

# КОМПЬЮТЕРНОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ СИСТЕМ ОБРАБОТКИ ДАННЫХ ЭЛЕКТРОННЫХ ЭКСПЕРИМЕНТОВ В ФИЗИКЕ ВЫСОКИХ ЭНЕРГИЙ

*В.Г.Иванов*

Объединенный институт ядерных исследований, Дубна

Компьютерное и программное обеспечение электронных экспериментов в физике высоких энергий представляет собой большие и сложные системы (среды), в которых взаимодействуют люди, вычислительные системы, экспериментальные установки, программы, алгоритмы и информация. Эти системы создаются на базе мощных вычислительных систем, рабочих станций, компьютерных средств связи, современных технологий программирования и средств искусственного интеллекта.

Целью настоящего обзора является рассмотрение проблем, связанных с созданием таких систем: специфика обработки и анализа данных в экспериментах по физике высоких энергий и требования к вычислительным ресурсам, характеристики вычислительных центров лабораторий; роль больших ЭВМ, распределенных вычислительных систем и рабочих станций; появление RISC-компьютеров (Reduced Instruction Set Computers) и их влияние на вычислительные системы электронных экспериментов; задачи и структура программного обеспечения и возможные пути его развития; компьютерная стратегия нового поколения электронных экспериментов в современных условиях.

Computer and software support of the electronic experiments in high energy physics is realized as environments in which different computer systems, detectors, programs, algorithms and human beings interact. These systems are created on the base of mainframes, distributed computing, power graphic workstations and networks using CASE (Computer Aided Software Engineering) and artificial intelligence tools.

The task of the review is a consideration of ways for solving the computing problem in the large-scale electronic experiments, the computer resources requirements, central computer facilities in the HEP centers, the arrival in the market place RISC-computers (Reduced Instruction Set Computers) and the influence of them on the computer systems in HEP, distributed computing and software development methodologies for the new generation of experiment in high energy physics.

Для обработки экспериментальных данных, получаемых в современных электронных экспериментах в физике высоких энергий, требуются не только мощные вычислительные системы и большие программные

комплексы, но также решение большого круга задач, связанных с огромными объемами получаемой информации, организацией работы больших международных коллабораций и обмена данными между их участниками, оснащением участников работ эффективными средствами анализа получаемых данных, позволяющими получать физические результаты вскоре после завершения набора статистики. Так, например, в проводившемся в Лаборатории им. Ферми (США, Батавия) эксперименте E687 в 1990—1991 гг. получено более 500 млн. случаев фоторождения. Их реконструкция была завершена через семь месяцев после окончания сеансов на ускорителе. Естественно, что обработка такого огромного количества данных невозможна без создания высокопроизводительной и эффективной компьютерной среды.

Суть проблемы, связанной с применением вычислительных и программных средств в экспериментальных исследованиях, была сформулирована при подготовке экспериментов на большом электрон-позитронном коллайдере (LEP) CERN следующим образом: «С самого начала планирования экспериментов на большом электрон-позитронном коллайдере CERN было известно, что существует проблема их обеспечения вычислительными ресурсами. Как показали дискуссии 1984 и 1985 гг., в течение ряда лет было широко распространено неправильное представление о том, что эта проблема носит исключительно финансовый характер. Поэтому дискуссии того времени в основном сводились к обсуждению расхождения между имеющимися и требуемыми вычислительными мощностями. Однако в настоящее время ситуация в понимании этой проблемы существенно изменилась. Стало ясно, что проблема является более сложной, чем казалось ранее, и ее решение заключается не только в наращивании вычислительных мощностей для пакетной обработки. Наряду с последними, необходимо также рассматривать вопросы массовой памяти; роль, которую могут играть процессорные фермы; сети передачи данных, как внутри экспериментальной лаборатории, так и за ее пределами; а также рабочие станции для интерактивного анализа получаемых результатов» [4].

К этому следует добавить программное обеспечение этих систем и весь связанный с ним комплекс указанных выше проблем. Иначе говоря, разработка компьютерного обеспечения современных экспериментов по физике высоких энергий заключается в создании среды, в которой взаимодействуют люди, вычислительные системы, экспериментальные установки, алгоритмы и информация. В дальнейшем эту среду будем называть компьютерингом.

Таким образом, под компьютерингом в физике высоких энергий (ФВЭ) понимается широкий спектр проблем, связанных с применением ЭВМ для решения различного рода задач в этой области исследований. О том,

что понимается в настоящее время под этим термином, можно судить по тематике докладов на специальных конференциях «Computing in High Energy Physics», которые регулярно проводятся начиная с 1980 г., а также конференциях, посвященных вопросам разработки детекторов [1—12].

Большое внимание на этих конференциях уделяется не только вопросам собственно компьютерного и программного обеспечения исследований, но и всему комплексу проблем, так или иначе связанных с применением ЭВМ в этой области научных исследований. Это и вычислительные комплексы физических центров, объединяемые линиями связи в распределенные вычислительные системы, и их программное обеспечение, языки и методы программирования, системы аналитических вычислений, алгоритмы и методы решения физических задач, и вопросы применения в исследованиях экспертных систем и других элементов искусственного интеллекта. Так, например, на компьютерной конференции 1992 г., помимо пленарных заседаний; посвященных рассмотрению общих проблем ФВЭ-компьютинга, работало 15 параллельных секций, посвященных рассмотрению следующих вопросов [12]:

- Триггирование и сбор данных.
- Библиотеки и пакеты.
- Вычисления, архитектура и опыт.
- Сети передачи данных.
- Компьютерные помощники для конструирования детекторов.
- Графика и визуализация.
- «Дешевая» пакетная обработка.
- Объектно-ориентированные приложения.
- Базы данных.
- Средства CASE (Computer Aided Software Engineering) технологии.
- Открытое программное обеспечение и распределенные вычисления.
- Нейронные сети.
- Высокоскоростные связи.
- Технология массовой памяти.
- Управление данными.

Целью настоящего обзора является рассмотрение процесса обработки данных в экспериментах по физике высоких энергий, с точки зрения их требований к компьютерингу, включая методы оценки необходимых вычислительных ресурсов, характеристики различных вычислительных центров лабораторий физики высоких энергий и специализированных вычислительных систем; определение роли больших ЭВМ и распределенных вычислительных систем. На примере экспериментов на LEP рас-

смачивается эволюция распределенных систем обработки, начиная с первоначальных замыслов и кончая реализацией. Вопросам компьютерной стратегии и методологии разработки программного обеспечения для нового поколения экспериментов посвящен заключительный раздел обзора.

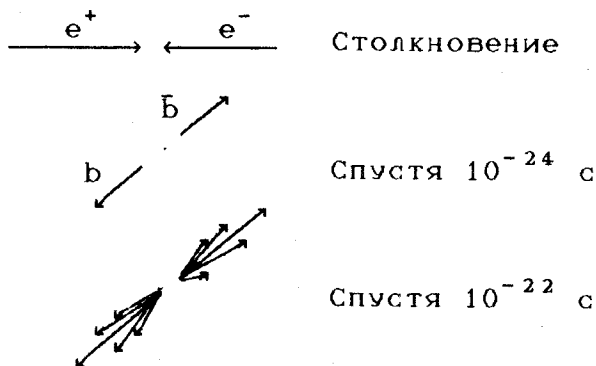
## 1. НЕКОТОРЫЕ ОСОБЕННОСТИ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ В ФИЗИКЕ ВЫСОКИХ ЭНЕРГИЙ

Объектом исследований в физике высоких энергий является квантовый мир элементарных частиц. Для проведения этих исследований необходимы мощные ускорители на высокие и сверхвысокие энергии; сложные и дорогостоящие детекторы; мощные вычислительные системы с устройствами массовой памяти, способные обрабатывать и хранить огромные потоки информации, сети передачи данных, обеспечивающие оперативную связь участников больших международных коллабораций, и соответствующее программное обеспечение.

Вследствие этих обстоятельств вычислительные системы, обеспечивающие обработку и анализ экспериментальных данных, и сети передачи данных стали рассматриваться, наряду с детекторами и ускорителями, как необходимые элементы экспериментов [20].

