

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ В ЯДЕРНО-ФИЗИЧЕСКИХ ИССЛЕДОВАНИЯХ*

В.И. Савельев

DESY, Hamburg, Germany



В работе рассмотрены вопросы создания систем математического моделирования и анализа для крупномасштабных ядерно-физических экспериментов. Подробно рассмотрены вопросы системного решения вычислительных средств систем, алгоритмы и структуры программного обеспечения. Рассмотрен конкретный пример эксперимента HERA-B, позволяющий продемонстрировать основные тенденции создания систем математического моделирования и анализа на современном этапе.

ВВЕДЕНИЕ

Процесс создания экспериментальной установки в физике высоких энергий как нельзя лучше характеризует основные особенности новой технологии научных исследований математического моделирования и вычислительного эксперимента, отличающихся многоцелевой направленностью и методологической универсальностью, позволяя на основе накопленного опыта математического моделирования быстро и эффективно решать задачи. Важнейшим результатом является возможность проведения вычислительного эксперимента уже на этапе создания экспериментальной установки с обоснованием оптимальных условий проведения эксперимента - структуры спектрометра, состава детекторов, необходимой статистики экспериментальных данных, а также сроков проведения и ожидаемых результатов.

При проведении экспериментальных исследований роль математических методов и математического моделирования дает возможность подойти по-новому к вопросу организации процесса проведения исследований с более тесной координацией экспериментальных и теоретических исследований в рамках новой методологии научных исследований, совмещения реального эксперимента (набора экспериментальных данных и их анализа) с математическим моделированием и вычислительным экспериментом на базе вычислительных средств эксперимента.

СТРУКТУРА СОВРЕМЕННЫХ ЯДЕРНО-ФИЗИЧЕСКИХ ЭКСПЕРИМЕНТОВ

Структура крупномасштабных экспериментов в физике высоких энергий полностью соответствует предложенной А.А. Самарским схеме, «модель-алгоритм-программа» (рис. 1), где прослеживаются основные этапы вычислительного и реального эк-

© В.И. Савельев, 2002

* Работа поддержана грантом РФФИ 02-01-01014

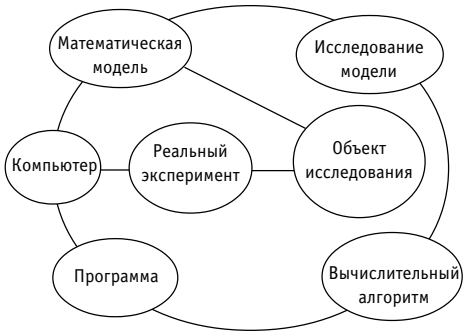


Рис. 1. Схема «модель- алгоритм-программа»

спериментов [1].
Примером может служить эксперимент HERA-B-исследование CP нарушения в системе В-мезонов на ускорительно-накопительном комплексе HERA при энергии протонов 920 ГэВ [2]. На рис. 2 приведена схема спектрометра эксперимента HERA-B, представляющего собой полнофункциональный спектрометр физики высоких энергий с развитой структурой вычислительных средств: системой математического моделирования, накопления статистики, предварительного и окончательного анализа данных.

Необходимо отметить, что современные эксперименты физики высоких энергий практически не имеют возможности использовать готовые технологические решения - их просто не существует. Принятие решения, как правило, обосновывается результатами полного монте-карло-моделирования и исследования спектрометра с алгоритмами и программами сбора и анализа данных.

Современное экспериментальное состояние эксперимента HERA-B характеризуется многими факторами, основными из которых являются следующие.

- Значительный объем данных (как экспериментальных, так и моделируемых). Для успешного проведения эксперимента в условиях HERA-B поток информации оценивается объемом 20 TB в год, при условии набора статистики 10^7 событий/год и объемом 100 kВ/событие. При этом, как правило, поток информации от моделируемых процессов значительно превышает потоки экспериментальной информации.
- Необходимость максимальной производительности при реконструкции и анализе событий. Высокая множественность и топологическая сложность событий является сильным ограничивающим фактором отбора и создания алгоритмов для анализа. Современный уровень вычислительных процессоров позволяет достичь производительности 2-4 с/событие. Важнейшим фактором для анализа становится визуализация

