиной межкрейтных обменов, в единую систему - "ферму". Этот принцип организации обработки данных в экспериментальной физике впервые реализовали Томас Нэш и сотрудники FERMILAB США/4/ для off-line обработки, но он может с успехом применяться и для фильтрации в реальном времени/4/.

Для увеличения полосы пропускания системы и объединения в единую систему практически неограниченного количества процессоров обработки/4/ необходима система межкрейтных связей, включающая в себя быстрые шины межкрейтных соединений, контроллеры управления быстрой шиной и переключатель быстрых шин.

В рассматриваемой схеме ТВУ быстрая шина /БШ/ /полоса пропускания - 20 мбайт/с, адрес и данные 32-разрядные/ объединяет секции магистрально-модульной архитектуры стандарта MULTIBUS2, каждая из которых содержит 8 одноплатных ЭВМ, программно совместимых с ЭВМ микроVAX. Контроллер быстрой шины /КБШ/, входящий в состав модуля триггера 3-го уровня, выполняет роль host-ЭВМ одной из быстрых шин. Переключатель быстрых шин /ПБШ/ обеспечивает соединение КБШ и одной из БШ, у которой хотя бы один из процессорных модулей фермы свободен и готов начать фильтрацию текущего события, отобранного триггером 3-го уровня для дальнейшего анализа. Количество процессорных модулей в ферме для одной быстрой шины и одного КБШ можно оценить по формуле

$$N_{ЭФ} = \left[\frac{T\Phi * C}{K_{ВВ}} \right] + 1, \quad /1/$$

где ТФ - время фильтрации в одном модуле, с; С - частота поступления данных на вход фермы, Кбайт/с; $K_{ВВ}$ - суммарный объем информации ввода-вывода в один процессорный модуль в Кбайт.

К переключателю быстрых шин подключены также БШ к основному /мастер-/крейту контроллера ТВУ ССОД и БШ от выходного буфера событий, в который собираются все обработанные события триггера 4-го уровня со всех ферм. Этот буфер полных событий является частью отдельной фермы, обеспечивающей вывод обрабо-

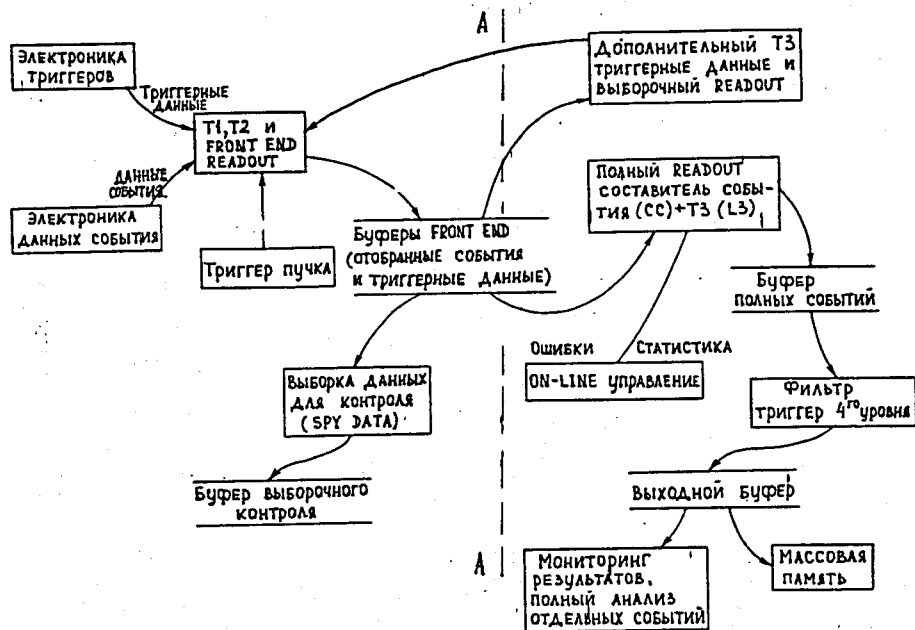


Рис. 4. Диаграмма прохождения данных через ССОД.