1.1. Специфика измерений. Требования, которые предъявляются к системам обработки экспериментальных данных, вычислительным ресурсам, устройствам массовой памяти и сетям передачи данных, постоянно растут из-за увеличивающихся объемов экспериментальных данных. Необходимость обработки все больших объемов информации возникает не только из-за сложности индивидуальных столкновений частиц при больших энергиях, но также и из-за недетерминистского характера наблюдаемых результатов. Последнее обусловлено тем, что даже при полном понимании физических процессов, происходящих при столкновении частиц высоких энергий, мы могли бы предсказать только вероятности различных конечных состояний (типы частиц, их импульсы и направления), которые могут возникать, например, в электрон-позитронных столкновениях при энергии  $\sim 100$  ГэВ в системе центра масс. Но так как в действительности физика таких явлений известна недостаточно хорошо, то для проверки теоретических предсказаний необходимо экспериментально измерять вероятности различных процессов, что требует регистрации десятков, а то и сотен миллионов исследуемых событий [16,143].

1.2. Искажение результатов взаимодействий в процессе измерений: необходимость моделирования. Специфика изучаемых процессов и

Рис. 1. Схема процесса  $e^+e^- \rightarrow b\bar{b}$ 

регистрирующая аппаратура не позволяют нам наблюдать картину того, что на самом деле происходит при столкновении частиц высоких энергий. Рассматривая кварки как фундаментальные частицы, мы тем не менее не можем наблюдать их непосредственно, т.к. из-за быстрого протекания физических процессов невозможно зафиксировать и непосредственно измерить характеристики первичных взаимодействий. Это обстоятельство иллюстрируется рис.1, где показаны три «моментальные» фотографии  $e^+e^-$ -столкновения, в результате которого образовалась пара кварк — антикварк ( $b\bar{b}$ ), которая «фрагментирует» в струи других частиц за такой интервал времени, в течение которого невозможно произвести какие-либо измерения. Но даже если мы и смогли бы обчислить этот процесс и измерить характеристики всех образовавшихся частиц, то все равно не смогли бы идентифицировать струи от фрагментации  $b\bar{b}$ -кварков для каждого рассматриваемого события. Эта задача решается на статистической основе с использованием специфических особенностей процесса фрагментации именно этих кварков [16,17].

Фрагментация кварков является характерным примером того, как физика процесса искажает первоначальную картину результата взаимодействия, которую нужно восстановить по результатам измерений продуктов вторичных реакций.

Несовершенство детекторов также вносит свой вклад в искажение регистрируемых данных, что обусловлено целым рядом причин («мертвые области», недостаточное разрешение, регистрация фона и т.п.) [17,19].

Поэтому естественным подходом в данном случае является моделирование физических процессов для изучения связи между ними и наблюдаемой в экспериментальной установке «картиной». Это позволяет оценить возможности обнаружения конкретных процессов (гипотез) и найти критерии для их статистического выделения.

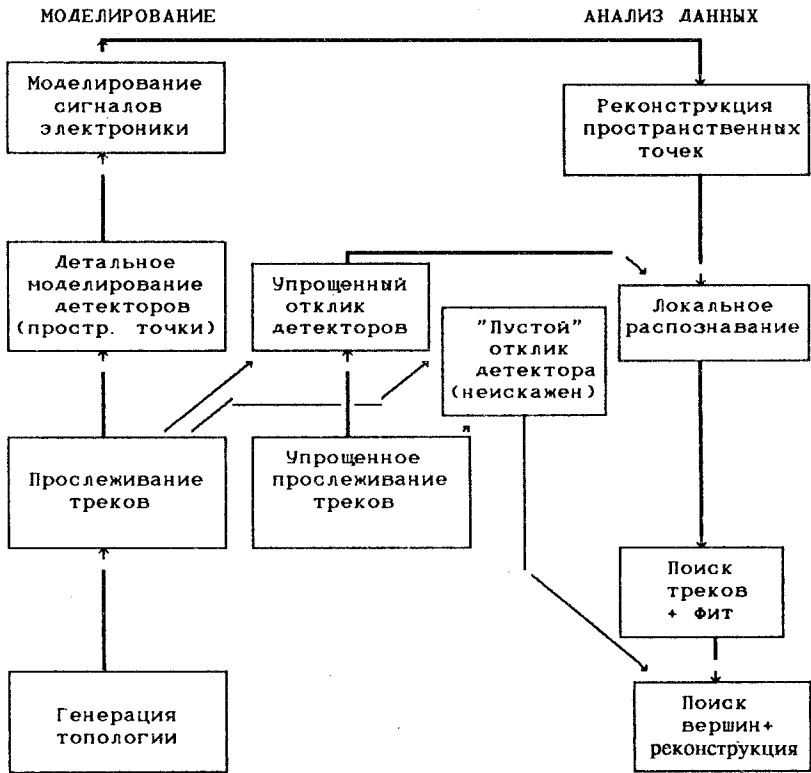


Рис. 2. Связь моделирования и обработки данных

В связи с этим детектор рассматривается как устройство, на вход которого поступают частицы, а на выходе появляются сигналы, а его свойства описывают с помощью функции отклика, которая определяет вероятность возникновения в детекторе сигнала с данными характеристиками при попадании в него частицы с заданными свойствами [18]. Для решения этой задачи в процессе моделирования генерируются изучаемые события и результаты их «измерений» путем проведения сгенерированных частиц через регистрирующую аппаратуру. Это позволяет определить связь между величинами, которые мы можем измерить, и той фундаментальной физикой, которую мы хотим понять. На основе сравнения результатов моделирования и эксперимента делаются выводы о характере наблюдаемых процессов. Вследствие этого в физике высоких энергий и возникло правило, которое, правда, не носит абсолютного ха-

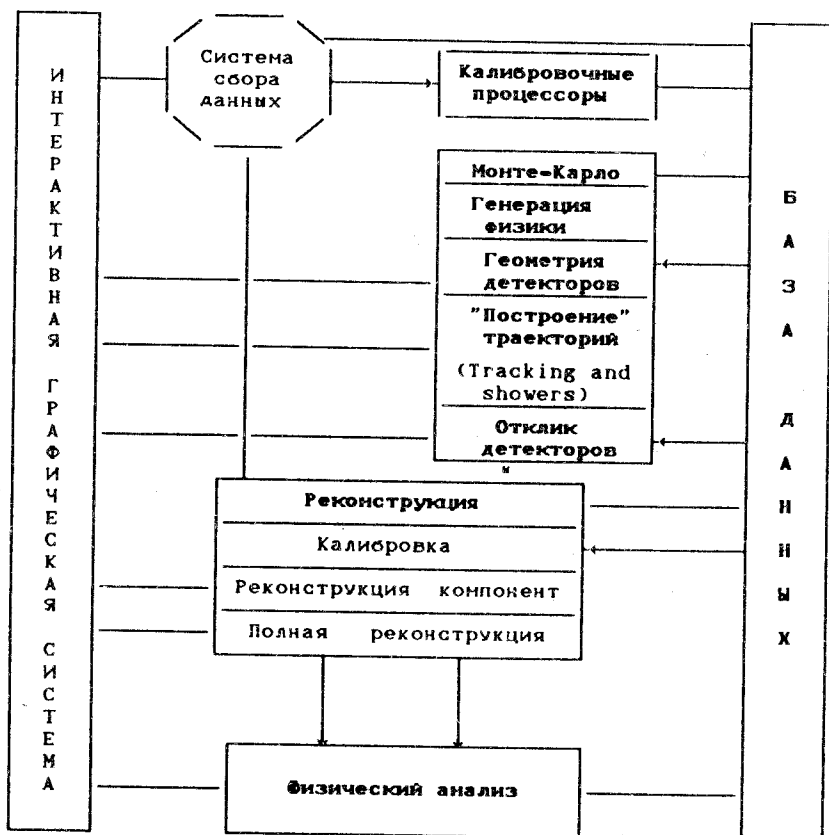


Рис. 3. Принципиальная схема анализа данных

рактера: «В экспериментах в физике высоких энергий нет результата без моделирования» [17].