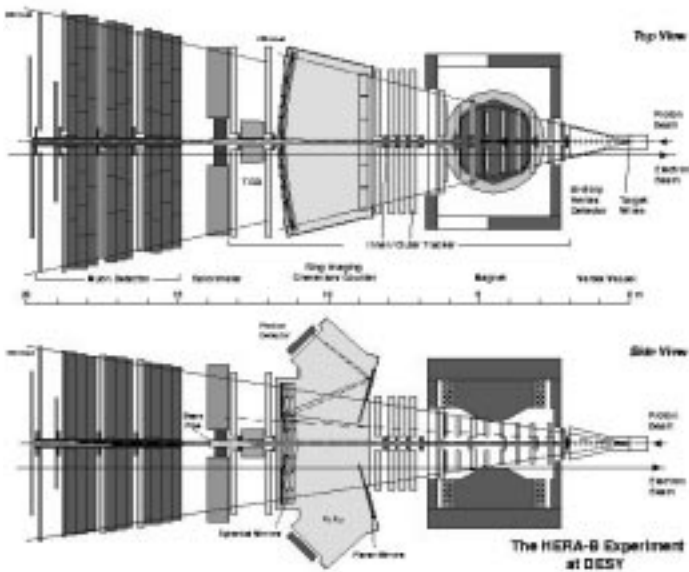


Рис. 2. Схематический вид спектрометра эксперимента HERA-B

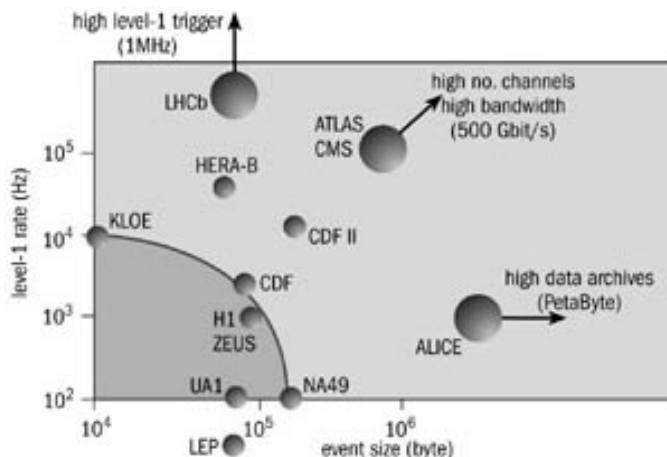


Рис. 3. Диаграмма производительности компьютерных систем анализа обработки информации для современных экспериментов в физике высоких энергий

ция в трехмерном пространстве результатов анализа.

На рис. 3 представлены достижения в области создания высокопроизводительных систем сбора и анализа данных для крупномасштабных экспериментов физики высоких энергий включая систему эксперимента HERA-B [3].

МЕТОД МОНТЕ-КАРЛО В ЭКСПЕРИМЕНТАЛЬНОЙ ФИЗИКЕ ВЫСОКИХ ЭНЕРГИЙ

Моделирование реальных физических процессов представляет собой наиважнейшую задачу, особенно для верификации новых физических явлений и создания экспериментального оборудования для физических исследований. Естественно, проверка математической модели и результатов моделирования должна быть выполнена на экспериментальной основе; в частности, важнейшим фактором является соответствие математического моделирования существующим физическим моделям и реальной постановке эксперимента. Основным методом моделирования, позволяющим получать значимые результаты, является метод Монте-Карло, который первоначально использовался успешно для задач нейтронной физики, где традиционные численные методы оказались малоприменимыми. Далее его применение распространилось на широкий круг задач статистической физики, в том числе на взаимодействие элементарных частиц и распространение электромагнитного излучения. Существенный вклад в развитие метода Монте-Карло для широкой области применения внесла школа российских математиков Г.И. Марчука, С.М. Ермакова [4].

Применение метода Монте-Карло оправдано, в первую очередь, в задачах, которые допускают теоретико-вероятностное описание, что объясняется естественностью получения результата с заданной вероятностью в задачах с вероятностным содержанием и существенным упрощением процедуры вычислений. Использование современных вычислительных средств и языков математического программирования существенно упростило программирование вычислений и перенесло центр тяжести на создание математического описания исследуемого явления, необходимые упрощения задачи, выбор подходящего численного метода, исследование его погрешности.

Общим инструментом анализа Монте-Карло в экспериментальной физике высоких энергий является комплекс программных библиотек CERNLIB и входящий в них программный пакет математического моделирования GEANT (описание детекторов и моделирование физических процессов [5]) включающий в себя

- описание физической установки (спектрометра), основанное на структуре геометрических образов, представляющих материальные объекты и учитывающих вещество как взаимодействующую среду;
- прохождение элементарных частиц и ионизирующих излучений в материале установки, моделируемое с учетом границ геометрических объемов и физических явлений в соответствии с природой частиц, взаимодействия с веществом, магнитными и электрическими полями;
- статистические исследования, выполняемые методом Монте-Карло;
- фиксирование параметров физических процессов, осуществляемое на основе баз данных.

Основными особенностями метода Монте-Карло являются возможность моделирования прохождения элементарных частиц через экспериментальную установку (физическую) и полное моделирование взаимодействия элементарных частиц с веществом с возможностью графического представления физических процессов. Необходимо отметить особенность ситуации, связанной с описанием явления переходного излучения, открытого российскими учеными [6]. Полное отсутствие описания и моделирования переходного излучения в стандартных программных пакетах физики высоких энергий значительно усложняет исследования по переходному излучению и создание детекторов на его основе, являющихся мощным средством идентификации элементарных частиц. Важнейшим вкладом автора является создание прикладной математической модели переходного излучения, совместимой с основными системами моделирования процессов в физике высоких энергий и адаптированной с основными программными библиотеками физики высоких энергий [7].

Программные пакеты моделирования взаимодействия элементарных частиц с веществом первоначально разработаны для физики высоких энергий, и в настоящее время находят широкое применение в таких областях как медицина, биология, астрофизика, защита от ионизирующих излучений.

ВЫЧИСЛИТЕЛЬНЫЕ СРЕДСТВА ДЛЯ ПРОВЕДЕНИЯ МОДЕЛИРОВАНИЯ И АНАЛИЗА ДАННЫХ В ЭКСПЕРИМЕНТАХ ФИЗИКИ ВЫСОКИХ ЭНЕРГИЙ

Структура вычислительных средств экспериментов

Основным направлением создания современных вычислительных систем для крупномасштабных систем моделирования и анализа информации, включающих в себя прямые численные методы, является направление использования многопроцессорных вычислительных систем, представляющих собой распределенную сеть вычислительных кластеров с соответствующей архитектурой аппаратных и программных средств, развиваемое Б.Н.Четверушкиным [8]. Основными задачами создания высокопроизводительных вычислительных комплексов и алгоритмов для таких систем с целью успешного адаптирования к архитектуре многопроцессорных вычислительных систем являются следующие.