танных данных из ТВУ, выборочную обработку отдельных классов событий, мониторинг, тестовые и калибровочные работы по выходным результатам. Эта входная ферма имеет непосредственный выход на собственную host-ЭВМ.

Диаграмма прохождения данных от аппаратуры детекторов через всю систему триггирования до записи результатов на носители массовой памяти представлена на рис. 4.

Суммируя сказанное, можно отметить следующее:

Первая стадия считывания данных с детекторов осуществляется контроллерами центрального триггера и триггеров 1-го и 2-го уровней в каждом детекторе. На этом этапе происходит форматирование и кодировка данных для уменьшения их объема. Контроллером каждого детектора, в зависимости от его типа и состава аппаратуры, проводятся некоторые виды обработки. Операции на этой стадии выполняются, в основном, аппаратно, и все алгоритмы обработки здесь жестко синхронизированы для обеспечения режима pipeline и заканчиваются записью (READIN) отобранных событий в буферы FRONT-END контроллеров каждого детектора.

Затем /ниже линии AA на рис. 2/ начинается работа ТВУ /триггеры 3-го и 4-го уровней/. Составитель события в ТВУ,

взаимодействуя с контроллером центрального триггера и контроллерами детекторов, начинает полный READOUT, который включает все операции по передаче данных в буфер полных событий: сборку события, форматирование и добавление служебной информации, необходимой для дальнейшей обработки. Полный READOUT осуществляется в асинхронном режиме по отношению к работе триггеров 1-го и 2-го уровней. При этом обязательным условием является обеспечение необходимой полосы пропускания на входе в ТВУ, чтобы средняя скорость считывания (READOUT) была не меньше скорости закачки (READIN) в буфер FRONT-END.

Особенность организации ввода данных и структура триггеров третьего уровня (L3) позволяют легко ввести в ТВУ "режим просмотра" (spying processing) для контроля за работой системы триггирования в целом.

ЗАКЛЮЧЕНИЕ

Выбор типов процессорных модулей и способа их объединения в единую вычислительную систему /ферму/ определяет базис для разработки архитектуры и функциональной схемы ТВУ и выбора структуры системного программного обеспечения.

Выделение этой задачи как основной на начальной стадии подготовки проекта триггера высокого уровня /ТВУ/ УКД обосновывается в "Предложениях по проекту триггера высокого уровня для экспериментов на УНК", представленных рабочей группой ЛВТА ОИЯИ, которые были рассмотрены и одобрены 12-м Рабочим совещанием по экспериментам на встречных пучках УНК.

Выбор стандарта МАЛТИБАСС II, принятого в июне 1987 года в качестве международного стандарта для многопроцессорных 32-разрядных вычислительных систем, объясняется тем, что он Постановлением ГКВТИ принят как перспективный стандарт для создания высокопроизводительных магистрально-модульных систем и в СССР.

ЛИТЕРАТУРА

1. Denisov S.P. - In: Proceedings of Experiment Facility for UNK Collider, IHEP, Serpukhov, 1988, p.343.
2. Proceeding of the Workshop on Physics at Future Accelerators. CERN 87-07, v.1, June 1987.
3. Technical Progress Report H1 Collaboration, October 4, 1987.
4. The ACP Branch Bus and Real Time Applications of ACP Multiprocessor System. R.Hance et al. Fermilab - Conf. - 87/76, May 8, 1987.

5. The MICROVAX - Based Data Acquisition System for D0, D.Cutts et al., 6/87.
6. The DELPHI Data Acquisition System Project, DELPHI-MI/DAS 87-04, September 1987.
7. MULTIBASII Product Data Book, INTEL Corporation, 1986.

Рукопись поступила в издательский отдел
27 мая 1989 года.