Затраты вычислительной мощности на моделирование по полной схеме, начиная с задания физических гипотез и кончая реконструкцией «результатов измерений», достаточно велики. Поэтому при решении ряда задач для их сокращения используют «частичное» или «быстрое» моделирование, когда моделируется только часть процесса. Например, для оценки вероятности регистрации тех или иных частиц в детекторе достаточно получить их характеристики и «провести» через установку. Принципиальная разница между «полным» и «быстрым» моделированием показана на рис.2 [4].

**1.3. Схема физического анализа данных.** Принципиальная схема процесса анализа данных в экспериментах по физике высоких энергий показана на рис.3 [16,17].

После сбора, фильтрации и предварительной обработки данных, регистрируемых детекторами установки, производимой системой сбора данных, информация, относящаяся к одному событию, объединяется в единый массив данных и записывается на машинные носители.

Необходимые для реконструкции событий параметры установки и взаимное расположение ее элементов относительно друг друга определяются в ходе калибровки установки и заносятся в базу данных (БД), к которой имеют автоматический доступ программы реконструкции, моделирования и анализа.

В ходе реконструкции восстанавливается пространственная картина события и находятся параметры составляющих его частиц.

На завершающей стадии физического анализа производится сравнение распределений и корреляций этих основных величин для реальных и генерированных событий.

Основным рабочим инструментом физиков, занимающихся анализом данных, являются мощные графические рабочие станции, с помощью которых производится отбор и анализ событий, изучается их пространственная картина.

**1.4. Некоторые вопросы организации коллабораций.** Вследствие увеличения размеров ускорителей и энергий сталкивающихся частиц экспериментальное оборудование, применяемое для исследований в области физики высоких энергий, стало большим, сложным и дорогостоящим. Вес современных детекторов составляет тысячи тонн, занимаемое пространство — тысячи кубометров, число каналов быстрой электроники — сотни тысяч, стоимость — десятки и сотни миллионов долларов [21].

Суммарные объемы исходных данных и результатов достигают десятков терабайт. Обработка и анализ таких огромных потоков информации требуют колоссальных вычислительных мощностей (измеряемых тысячами, а для планируемых в настоящее время экспериментов и миллионами машин производительностью 1 MIPS), привлечения сотен высококвалифицированных специалистов [16,20].

Вследствие этих обстоятельств экспериментальные исследования в физике высоких энергий проводятся в небольшом числе ядерных центров [20]. Так, например, в настоящее время планируется сооружение только одного большого адронного коллайдера (LHC) в CERN с энергиями сталкивающихся пучков протонов  $7 + 7$  ТэВ, хотя еще в 1993 г. речь шла о создании двух коллайдеров такого типа.

В подготовке и проведении экспериментов участвуют сотни специалистов из десятков институтов и организаций различных стран.



Так, например, в L3-эксперименте участвуют около 500 физиков из 42 институтов и университетов 14 стран [16].

Программное обеспечение экспериментов составляет сотни тысяч кодов, а его разработка требует нескольких сотен человеко-лет.

Для того чтобы обеспечить обработку и анализ данных силами больших международных коллабораций, необходимо сначала зарегистрировать исходные данные, произвести реконструкцию событий, обработать результаты реконструкции и распределить их между участниками коллаборации для физического анализа.

Таким образом, с одной стороны, имеется база данных и десятки тысяч носителей информации, а с другой — сотни физиков из десятков институтов различных стран. Проблема заключается в организации распределенной системы обработки данных, охватывающей десятки стран различных континентов. Естественно, что при решении этой проблемы возникает большой комплекс технических, организационных и социологических задач [17].

Для организации эффективной работы больших международных коллабораций в 1989 г. предложено разделить их участников на следующие три категории: основной центр или экспериментальная часть, региональные или национальные центры и университеты [2,4].

Основной центр (как, например, в эксперименте L3) обеспечивает работу ускорителя и часть наиболее важной технической поддержки эксперимента, но лишь в небольшой степени — людские ресурсы [16]. Что касается реконструкции событий, то первоначально предполагалось, что в таком основном центре, каким является CERN, будут обрабатываться 1/3 всех зарегистрированных событий, храниться исходные данные, суммарные результаты Data Summary Tape (DST), базы данных результатов калибровки, а также он будет распространять информацию между участниками коллаборации.

Региональные или национальные центры должны вносить существенный вклад в процесс подготовки и проведения эксперимента, обработку и анализ результатов [11].

Что касается университетов или небольших институтов, то они решают отдельные задачи в соответствии с имеющимися у них ресурсами и обеспечивают подключение к исследованиям интеллектуального потенциала вузовской науки.

Однако, благодаря появившейся вскоре возможности создания сравнимых по мощности с большими ЭВМ недорогих вычислительных систем на базе рабочих станций, производительность которых возрастала быстрыми темпами, ситуация вскоре коренным образом изменилась. Так, например, в экспериментах на LEP появились собственные

вычислительные системы, позволяющие производить реконструкцию событий в течение нескольких часов после их регистрации [13,16].

## 2. ОБЪЕМЫ ДАННЫХ И ПОТРЕБНОСТИ В ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСАХ В ЭКСПЕРИМЕНТАХ ПО ФИЗИКЕ ВЫСОКИХ ЭНЕРГИЙ

Компьютинг в физике высоких энергий отличается от многих других типов «научного компьютеринга» тем, что здесь большое внимание необходимо уделять не только собственно вычислениям, но и проблеме организации процесса обработки [16]. Необходимость этого диктуется огромными потоками исходных данных и получаемых результатов, большими объемами вычислений и участием в обработке и анализе данных сотен специалистов из многих институтов различных стран. Для того чтобы оценить и понять те проблемы, с которыми придется столкнуться в процессе обработки больших объемов данных, необходимо научиться его моделировать и, в первую очередь, оценивать объемы получаемых результатов и необходимые вычислительные ресурсы.

Следует отметить, что в экспериментальной физике достаточно сложно планировать необходимые вычислительные ресурсы, особенно тогда, когда речь идет о новом поколении экспериментов. Это обусловлено тем, что многое может измениться не только в ходе подготовки эксперимента, но также и во время его проведения. Поэтому результаты моделирования и различного рода оценки носят ориентировочный характер, но тем не менее они позволяют заранее оценить те проблемы, с которыми придется встретиться, и найти способы их решения [4].

**2.1. Методика оценки вычислительных ресурсов для off-line обработки данных электронных экспериментов.** Рассмотрим принципиальный подход к методике оценки вычислительных ресурсов для off-line обработки данных в экспериментах на примере электрон-протонного коллайдера HERA (DESY), который приведен в книге «Computing at CERN in the 1990s» [12].

В основу положены следующие рассуждения.

1. Установка в среднем будет регистрировать одно событие в секунду и работать шесть месяцев в году. Тогда общее число событий, получаемых на ней ежегодно, составит  $1,5 \cdot 10^7$ . Если на DST будет отбираться  $1,5 \cdot 10^6$  полезных событий, а средние длины событий на лентах исходных данных и DST — 100 и 20 кбайт соответственно, то суммарный объем получаемых ежегодно исходных данных составит 1500 Гбайт, а DST — 30 Гбайт. При затратах машинного времени 20 с на реконструкцию одно-

го события на ЭВМ типа IBM 370/168 на реконструкцию всех событий потребуется 12 таких машин (12 «церновских» единиц вычислительной мощности), если каждая ЭВМ будет работать 290 дней в году.

2. Методом Монте-Карло будет генерироваться столько же событий, сколько реальных событий будет отбираться на DST, т.е.  $1,5 \cdot 10^6$  монте-карловских событий в год. Из них 30% — по полной схеме и 70% — по быстрой. При длине событий, генерируемых по полной схеме, 100 кбайт, а по частичной — 20 кбайт, средняя длина одного монте-карловского события на DST составит 44 кбайт, а ее объем — 66 Гбайт. При среднем времени генерации одного МК события 60 с на генерацию и последующую реконструкцию событий потребуются 4 единицы вычислительной мощности. Процедура моделирования указанного числа событий будет повторяться 2,5 раза.