1. Наличие внутреннего параллелизма для возможности одновременного решения ряда подзадач, число которых оптимально равно числу процессоров вычислительной системы. Свойство внутреннего параллелизма предполагает, что разбиение на такие задачи может быть произведено оптимальным образом как опытным программистом, так и автоматически, с помощью соответствующего программного обеспечения. Это требование является ключевым для возможности использования многопроцессорной вычислительной системы.

2. Вторым требованием к параллельным процессам является обеспечение равномерной загрузки (load balancing) процессоров. В случае, когда однородные процес-

соры загружены неравномерно, наименее эффективный процесс (например, из-за большого объема вычислений) будет определять эффективность всей вычислительной системы в целом. Требование равномерной загрузки процессоров неразрывно связано с требованием внутреннего параллелизма метода.

3. Третьим важнейшим свойством является минимизация обмена информацией между процессорами в процессе вычислений. Это требование, в первую очередь, актуально для многопроцессорных вычислительных систем с распределенной памятью. При большом объеме обмениваемой информации многопроцессорная система с распределенной памятью значительную часть своего времени будет тратить на обмена по относительно медленным каналам связи. Это, естественно, приводит к снижению эффективности работы всей системы.

4. Четвертым свойством алгоритмов, позволяющим добиться успешной параллельной адаптации, является простота алгоритмов. Это требование особенно важно в связи с бурным прогрессом в области многопроцессорной вычислительной техники, сопровождаемым неизбежным изменением математического обеспечения.

Данные принципы определяют основу планирования и создания вычислительной среды для крупномасштабных вычислительных и натурных экспериментов в науке и технике.

Визуализация больших объемов данных на многопроцессорных системах

Наряду с общими принципами хочется отметить некоторые специфические области, влияющие на структуру вычислительных средств для экспериментов в физике высоких энергий. Важнейшим средством повышения эффективности проведения экспериментальных исследований в крупномасштабных экспериментах физики высоких энергий в настоящее время становятся процессы визуализации больших объемов данных [9]. Обработка больших объемов информации на высокопроизводительных многопроцессорных системах создает ряд новых проблем, связанных с визуализацией результатов научных расчетов и моделирования. При анализе новой ситуации представляется важным то условие, что данные проблемы приходится решать научным коллективам, занятым непосредственно научными исследованиями.

В данный момент практически отсутствуют доступные системы визуализации и хранения больших объемов данных, в том числе и интерактивные средства, ориентированные на работу в неоднородных вычислительных локальных и глобальных сетях. В результате, при работе в таких системах проблемы начинаются уже с момента записи данных на диски. На многопроцессорных вычислительных системах общего назначения графические подсистемы и устройства отображения обычно недостаточно развиты под специфические задачи. В результате передача больших потоков информации по локальным и глобальным сетям для визуализации катастрофически увеличивает время отклика, визуализация практически перестает быть интерактивной и нарушается нормальное функционирование вычислительной сети в целом.

В полной мере сложности проявляются при попытке визуализировать данные с помощью привычных интерактивных систем, как правило, ориентированных на технику персонального применения и обладающих в силу этого ограниченными ресурсами. Большинство программ визуализации, ориентированных на персональные компьютеры, составляющие основной парк рабочих мест, не в состоянии даже прочесть требуемые объемы информации. В этом смысле привлекательной является идея построения алгоритмов, позволяющих работать с данными, хранимыми в сжатом виде в оперативной памяти. Время, требуемое на распаковку при многократном обращении к одним и тем же данным, может с лихвой окупиться за счет снижения затрат на

присущий современным многозадачным системам страничный обмен (swapping) и относительного увеличения числа обращений к данным, находящимся в процессорной памяти.

Определенные сложности возникают на пути обеспечения равномерной загрузки процессоров в процессе построения изображений. Например, при построении изопрооекций число операций сильно зависит от градиента функции - величины, как правило, распределенной по изображению крайне неравномерно. Средства визуализации должны поддерживать возможность увеличения фрагмента изображения, интересующего пользователя. Простое выделение для увеличения одной четверти изображения при равномерном распределении области визуализации по процессорам приводит к тому, что три четверти процессоров в дальнейшей обработке изображения принимать участия не будут.

Подобное состояние в области визуализации больших массивов данных, в первую очередь, связано с новизной проблемы. Одним из перспективных направлений преодоления является технология клиент – сервер [10]. В рамках этой технологии сервер, запущенный на многопроцессорной системе, производит основную обработку данных, а клиент поддерживает диалог с пользователем и производит окончательные формирование и вывод на экран. Важным преимуществом технологии клиент-сервер является преемственность программного обеспечения. При переходе на новую вычислительную систему не нужно переносить все программное обеспечение. Вся графическая часть остается у клиента и будет работать без изменений с новым сервером в новой системе. Технология клиент-сервер универсальна по отношению к операционным системам. Под каждую операционную систему можно создать свою клиент-часть, которая будет использовать эффективно возможности данной системы. В то же время сложное прикладное математическое обеспечение останется на сервере и не подвергнется изменениям, которые могут внести трудноидентифицируемые ошибки. При этом пользователь может анализировать результаты вычислений в привычной и удобной для него операционной системе.

Еще одним преимуществом технологии клиент-сервер является возможность удаленного доступа. Файлы задач зачастую имеют очень большой размер; передавать их по сети в исходном виде неудобно и для этого требуются значительные ресурсы. В то же время важно получить результаты в месте, удаленном от рабочего места пользователя. Кроме того часто возникает ситуация, когда пользователи используют распределенную систему вычислений. В этой ситуации использование технологии клиент-сервер просто незаменимо. При таком подходе все данные находятся на удаленном компьютере, а пользователь получит визуальное представление, используя клиент-часть системы визуализации и ресурсы глобальной сети, по которым передаются лишь данные, необходимые для построения изображения на пользовательском терминале.

Конкретные результаты представленных методов и средств визуализации будут представлены далее при рассмотрении эксперимента HERA-B.

СИСТЕМА ВЫЧИСЛИТЕЛЬНЫХ СРЕДСТВ ЭКСПЕРИМЕНТА HERA-B

Для достижения эффективного процесса моделирования и анализа эксперимента HERA-B принята концепция вычислительной структуры на основе иерархии многопроцессорных кластеров нескольких уровней для обработки потока данных в реальном масштабе времени, окончательной обработки и получения физических результатов на базе комплекса суперкомпьютеров [2].

Основные требования к вычислительной системе эксперимента определяются потоками данных, а различные уровни иерархии системы – задачами, решаемыми на

них, уровнем сложности получения данных из сырых экспериментальных и модельных данных. Объем данных и требование временных ограничений на время процессинга варьируется очень широко.

Важным моментом является то, что поток данных невозможно существенно снизить на одном уровне, и это определяет структуру с несколькими уровнями. Актуальный уровень для каждого уровня системы определяется компромиссом между получаемой эффективностью и техническими возможностями современной вычислительной техники. Поток данных каждого уровня должен быть согласован с входным потоком следующего уровня, чтобы обеспечить соответствующую скорость потока данных на выходе для записи в устройства массового хранения и визуализации. Общая структура иерархии потоков данных и соответствующих уровней многопроцессорной вычислительной системы эксперимента представлена на рис. 4.

Наиболее эффективным методом максимального снижения потока данных является иерархичная структура - достаточно простые по действию, но тщательно отработанные алгоритмы на низших уровнях сокращают поток данных на возможно ранних этапах, оставляя более интеллектуальную обработку и анализ на высших уровнях системы. Тип процессоров для первого уровня и передачи данных на следующий уровень должен быть достаточно быстрым и, как правило, строится на специализированных цифровых процессорах - транспьютерах.

На более высоких уровнях возможности по времени и объему анализируемых данных увеличиваются благодаря более точному анализу и обработке, а также использованию мощнейшего метода пространственной реконструкции и визуализации.

Общая структура специализированных кластеров эксперимента и их соответствие иерархии представлена на рис. 5.

Основные структурные элементы.

- Первый уровень системы состоит из сетевого кластера, состоящего из 1200 специализированных параллельных процессоров с последовательными конвейерами данных, в качестве которых используются транспьютеры - специализированные высокопроизводительные процессоры (DSP-Digital Signal Processor) на базе ADSP21060 (SHARC) [11]. Максимальная скорость пропускания данных при полном описании события составляет 10 MHz. Максимальное время для принятия решения 12 μ s, и фактор снижения потока данных до 200. Основные критерии отбора событий - кинематические параметры события, жесткие пороги, идентификация лептонов для реконструкции инвариантной массы и реконструкция инвариантной массы дилептонов.

- Второй уровень системы базируется на кластере из 240 PC-процессоров (Pentium III, 600-800 MHz). Входной поток данных до 50 kHz включая возможность изменения потока до 100 kHz, т.к. на этом уровне наиболее велика неопределенность. Доступ к данным осуществляется через быстрый сетевой канал. Среднее время обработ-

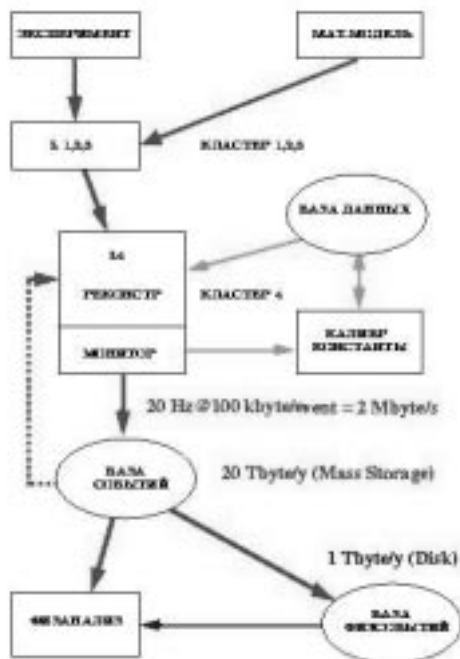


Рис. 4. Общая структура потоков данных в системе анализа данных эксперимента HERA-B

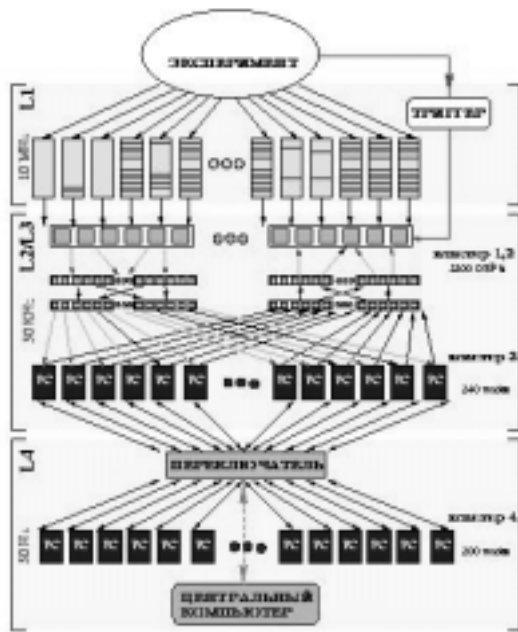


Рис. 5. Общая структура специализированных вычислительных кластеров эксперимента HERA-B

конструкцию события, классификацию и селекцию событий, используя полную информацию о событии. Позволяет получить дополнительный фактор 2.5 снижения потока данных перед записью на устройство долговременного хранения. Существенной особенностью четвертого уровня является возможность полной геометрической реконструкции события в 3D-пространстве и визуализация.