3. На чтение DST для обеспечения работы одного физика потребуется около часа работы машины, а 200 физиков с учетом эффективности работ ЭВМ (0,85) — около 200 часов в день или 10 «церновских» единиц мощности (200 часов/24/0,85).

Таким образом, для off-line обработки данных при таких предположениях потребуется 32 единицы вычислительной мощности и ~ 1800 Гбайт памяти для хранения исходных данных и полученных результатов.

Этот пример иллюстрирует принципиальный подход к оценке вычислительных ресурсов и набор предположений (моделей) и величин, которые нужно задать для получения искомым значений.

Естественно, что в реальной жизни ситуация может быть сложнее. Поэтому для моделирования этого процесса и исследования влияния различных факторов на интересующие физика величины необходимо сформировать алгоритмическую модель процесса, включив в нее различные подмодели, задать исходные данные и выбрать программное средство для решения этой задачи, желательно, в интерактивном режиме.

Удобным инструментом для решения такого рода задач являются электронные таблицы [24]. Применение последних позволяет решать эту задачу даже на персональном компьютере типа IBM PC, получая результаты спустя несколько секунд после формирования в интерактивном режиме модели процесса и задания всех нужных данных.

Включив в таблицу данные о стоимости вычислительных систем и различных типов компьютеров, скоростях передачи информации по линиям связи, максимальных объемах данных, записываемых на носители различных типов, и т.п., можно решать весь круг задач, связанных с моделированием процесса off-line обработки данных, выбором

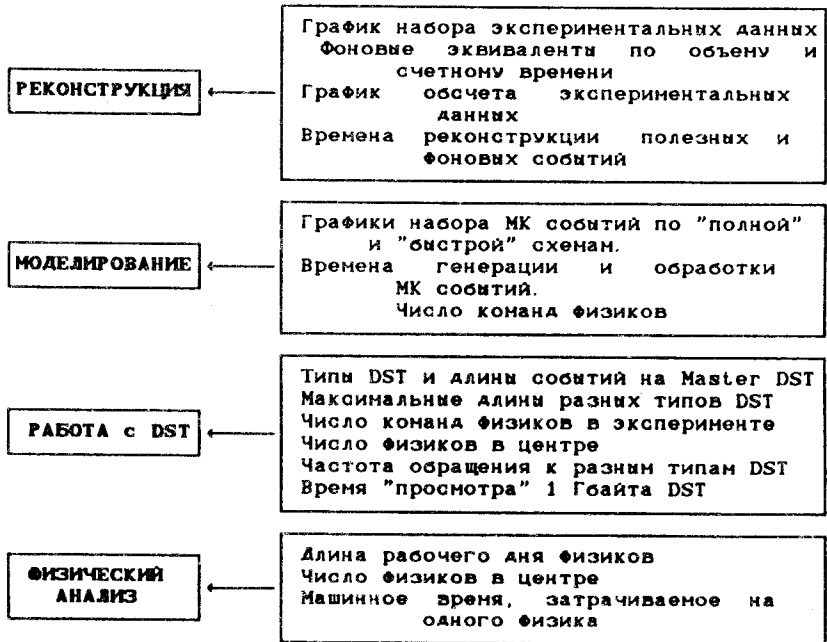


Рис. 4. Принципиальная схема модели процесса off-line обработки данных электронного эксперимента

оптимального варианта вычислительных систем с учетом их стоимости и т.п. Внося изменения в подмодели и исходные данные, можно исследовать влияние различных факторов на конечные результаты, представляемые в виде таблиц или графиков, через несколько секунд после ввода в таблицу соответствующих данных. Это позволяет эффективно и быстро решать задачу моделирования процесса off-line обработки, исследовать различные модели процесса и подбирать оптимальный вариант вычислительной системы [25]. Принципиальная схема алгоритмической модели этого процесса представлена на рис.4.

2.2. Основные требования к вычислительным ресурсам для экспериментов на LEP. Рассмотрению проблем, связанных с off-line обработкой экспериментальных данных и оценками компьютерных ресурсов, необходимых для проведения экспериментов на большом электрон-позитронном коллайдере CERN, было посвящено специальное совещание «Meeting to Understand the Specific Computing Needs of the LEP Experiments» (апрель 1987 года) [4].

Таблица 1. Рост вычислительных мощностей ВЦ CERN  
в период 1988—1991 гг.

ЭВМ	1988	1989	1990	1991
Совместимые с IBM	36	60	57	54
CRAY	32	32	32	90
DEC	6	10	10	10
Суммарная мощность	74	102	99	154

Таблица 2. Затраты CERN на различные проекты  
вычислительных средств в период 1989—1991 гг.  
в миллионах швейцарских франков

Проект	Стоимость		
	1989	1990	1991
Новые проекты систем реального времени	0,5	1	1
Диски для устройств массовой памяти	9	11	11
Автоматический загрузчик картриджей	1,4	0,6	2
Мощность больших CPU	2	21	25
Параллельная обработка	1	3	6
Территориальная сеть передачи данных			
Интерфейсы в CERN	1	1	1
Стоимость линий	4	10	12
Служба копирования лент	0,5	0,5	0,5
Инфраструктура (здания, архивы лент)	1	1	—
Итого	20,4	49,1	58,5

Основным результатом его работы явилось четкое понимание необходимости комплексного подхода к решению проблемы off-line обработки данных и масштаба требующихся вычислительных ресурсов.

В итоговом документе совещания и ряде последующих публикаций рассмотрена модель этого процесса для экспериментов на LEP и приведены сделанные на ее основе оценки необходимых вычислительных ресурсов [4,12,23].

Специалисты CERN подсчитали, что для реализации физической программы в 1991—1992 гг. необходимо 600 «церновских» единиц вычислительной мощности вместо имевшихся к концу 1988 г. 74 единиц.

Ориентировочная стоимость одной «церновской» единицы вычислительной мощности для больших ЭВМ (CRAY, IBM, VAX) в то время составляла ~ 300 тыс. долларов. В табл. 1 приведены данные о развитии вычислительного центра CERN в период 1988—1991 гг., а в

табл. 2 — стоимость различных вычислительных средств (в миллионах швейцарских франков), приобретенных CERN в 1988 и 1989 гг. и запланированных на 1990 и 1991 гг.

Из приведенных в таблицах данных видно, что даже значительное увеличение инвестиций на приобретение больших ЭВМ все равно не позволило бы CERN в течение нескольких лет выйти на требуемый уровень вычислительной мощности по большим ЭВМ.

Поскольку требования физической программы к вычислительным ресурсам значительно превосходили уровень планируемых расходов на эти цели, была предложена концепция распределенной обработки данных, в соответствии с которой CERN взял на себя обязанность обрабатывать не менее трети получаемых данных и обеспечивать остальных участников коллабораций информацией, необходимой для обработки остальной части данных и анализа DST [12].

CERN обязался для этой цели к 1992 г. увеличить мощность больших ЭВМ вычислительного центра до 150 единиц.

Остальные участники коллабораций должны были обеспечить 300 единиц, а оставшиеся 150 — параллельные вычислительные системы.

Наряду с этим были указаны следующие тенденции в эволюции организации распределенных вычислений.

1. Следующие пять лет будут периодом прогресса в развитии распределенной вычислительной среды, эксплуатации быстро развивающихся рабочих станций и сетей передачи данных.

2. Основными элементами в эволюции компьютеринга в CERN, национальных и региональных центрах и университетах будут следующие:

— рост вычислительной мощности для пакетной обработки данных, включая обычные ЭВМ общего назначения, эксплуатация в широких масштабах дешевых параллельных компьютеров и их интеграция с картриджами и дисковой памятью в масштабах, необходимых для физики высоких энергий;

— становление персональных рабочих станций как основного средства для интерактивного анализа данных. Развитие распределенной совместной обработки для интерактивного анализа данных, когда различные компоненты программ одновременно выполняются на мощных рабочих станциях и одной или нескольких центральных ЭВМ и файловых станциях;

— введение высокопроизводительных территориальных сетей передачи данных, которые позволят физикам эффективно работать на своих рабочих местах.

В качестве ключевого элемента компьютерного плана CERN было определено улучшение на порядок коммуникационных возможностей

между CERN и сотрудничающими с ним институтами для обеспечения децентрализованной обработки и анализа данных.

На основании проведенных оценок были даны следующие рекомендации для обеспечения потребностей CERN в устройствах массовой памяти на 1991/1992 гг.

— Обеспечить достаточное количество картриджей и дискового пространства для обработки не менее 50 Тбайт данных (это ~ 250 тыс. картриджей).

— Снабдить CERN автоматическими загрузчиками картриджей, способными обрабатывать 8 Тбайт (40 тыс. картриджей).

— Обеспечить копирование картриджей на ленты и их распространение между участниками коллаборации.

— Обеспечить ~ 1 Тбайт дискового пространства в CERN для организации быстрого доступа к наиболее активной части данных из центрального хранилища.

В связи с вышеизложенным встал вопрос о создании в CERN общей вычислительной инфраструктуры для обеспечения работы распределенного вычислительного комплекса, охватывающего различные типы ЭВМ, объединяемых сетями передачи данных.

Для того чтобы внешние центры могли активно участвовать в анализе экспериментальных данных экспериментов на LEP, к их вычислительным ресурсам также были предъявлены соответствующие требования.

Вычислительная мощность — 20 единиц.

Размер дискового пространства — 100 Гбайт.

Автоматический загрузчик на 5 тысяч носителей.

Поскольку проведенный анализ потребностей в вычислительных ресурсах экспериментов на LEP показал, что они значительно превышают возможности вычислительного центра CERN, физики стали искать возможности наращивания вычислительных мощностей за счет создания специализированных вычислительных систем. Это, в конечном итоге, привело к созданию для крупных экспериментов вычислительных систем, обеспечивающих реконструкцию событий и организацию физического анализа [14,17,143].

Таким образом, в развитии вычислительных систем для поддержки электронных экспериментов в 1988—1989 гг. наметились две тенденции. С одной стороны, это наращивание мощностей больших ЭВМ в основном центре и центрах участников коллаборации, а с другой — создание параллельных вычислительных систем, вписывающихся в вычислительные инфраструктуры основных центров.

**2.3. Объемы данных, регистрируемых детекторами.** Большие детекторы, с помощью которых проводятся измерения в интервале энергий

Таблица 3. Основные компоненты детектора UA1 и размер события [26]

Компоненты детектора UA1	Число каналов	Данные, кбайт	
		«сырые»	форматированные
Центральный детектор — дрейфовая камера	6200	1600	~80
Адронный калориметр	1200	2,4	2,4
Электромагнитный калориметр	2200	4,4	4,4
Калориметрический детектор положения	4000	8	8
Передняя камера	2000	32	8
Дрейфовая (мюонная) камера	6000	1	1
Стримерные трубки	50000	50	~4
Урановый калориметр	20000	80	8
Итого		~1770	~140

Таблица 4. Данные о размерах событий для различных экспериментальных условий [27]

Ускоритель	Светимость, $1/\text{см} \cdot \text{с}$	Сечение	Среднее число каналов детектора	Частота	Размер события	Скорость взаимодействий
LEP	$10^{31}$	32 нб	150000	45 кГц	50—100 кбайт	1000 Гц
HERA	$2 \cdot 10^{31}$	15 мкб	260000	10 МГц	125 кбайт	10000 Гц
LHC	$10^{34}$	60 мб	$10^7$	65 МГц	1—10 Мбайт	65 МГц

100 + 1000 ГэВ в с.ц.м. имеют ~ 200000 чувствительных элементов, что позволяет им легко различать энергетические вклады соседних частиц, образующих струи и ливни. Считывание с этих элементов производится каждый раз, когда триггерная система установки обнаруживает интересное событие. При этом один чувствительный элемент может выдавать от одного до нескольких сотен байт информации. Таким образом, «сырые» данные с больших детекторов содержат свыше 1 Мбайт на одно событие. Однако, вследствие того, что не во все чувствительные элементы попадают образующиеся в результате столкновения частицы, и выход с элементов, дающих нулевой результат, отбрасывается, объем регистрируемых на одно событие данных уменьшается до 100+200 кбайт.



В табл.3 приведены данные, показывающие результаты сжатия информации для установки UA1 с 4л-геометрией.

Будущие детекторы для нового поколения ускорителей на более высокие энергии будут иметь большее количество чувствительных элементов, и ожидается, что в этом случае размеры события после первичного сжатия составят ~ 1 Мбайт.

В табл. 4 приведены данные, показывающие длины событий для различных экспериментальных условий [27].

**2.4. Единицы измерения производительности ЭВМ.** До недавнего времени в физике высоких энергий, а также вообще в вычислительной технике, применялась исторически сложившаяся тенденция оценивать быстродействие (производительность, мощность) ЭВМ в миллионах инструкций, выполняемых процессорами в секунду (MIPS). Однако в настоящее время из-за большого разнообразия наборов команд различных процессоров, существенно отличающихся по архитектуре, этот показатель перестал быть объективным критерием. Поэтому для объективной оценки производительности ЭВМ лучше всего использовать набор вычислений, характерных для типичных задач конкретных приложений. Поэтому оценка производительности ЭВМ в настоящее время производится с помощью специальных эталонных тестов (Benchmarks), которые состоят из некоторой смеси операций, определенной для заданного теста и характерной для некоторого класса задач [29—31].

В настоящее время, наряду с MIPS, используются также единицы HER-MIPS (CERN), SSCUPS (SSCL), VUPS (VAX units of performance) и другие единицы [16,32].

При сравнении производительностей ЭВМ, выраженных в различных единицах, можно использовать следующее приближенное соотношение [32]:

$$1 \text{ SSCUP} \approx 1 \text{ SPECmark} \approx 0,25 \text{ CERN unit} \approx 1,4 \text{ MIPS.}$$

Примеры производительности компьютеров для различных тестов приведены также в табл.5 и 6.

**2.5. Затраты вычислительной мощности на реконструкцию и моделирование событий.** Реконструкция типичного события, образуемого при столкновении частиц с энергией 100 + 200 ГэВ, требует от 50 до 150 с на ЭВМ с производительностью 1 HER-MIPS и производится в настоящее время в большинстве случаев в течение нескольких часов после его регистрации. В ходе последующего анализа уточняются значения калибровочных констант, совершенствуются алгоритмы и программы. Это может привести к повторной реконструкции всех или значительной части зарегистрированных событий, что, как правило, и делается [16].

**Таблица 5. Примеры относительной производительности компьютеров для типичных программ в единицах НЕР-MIPS и MIPS [16]**

ЭВМ	Производительность	
	MIPS	НЕР-MIPS
IBM 3090-600E (1 процессор)	6,5	25
Cray XMP-48 (1 процессор)	8,0	30
Apollo DN 10000 (1 процессор)		15—25
HP 9000/720		24—35
IBM RS 6000-320		10—13
VAX 11/780*	1,0	1

\*С ускорителем команд с плавающей запятой.