Окончательная обработка событий и получение физических результатов осуществляются на комплексе суперкомпьютеров SGI - сервер SGI Challenger (16 суперпроцессоров R10000 CPU, 250 MHz, 1536 MB RAM, производительностью 10 MGF, операционная система IRIX 6.5, 2 TB объем диска, 2 GB Ethernet-интерфейс).

Специализированный вычислительный кластер первого уровня

Информация о событии перед обработкой составляет объем 450 kB, и требуется передача для процесса обработки на уровне 22 GB/s при скорости поступления событий 50 kHz на первый уровень вычислительной системы. Первый уровень системы сохраняет информацию о событиях в течение обработки и запрашивает часть информации для анализа. Данный кластер состоит из 134 SHARC кластер процессоров, в каждом из которых 6 DSP. Структура отдельного модуля SHARC состоит из быстрой цифровой шины передачи данных 240 MB/s с возможностью коммуникации 6 процессоров с быстрым доступом к 512 kB памяти на каждый процессор. Каждый процессор имеет шесть каналов связи (SHARC-links) с частотой передачи 80 MHz, обеспечивая полосу передачи данных до 40 MB/s. Все 36 каналов связи SHARC модуля доступны по шине передачи и соединяются друг с другом посредством специальной сети, фиксированной аппаратно. Буферирование событий осуществляется жестко по принципу «первый пришел - первый ушел». Для ожидаемого среднего времени 3.8 ms 190 процессоров необходимо для процессинга событий и обеспечения 50 kHz выходного потока данных, соответствующего возможностям следующего уровня системы.

Специализированный сетевой переключатель для первого уровня системы также

ки события до 4 ms, возможный фактор снижения потока данных до 100 и выходной поток данных на уровне 500 Hz.

- Третий уровень системы (практически впервые) имеет доступ к полной информации о событии в локально распределенной памяти. Алгоритм более сложный, с возможностью адаптации и анализа нефиксированной информации, например, множественность, тип события. Третий уровень системы реализован на кластере 200 PC Pentium III, 800 MHz. Время обработки события до 100 ms и соответственно фактор снижения потока данных на этом уровне 10.

- Четвертый уровень позволяет обрабатывать поток данных до 50 Hz. Реализован в виде третьего кластера на 200 PC-процессорах (Pentium III, 800). Алгоритм выполняет он-лайн ре-

был реализован на основе SHARC-процессоров. Поток данных передается через процессоры каждого модуля к финальному и далее через канал связи. Решение на каждом этапе о передаче события принимается через таблицу назначений. Среднее время от запроса второго уровня системы до поступления события составляет 70 μ s.

Система вычислительных кластеров анализа второго, третьего уровней

Важнейшим преимуществом организации кластера, анализа и визуализации в режиме реального времени является сокращение времени отклика от анализа данных к процессу набора экспериментальных данных. Непосредственно после набора экспериментальных данных или генерации моделированных событий информация о качестве событий доступна, что дает быстрый отклик на получение данных. Другим преимуществом является возможность использования процессорного кластера в качестве интеллектуального фильтра событий. При этом возможны доступ к полной информации о событии, классификация и отбор событий, обеспечивающих доступ физиков к определенной базе событий. Это особенно важно для современных физических экспериментов, где поток значащих событий составляет величину порядка 10^{-5} от общего потока событий.

Реализация вычислительного кластера предусматривала решение соответствующих проблем, прежде всего создание комплекса программ анализа и реконструкции, работающих в режиме реального времени, в том числе имеющих возможность непосредственно получать данные от экспериментальной установки или генератора моделированных событий. Это требует распределения потока данных по процессорам для реконструкции, анализа и сбора в единый поток на выходе для архивирования.

Ключевой проблемой является задание калибровочных и стабилизирующих констант в процессе обработки событий, что требует соответствующего программного обеспечения и адекватной базы данных.

Для упрощения структуры и программного обеспечения для вычислительного кластера выбрана параллельная стратегия, т.е. процессоры не имеют взаимодействия в процессе обработки информации; каждый процессор загружается копией программы реконструкции и анализа и программой внешнего управления. Входной поток вычислительного кластера составляет 5 MB/s, что соответствует характеристикам стандартной сетевой технологии.

Требование относительно большого количества вычислительных процессоров, естественно, коррелирует со стоимостью реализации и определяется оптимальным соотношением стоимость/характеристики и возможностями модернизации на основе дополнительных процессоров или заменой процессоров более современными и производительными. Анализ показывает что оптимальным с этой точки зрения является использование стандартных доступных в промышленном производстве процессоров. Программные средства и операционные системы должны быть совместимыми с операционной системой, используемой в центральном вычислительном комплексе, а именно UNIX, принятой для эксперимента.

Структура и оборудование вычислительного кластера четвертого уровня

Вычислительный кластер создан на базе стандартных PC-процессоров с операционной системой Linux. Это позволяет разместить оборудование практически независимо от расположения источников информации. Каждый PC-процессор оборудован 13 GB постоянным диском, 128 MB операционной памятью на процессор и упрощенной графической картой. В результате анализа было выбрано наиболее эффектив-

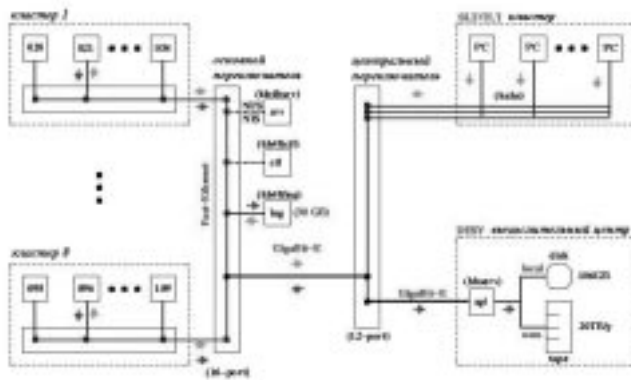


Рис. 6. Общая структура сетевой организации вычислительных средств эксперимента

ное решение - использовать платы с двумя процессорами Pentium III с тактовой частотой 500 MHz. Оба процессора используют одну и ту же интерфейсную карту для подключения к системе. Принцип распределения общей памяти неисключительный. Снижение производительности за счет данного принципа незначительное.