**Таблица 6. Производительности ЭВМ для различных эталонных тестов [32]**

ЭВМ	SSCUPs	SSCUPs (оптимизир.)	SPECmarks	MIPS	CERN
VAX-11/780			1,0	1,0	
VAX 6410	7,0	8,4		7,5	1,9
SGI 4D/35S	15,3		23,0	33,0	
SGI 4D/310	13,5	21,6	18,5	30,0	5,3
IBM RS/520	11,9		22,0	29,5	
IBM RS/530	14,3	21,3	28,6	33,0	4,9
SUN Sparc 2	11,8	16,3	21,0	28,5	3,5
DEC Stat. 3100	6,4	10,0	10,8	14,0	2,8
DEC Stat. 5000	10,5		18,5	24,0	4,4
Apollo DN 10000	11,6	18,6	17,4	22,0	4,9
HP Apollo 9000/720	22,7	29,4	55,7	57,0	

Моделирование современных экспериментов требует очень больших вычислительных ресурсов, которые за последние 15 лет возросли на 3—4 порядка. С одной стороны, этот рост обусловлен проведением очень точных поисковых экспериментов, требующих моделирования больших объемов сложных событий; а с другой — резким снижением стоимости вычислительных ресурсов, что позволило ставить перед моделированием новые задачи, которые раньше было невозможно решать. Основное время ЭВМ тратится на моделирование ливней, образующихся при прохож-

Таблица 7. Объемы данных и потребности в вычислительной мощности для настоящих и будущих больших экспериментов в физике высоких энергий

Частицы	Настоящие			Будущие	
	$e^+e^-$	$e^-p$	$pp$	$pp$	$e^+e^-$
Энергия в системе центра масс, ГэВ	100	300	2000	(16—40) К	(0,5—2) К
«Сырая» частота интересных событий, Гц	$\leq 1$	1—1000	$1-10^6$	$1-10^8$	$\leq 0,001$
Частота записи на «ленту», Гц	$\sim 1$	$\sim 1$	$\sim 1$	10—10000	$\sim 1$
Размер события, кбайт	100—200	$\sim 100$	$\sim 100$	(1—10) К	(1—10) К
Время реконструкции события, HER-MIPS секунды	50—100	$\sim 100$	$\sim 200$	500—5000	500—5000
Время детального моделирования события, HER-MIPS секунды	(2—5) К	(2—5) К	$\sim 5$ К	(20—200) К	(20—100) К
Годовой объем данных, Тбайт	1—10	1—10	$\sim 10$	$\geq 1000$	$\sim 10$
CPU ресурс, HER-MIPS	50—150	30—100	30—100	$\geq 100$ К	$\sim 200$
Реконструкция и анализ Моделирование	200—2000	100—1000	50—500	$\geq 100$ К	$\sim 200$

дении частиц через установку. Если при этом требуется высокая точность, то затраты времени становятся примерно пропорциональными полной энергии, оставляемой в детекторе. Так, например, при 40 ГэВ оно может составить сотни часов на ЭВМ с производительностью 1 HER-MIPS на моделирование одного события [16]. Что касается стоимости вычислительных мощностей, то из данных, приведенных в [2, 20, 33], стоимость 1 MIPS большой ЭВМ (типа IBM) с 1987 по 1991 г. снизилась почти в четыре раза и в 1991 г. составила 40 тыс. долларов. В то же время стоимость 1 MIPS для рабочих станций была в десять и более раз ниже [33].

В табл. 7 приведены данные, показывающие годовые объемы информации и потребности в вычислительной мощности для текущих и планируемых экспериментов в физике высоких энергий, исходя из частоты регистрации событий, времени работы установки на ускорителе и характеристик изучаемых событий. При вычислениях предполагалось, что время работы различных установок в пучках примерно одинаково и составляет  $\sim 2000$  часов в год. Эта цифра получена на основании следующих данных. Обычно на физические исследования на ускорителях по расписанию отводится половина года. В течение этого времени ус-

коритель работает по 24 часа в сутки, но, по мнению физиков, эффективность работы ускорителя и установки не более 50% [16].

### 3. КОМПЬЮТЕРНОЕ ОКРУЖЕНИЕ ЭКСПЕРИМЕНТОВ В ФИЗИКЕ ВЫСОКИХ ЭНЕРГИЙ

Вследствие большого разнообразия требований к вычислительным ресурсам, время, которое тратится на обработку экспериментальных данных и получение физических результатов, в значительной степени определяется их «компьютерным окружением» («средой»).

3.1. Принципиальная схема компьютерного окружения. Принципиальная схема такой среды, реализованной для L3-эксперимента (LEP, CERN) в 1989—1992 гг., приведена на рис.5. В нее входят большие ЭВМ, системы специальных процессоров, предназначенные для решения различных задач, рабочие станции и большой набор периферийных устройств для регистрации, хранения и выборки данных. «Суперминикомпьютеры» (в основном VAX) применяются для сбора данных, калибровки детекторов и запуска небольших тестовых задач. Основ-

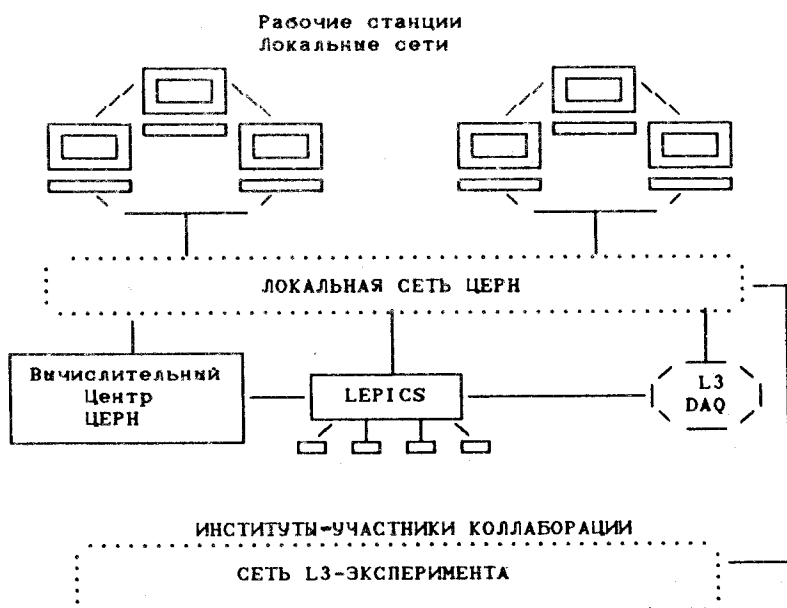


Рис. 5. Принципиальная схема компьютерной среды L3-эксперимента

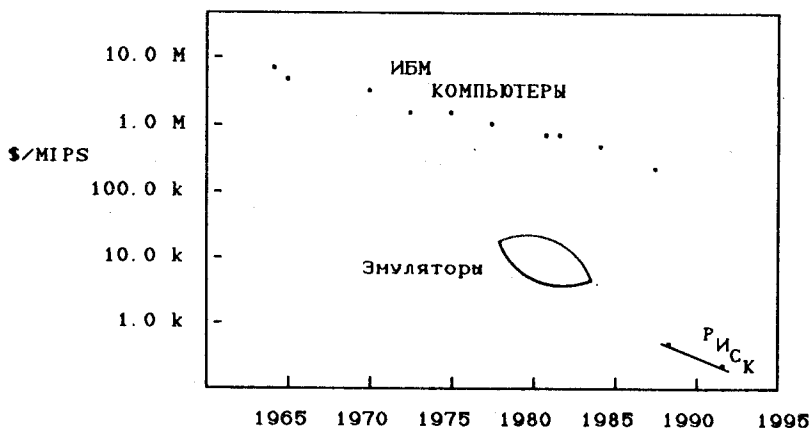


Рис. 6. Падение со временем стоимости различных вычислительных систем (ЭВМ фирмы IBM, эмуляторов и рабочих станций RISC)

ное развитие программного обеспечения и физический анализ получаемых результатов производится на персональных рабочих станциях.

Элементы компьютерной среды экспериментов основной лаборатории соединяются между собой локальными сетями передачи данных, а с вычислительными системами участников коллаборации — посредством территориальных или глобальных сетей передачи данных.

В 70-х и в начале 80-х годов большие лаборатории постоянно приобретали доступные и наиболее мощные ЭВМ, поскольку никаких других возможностей для наращивания вычислительных мощностей тогда не было.

Примерно в 1976 г. несколько небольших групп физиков поняли, что хотя большие ЭВМ и очень хорошо подходят для вычислений с большими объемами ввода/вывода, но имеются задачи, которые могут решаться быстрее с помощью более дешевых вычислительных средств (эмуляторов) [118—122]. Хотя их использование и создавало определенные неудобства, они, тем не менее, нашли применение в ряде экспериментов [16].

К 1985 г. физикам стало ясно, что нужно искать пути получения относительно дешевых вычислительных ресурсов, чтобы избежать в будущем значительных затруднений, связанных с обработкой больших массивов данных. Появление на компьютерном рынке RISC-компьютеров (Reduced Instruction Set Computers) со значительно лучшим соотношением стоимость/производительность, чем у больших ЭВМ, обеспечило решение компьютерной проблемы за счет создания на их основе мощных вычислительных систем (рис.6).

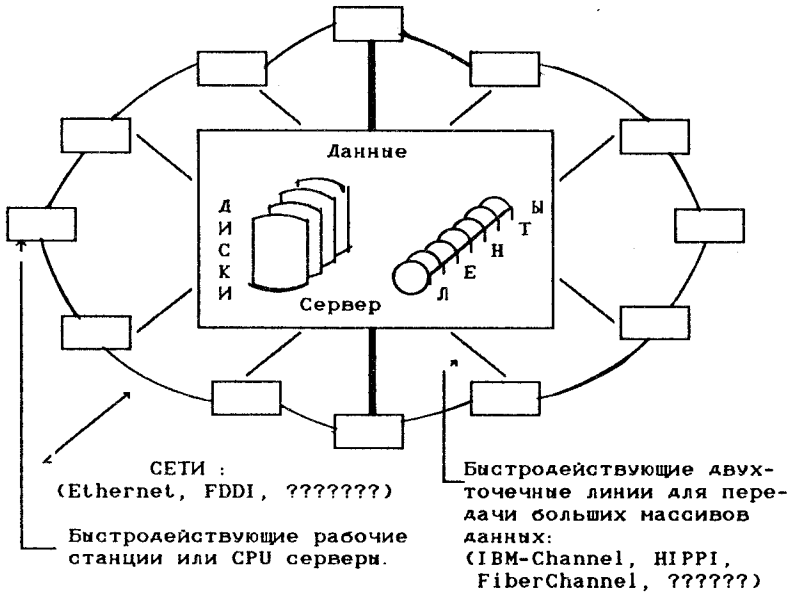


Рис. 7. Модель компьютеринга. Показаны сервер данных и мощные рабочие станции

Работы по созданию такого рода систем были начаты в FNAL в то время, когда большие ЭВМ все еще преобладали [126,132].

Вторая генерация систем использовала семейство процессорных чипов R3000 корпорации MIPS. Каждый процессор мог иметь от 8 до 32 Мбайт памяти и мощность 10—12 МЕР-MIPS. Было построено свыше 200 АСР/R3000 процессоров, и созданные на их основе системы стали широко применяться в FNAL и ряде университетов для обработки данных начиная с 1991 г. [16].

Создание распределенных вычислительных систем на базе недорогих процессоров существенно изменило компьютерную среду экспериментов и, в первую очередь, саму модель ФВЭ-компьютинга, которая была предложена в 1987 г. [4].

### 3.2. Примеры моделей для компьютеринга в физике высоких энергий.

Создание относительно недорогих вычислительных систем на базе микрокомпьютеров и их подключение с помощью локальных сетей передачи данных к вычислительным системам и мощным ЭВМ центральных вычислительных комплексов лабораторий привели к созданию новой модели компьютеринга для экспериментов, принципиальная схема которой приведена на рис.7. Ее центром является мощный сервер данных с

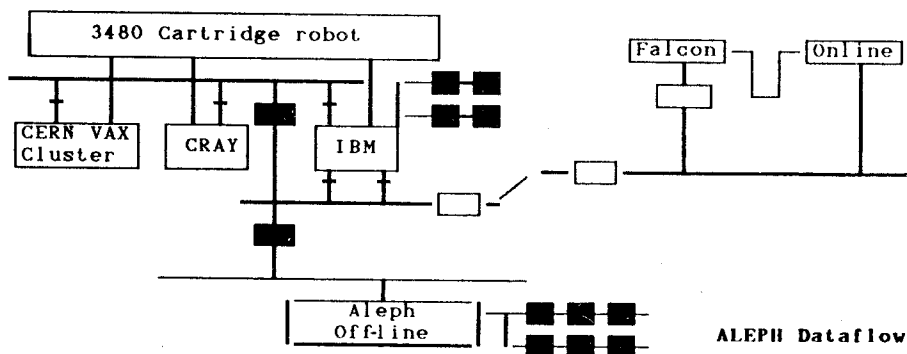


Рис. 8. Схема движения данных, используемых системами реконструкции и off-line обработки детектора ALEPH

большим объемом дисковой памяти и большой библиотекой носителей информации, окруженный сетью мощных рабочих станций [130].

Эта модель является идеализированной и упрощенной схемой, построенной на основе вычислительных систем нескольких экспериментов. Поэтому рассмотрим несколько примеров ее конкретной реализации.

Так, например, LEPICS (L3 Parallel Integrated Computer System) включала IBM 3090, соединенную с рядом HP/Apollo DN 1000 RISC рабочих станций при помощи интерфейса Stollmann GmbH. Присоединенные процессоры могли параллельно решать задачи реконструкции, моделирования и анализа. Доступ к тысячам лент, монтируемых роботом, и десяткам тысяч лент, монтируемых операторами, обеспечивался IBM 3090. Общая сеть обеспечивалась подсетями FDDI, ETHERNET и Apollo Token Ring (12 Мбит/с), а высокоскоростные (1—2 Мбайт/с) связи — интерфейсом Stollmann, соединяющим DN 10000 с IBM 3090. Полная мощность LEPICS — около 500 HEP-MIPS [16].

На рис.8 показана схема движения данных, используемая системами реконструкции и off-line обработки с детектора ALEPH (LEP, CERN).

Система сбора данных (VAXcluster) пересылает файлы исходных данных через двухпортовые диски в систему реконструкции событий Falcon. В 1989—1990 гг. система базировалась на 12 процессорах VAX-станций 3100/30 без дисков и экранов, и ее производительность составляла 24 HEP-MIPS, а в 1992 г. была поднята до 90 HEP-MIPS за счет использования более современных VAX-станций [13].

Результаты реконструкции, готовые для физического анализа, передавались через сеть ETHERNET на диски IBM 3090/600 и ALEPH off-line VAXcluster. На IBM производилось сжатие данных, их классификация и

запись на картриджи для распространения среди участников коллаборации. Управление процессом движения данных между различными системами было полностью автоматизировано и позволяло получать результаты реконструкции для физического анализа спустя несколько часов после их регистрации.

Для повторной обработки данных при введении усовершенствований в алгоритмы и константы в качестве сервера данных использовалась ЭВМ CRAY, обеспечивавшая доступ системы Falcon к лентам исходных данных.

Ферма для реконструкции событий в эксперименте DELPHI — DELFARM является смешанной системой, состоящей из 9 VAX-станций и 13 DEC-станций 5000. Первые организуют работу системы и ввод-вывод данных, а вторые обеспечивают вычислительный ресурс. О возможностях системы можно судить по результатам ее работы в 1992 г. В течение этого года было получено 9000000 триггеров, реконструкция которых завершалась в течение дня после сбора данных. Повторная обработка производилась дважды по мере уточнения констант. DST хранилась как в вычислительном центре CERN, так и на специальной DST-ферме. Последняя состоит из шести DEC-станций с 70 гигабайтным дисковым пространством, которого достаточно для хранения всех данных 1992 г. и обеспечения простого и легкого доступа к ним. Когда любая из этих двух ферм становится свободной, то она автоматически переключается на моделирование событий и также автоматически переключается на решение основных задач по мере необходимости. Данные, которые нужны институтам коллаборации, копируются на картриджи 3480 в вычислительном центре или на кассеты Exabyte на DST-ферме [15, 143].

Разработка и применение такого рода систем существенно сократили сроки получения физических результатов. Об этом, в частности, свидетельствует беспрецедентная скорость, с которой были получены физические результаты в экспериментах на LEP [133]. Есть и другие примеры. Так, например, в FNAL в эксперименте E687 было получено более 500 млн. событий, реконструкция которых была закончена через 7 месяцев после окончания последнего сеанса на ускорителе, а вскоре после этого завершён отбор кандидатов в очарованные частицы и формирование DST [134].

В настоящее время распределенные вычислительные системы являются основным поставщиком вычислительных ресурсов в физике высоких энергий и вытесняют большие ЭВМ. Так, например, в 1992 г. около 90% всех требовавшихся экспериментам вычислительных мощностей было обеспечено распределенными вычислительными системами [16].