В качестве сетевой технологии выбран быстрый Ethernet - 100 MB/s. Группа из 16 PC процессоров объединены в миникластеры (рис. 6) и объединяются общим сетевым переключателем, далее миникластеры объединяются для соединения к сетевому оборудованию DESU с помощью кластер-переключателя, и далее к трем центральным вычислительным системам. Одна из вычислительных систем - сервер - обеспечивает пользовательское программное обеспечение на основе NSF-сервера (Network File System). Две другие центральные вычислительные системы используются как пользовательские и для контроля кластеров.

Система из 150 PC (по два процессора в каждом) обеспечивают компактную приборную структуру, не требующую значительного обслуживания.

СТРУКТУРА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ МОДЕЛИРОВАНИЯ И АНАЛИЗА ЭКСПЕРИМЕНТА HERA-B

Принципы и система программного обеспечения

Описание структуры и программных средств крупномасштабного эксперимента - достаточно сложная задача. Программная система, над которой работают несколько десятков системных и программных экспертов высокого уровня, распределенных во времени и пространстве по всей планете, и предназначенная для обработки огромного количества информации, представляет динамическую систему, создаваемую и совершенствуемую в течение всего времени активности эксперимента. Приведем несколько примеров структурных элементов системы программного обеспечения эксперимента HERA-B, основу которого составляет программный пакет ARTE [12].

ARTE - система программных средств анализа и реконструкции - представляет собой управляющую программу, контролирующую как процесс работы всех составляющих элементов программной системы эксперимента HERA-B в реальном времени, так и полную обработку и анализ информации. Область контроля составляет уровни III-IV триггеров эксперимента, реконструкцию в реальном времени и окончательный анализ, монте-карло- моделирование и физический анализ. Система осуществляет управление входными и выходными потоками информации, динамическое управление распределенной памятью, контроль над процессами в реальном времени, контроль над графическими процессами отображения информации. В ARTE включе-

ны интерфейсы к основным программным пакетам физики высоких энергий. ARTE обеспечивает общую структуру данных эксперимента и определения, включающие интерфейс к программам на базе языков программирования Fortran, C, C++.

Основные данные, характеризующие систему.

- Операционная система UNIX - ARTE создана и протестирована под операционной системой UNIX.

- Базовые платформы - Silicon Graphics, Hewlett-Packard и IBM RS6000, кластеры на базе процессоров PIII.

- Языки программирования ARTE были разработаны на основе Fortran77, являющегося официальным программным языком в физике высоких энергий. Структура позволяет эффективно использовать программы на основе программных языков C и C++. В настоящее время ядро системы переносится на C++.

- Управление кодами - на основе CMZ- и CVS-систем управления и распределения кодов.

- Основные программные пакеты и библиотеки: CERNLIB, HBOOK, PAW, HIGZ, LUND - пакет генераторов физических событий (теоретического отдела университета Лунд, Швеция), FRITIOFF, JETSET, GEANT, ZEBRA - система динамического управления памятью.

В рамках данной структуры были предложены оптимальные форматы данных для монте-карло-анализа и детектора переходного излучения, а также оптимальные алгоритмы математического моделирования эксперимента HERA-B на основе метода Монте-Карло, позволяющие производить анализ моделирования идентичным анализу экспериментальных данных не только на программном уровне, но и с использованием идентичных компьютерных средств. Данный метод позволяет исследовать эффективность практически всех уровней анализа данных включая триггерные уровни, повысить производительность генерации и анализа Монте-Карло данных на порядок по времени.

Математическая модель спектрометра HERA-B

Одним из основных элементов системы моделирования и анализа эксперимента HERA является интерфейс с программным пакетом моделирования GEANT, который содержит описание и моделирование экспериментальных установок в физике высоких энергий. HBGEANT состоит из комплекса подпрограмм и функций, являющихся дополнением к пакету программ CERN GEANT, который обеспечивает

- полное описание структуры геометрии спектрометра HERA-B для использования в ARTE и GEANT;

- трассировку элементарных частиц через спектрометр с учетом взаимодействий с веществом детекторов;

- 3D-графическое изображение спектрометра и отдельных элементов и треки частиц, проходящих спектрометр как в режиме реального времени, так и при окончательном анализе,

- CAD-файлы для проектирования элементов спектрометра.

Как было отмечено, математическое описание переходного излучения представляет собой специфику в силу отсутствия как такового в программах моделирования включая GEANT. Для включения этого раздела в математическую модель эксперимента было проведено тщательное исследование, в результате которого было создано программное обеспечение, включенное в официальное программное обеспечение эксперимента HERA-B [n].

Система анализа и визуализации данных в эксперименте HERA-B

Обеспечение процесса интерактивной обработки и анализа данных (экспериментальных и моделируемых) и их сравнение приобретает все большее значение в экспериментах по физике высоких энергий.

Основные требования.

1. Значительные потоки информации, которые необходимо размещать на постоянные носители долговременного хранения в течение времени набора статистики. Архивация данных, поступающих с эксперимента, как было отмечено выше, достигает при скорости набора статистики 100 кВ/событие и среднем времени набора статистики 10^7 s/год объема 20 TB в год. Моделированные события того же класса значительно превышают по объему экспериментальные данные, т.к. необходимо сохранять полную информацию о генерированном событии. Обработка событий в этом случае может значительно замедлиться из-за ограничений в доступе к центральному архиву на магнитных лентах 5 MB/s.

2. Значительные затраты процессорного времени для обработки экспериментальных и моделированных событий. События с адронным взаимодействием, имеющие многопараметрические пространственные распределения, полученные с экспериментальной установки, которые существенно влияют на сложность алгоритмов, используемых при реконструкции - поиск треков, восстановление треков, идентификацию частиц. Адаптивные алгоритмы, разработанные для реконструкции экспериментальных событий, требуют процессорного времени 2-5 s/событие на современных процессорах. Поскольку моделированные события значительно превышают по объему экспериментальные данные, как было отмечено ранее, то время обработки также значительно возрастает.

3. Интерактивная визуализация реконструируемых событий имеет огромное значение при реконструкции физических событий, т.к. до сих пор не найдены алгоритмы формализации принципиально новых классов взаимодействий, которые и представляют наибольший интерес для исследования.

Приведенные условия требуют огромного количества процессоров для обеспечения приемлемого соотношения между скоростью набора информации, архивированием и реконструкцией событий. С другой стороны, повторное считывание событий из архива и распределение их между процессорами является нелогичным. С целью

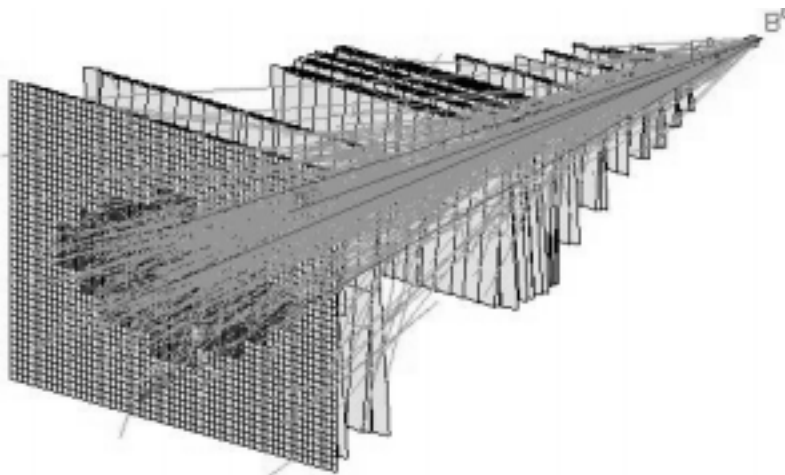


Рис. 7. Трехмерное представление моделируемого физического события эксперимента HERA-B: представлены основные компоненты системы координатных детекторов и электромагнитный калориметр, треки частиц взаимодействия протонов 920 GeV/c с углеродной мишенью

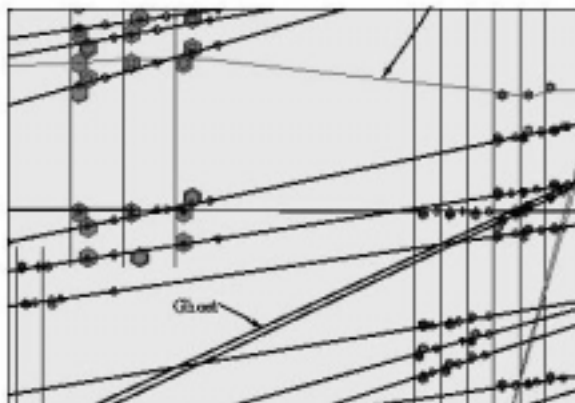


Рис. 8. Визуализация областей координатного детектора эксперимента HERA-B с реконструированными треками частиц и выделенными параметрами детектора: сотовая форма представляет собой отдельный координатный детектор и окрестность, внутри измеряемый параметр - время дрейфа

оптимизации процессов обработки и анализа данных было принято решения создать наряду со стандартной системой обработки на основе центрального вычислительного комплекса специализированный кластер с оптимизированными характеристиками для он-лайн и интерактивной обработки и классификации физических и моделированных событий на основе технологии клиент-сервер, приведенной выше.

На рис. 7 приведено 3D-представление одного события эксперимента HERA-B с полной реконструкцией, выделена топология полезного события с распадом В-мезона в определенное CP конечное состояние через рождение резонансного состояния очарованных кварков.

На рисунке представлена часть основного координатного детектора - основные плоскости и электромагнитный калориметр. Событие представляет собой взаимодействие протонов с энергией 920 GeV с веществом мишени - углеродом. Показаны вторичные треки взаимодействия в полном объеме детектора с учетом всех процессов взаимодействия в детекторах. Количество треков в детекторе достигает 1500 при требовании выделения и анализа физического канала распада, состоящего из 4 треков.

Возможности и методы работы с системой визуализации приведены на рис. 8, где показаны фрагменты того же события при визуализации более подробной информации - диапазон дискретизации системы визуализации составляет 2^{16} .

На данном уровне можно видеть также введение дополнительной информации о положении точки взаимодействия конкретного элемента экспериментальной установки - в данном случае отдельные ячейки дрейфовых камер, дающих координатную информацию, отображенную в виде эквипотенциальной окружности.

ЗАКЛЮЧЕНИЕ

Рассмотрены тенденции создания современных вычислительных систем и систем программного обеспечения для крупномасштабных экспериментов в физике высоких энергий в рамках «триады» А.А.Самарского «модель-алгоритм-программа».

Рассмотрены основные особенности современного состояния как компьютерного оборудования, так и программного обеспечения на примере эксперимента HERA-B, представляющего собой один из наиболее развитых экспериментов в данной области. Параметры системы сбора и обработки данных эксперимента HERA-B сравнимы с аналогичными системами экспериментов следующего поколения в физике высоких энергий, в частности экспериментов на LHC CERN (Европейский центр ядерных ис-

следований, Женева).

В рамках структуры компьютерной системы эксперимента HERA-B были предложены оптимальные форматы данных для монте-карло-анализа и детектора переходного излучения, а также оптимальные алгоритмы математического моделирования эксперимента HERA-B на основе метода Монте-Карло. Это позволяет производить анализ данных моделирования методами и средствами, идентичными анализу экспериментальных данных не только на программном уровне, но и с использованием идентичных компьютерных средств. Данный метод позволяет исследовать эффективность практически всех уровней анализа данных, включая триггерные уровни, повысить производительность генерации и анализа монте-карло-данных на порядок по сравнению с традиционными методами.

Система компьютерного и программного обеспечения выполняет сбор, анализ экспериментальных данных, математическое моделирование и анализ данного эксперимента на планируемом уровне. Система обеспечила набор экспериментальных данных в течение 1999-2001 гг. на уровне 10 ТБ, и в настоящее время производится их физическая обработка по анализу сечения рождения В-мезонов. Статистика для моделируемых событий составляет суммарно несколько миллионов событий в соответствии с физической программой эксперимента.

Литература

1. Самарский А.А., Вабищевич П.Н. Математическое моделирование и вычислительный эксперимент. - М., 2000.
2. HERA-B An Experiment to Study CP Violation in the B-System Using an Internal Target at the HERA Proton Ring. Proposal // DESY-PRC 94/02, DESY, 1994.
3. Pulling the trigger on LHC electronics // CERN COURIER. – 2001. – V. 41. - 1.
4. Ермаков С.М. Метод Монте-Карло и смежные вопросы. - М.: Наука, 1975.
5. CERN Programm Library, CERN, CN Division. /R. Brun, F. Bruyant, A. C. McPherson, P. Zancarini DD/EE/84-1, 1987.
6. Гинзбург В., Франк И. Излучение равномерно движущегося электрона, возникающее при его переходе из одной среды в другую // ЖЭТФ. – 1946. - Т. 16. – Вып. 1.
7. Егорычев В.Ю., Савельев В.И. Монте-Карло моделирование переходного излучения и идентификация электронов для эксперимента HERA-B: Препринт ИТЭФ. – 1999. - 11.
8. Самарский А.А., Четверушкин Б.Н. Использование и перспективы применения многопроцессорных транспьютерных систем для решения задач науки и техники: Сб. трудов институтов ОИВТА РАН, 1994.
9. Cerbal J.R. Collaborative visualization for parallel multidisciplinary applications, Elsevier Amsterdam, 1998.
10. Defert Ph., Cancio G., Cons L. at all. ASIS, mamage and distribute application software in the HEP community: Proc. of the Intern. Conf. on Computing in High Energy Physics (Berlin, 1997).
11. ADSP-21060x User's Manual, Analog Devices, 1995.
12. HERA-B Software Group \\ APTE - The Event Reconstruction and Analysis Tool for HERA-B \\ DESY, 1996.

Поступила в редакцию 2.02.2002

nation of the absorbed dose in the thyroid. The thyroid dose coefficient calculations for the ^{131}I , ^{132}I and ^{133}I . The proposed models and methods are based on the data obtained during the analysis of the case histories of about 400 patients, which have passed the radioiodine therapy course.

УДК 519.7: 621.039.643

Automodelling solutions of two-temperature magnetohydrodynamics equations \ P.P. Volosevich, E.I. Levanov; Editorial board of Journal "Izvestia visshikh uchebnikh zavedeniy, Yadernaya energetika" (Communications of Higher Schools. Nuclear Power Engineering) – Obninsk, 2002. – 10 pages. – References, 20 titles.

The equations of two-temperature magnetohydrodynamics are essential in mathematical simulation of some actual problems of plasma physics. The analysis of automodelling solutions of two-temperature magnetohydrodynamics equations including electrons thermoconductivity, electron-ion relaxation and electroconductivity is given in the article.

УДК 519.7: 621.039.51

Mathematical simulation and numerical experiment for nuclear-physics researches \ V.I. Saveliev; Editorial board of Journal "Izvestia visshikh uchebnikh zavedeniy, Yadernaya energetika" (Communications of Higher Schools. Nuclear Power Engineering) – Obninsk, 2002. – 14 pages, 8 illustrations. – References, 12 titles.

Some aspects of creation of systems for mathematical simulation and analysis of large-scale nuclear-physics experiments are given in this work. The questions of systems solution of hardware choices, algorithms and structure of software are considered in details. There is also considered the specific example of experiment – HERA-B which permits to demonstrate the main tendencies of creation of system for mathematical simulation and analysis at the modern stage.

УДК 519.7: 539.1.03

Simulation of transition radiation for periodical and spatial distributed structures \ S.J. Aplin, V.I. Saveliev; Editorial board of Journal "Izvestia visshikh uchebnikh zavedeniy, Yadernaya energetika" (Communications of Higher Schools. Nuclear Power Engineering) – Obninsk, 2002. – 12 pages, 8 illustrations. – References, 12 titles.

The mathematical model for description of transition radiation appearing when charged particle crosses the border of two mediums with different dielectric properties is given in the article. This model can be applied for investigation of transition radiation from moving charged particles in periodical ultra dense mediums – intranuclear arias.