**3.3. Роль персональных рабочих станций в компьютерной среде.**  
Под рабочей станцией в физике высоких энергий понимается компьютер,



который имеет память и мощность центрального процессора, необходимые для того, чтобы поддерживать развитие программного обеспечения экспериментов и выполнение соответствующих программ [16].

Графические рабочие станции осуществляют основной вклад в развитие программного обеспечения, особенно при использовании для этих целей станций с экранами высокого разрешения, многооконными операционными системами, системами управления файлами и процессами их совместной обработки с помощью локальных сетей и расширенными графическими возможностями. Мощные графические рабочие станции с трехмерной графикой также стали основой для обработки очень сложных событий, представления данных и интерактивного анализа [20].

Когда в 1982 г. началась подготовка экспериментов на LEP, то многих физиков очень беспокоили вопросы точных измерений и программного обеспечения, обусловленные появлением значительно более сложных проблем, с которыми им раньше не приходилось встречаться. К счастью, появление рабочих станций и относительно недорогих по сравнению с большими ЭВМ вычислительных мощностей обеспечило более эффективное, чем в 70-гг. [16,113], решение многих задач проектирования детекторов.

Персональные компьютеры позволили значительно сократить цикл редактирование — трансляция — загрузка — выполнение программы и обеспечили проведение интерактивной отладки, намного превышающей возможности простого терминала. Наиболее важным обстоятельством здесь является то, что персональный компьютер позволяет использовать графику для лучшего понимания работы программы. Каждый программный элемент можно сконструировать таким образом, чтобы он выдавал результаты в подходящем для визуализации виде, позволяя тем самым оценивать получаемый результат не по таблицам чисел, а по изучению его «картины» в детекторе в двух- или трехмерном пространстве. Естественно, что при таком подходе в сочетании с редактированием и выполнением программ в интерактивном режиме создаются качественно новые условия для разработки программного обеспечения экспериментов. Преимущества, которые обеспечивают графические возможности ПК, стали доступны также и потому, что с самого начала прикладные программы были ориентированы на их использование. Графика позволяет понять, как реально работают детекторы и программы.

В 1992 г. модели персональных компьютеров стоимостью около 7 тыс. долларов имели монохромный экран, объем памяти 16 Мбайт, диска 400 Мбайт и мощность 12 МЕР-MIPS. Более мощные машины стоимостью около 300 тыс. долларов имели 1280×1024 цветной экран, трехмерную

графику, диск свыше 1 Гбайт и, по крайней мере, мощность 30 MIPPS.

В будущем персональные компьютеры будут, по-видимому, использоваться в качестве усложненных терминалов, на которых будет производиться редактирование, частичная компиляция и физический анализ, а все задачи, требующие больших мощностей, будут пересылаться на более мощную машину. В этом случае покупка одной мощной ПК и нескольких X-терминалов сможет обеспечить участие небольших групп физиков в экспериментах [16].

X-терминалы — это нечто среднее между высокопроизводительными рабочими станциями и дешевыми терминалами. Они работают с графическим протоколом, который позволяет разделить вычислительные функции, выполняемые на головном компьютере, и функции отображения информации, реализуемые X-терминалом. X-терминалы, выпускаемые фирмой IBM, называются X-станциями.

X-станции представляют собой устройства, входящие в состав локальной сети, которые обеспечивают функциональные возможности и производительность полноценных рабочих станций, но стоят при этом, в расчете на одного пользователя, гораздо дешевле [136].

3.4. Сети передачи данных. По размерам охватываемых географических площадей сети делятся на глобальные (или территориальные), региональные и локальные [114].

Глобальные сети связывают абонентов, расположенных в пределах страны, континента или всего земного шара. В них широко используется спутниковая связь. Примерами таких сетей являются международные сети национальных авиакомпаний, вычислительные сети, соединяющие центры, ведущие фундаментальные исследования в определенных областях науки, например, в физике высоких энергий, сети, объединяющие академические центры и т.п.

Региональные сети соединяют вычислительные системы в пределах отдельных регионов (страна, область, группы учреждений и т.п.).

Локальные вычислительные сети (ЛВС) охватывают ЭВМ, расположенные на площадях, ограничивающихся одним институтом, предприятием или даже одним зданием.

В физике высоких энергий локальные вычислительные сети служат для соединения устройств как в рамках одного эксперимента, так и в масштабе всей лаборатории.

Локальные или глобальные сети передачи данных являются основными коммуникационными средствами экспериментов [115,123].

Принципиальная схема сетей передачи данных типичного эксперимента приведена на рис.9. Элементы вычислительной среды эксперимента соединяются между собой и с центральным вычислительным центром с

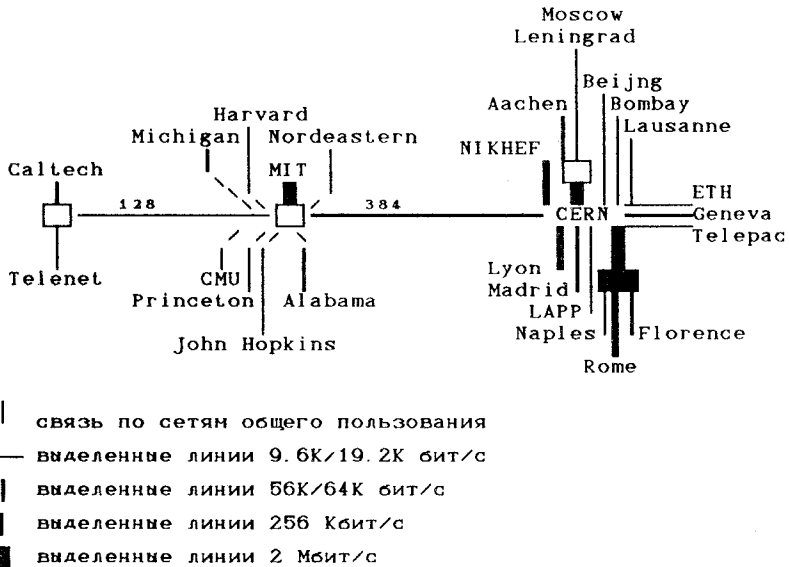


Рис. 9. Топология сети L3-эксперимента

помощью локальных сетей передачи данных, как правило, с высокими скоростями передачи информации (от 10 кбит/с до 100 Мбит/с). Если в эксперименте используется несколько локальных сетей, то они, как правило, соединены друг с другом и имеют выход на глобальные сети.

Сети передачи данных предоставляют физикам следующие возможности:

- 1) доступ с терминалов к удаленным вычислительным ресурсам, имеющимся как в месте проведения экспериментов, так и в организациях, участвующих в коллаборации;
- 2) связь посредством электронной почты с коллегами во всем мире;
- 3) обмен файлами между двумя любыми компьютерами участников коллаборации для:
  - совместной разработки проектов развития программного обеспечения и анализа данных;
  - доступа к удаленным базам данных и анализирующим программам;
  - доступа к удаленным библиотекам программ;
  - распространения небольших массивов событий для их анализа в удаленных лабораториях;
- 4) запуск задач на удаленных ЭВМ;

5) автоматическое сопровождение библиотек программ и баз данных среди участников коллаборации;

6) обмен информацией между задачами, считающимися на разных ЭВМ для:

— управления ходом экспериментов из удаленных лабораторий;

— распределенной обработки данных для эффективного использования доступных вычислительных ресурсов;

7) передача больших объемов экспериментальной информации;

8) организация телеконференций [20].

В последнее время большой интерес проявляется к волоконно-оптическим системам передачи данных в стандарте FDDI (Fibre Distributed Data Interface) со скоростью 100 Мбит/с. Основные области применения системы FDDI в CERN: использование в качестве быстрой центральной сети для соединения локальных сетей ETHERNET и обеспечение связи с высокой пропускной способностью рабочих станций с мощными ЭВМ.

Большинство рабочих станций и персональных ЭВМ еще долгое время будут, по-видимому, базироваться на сети ETHERNET, но будущие мощные рабочие станции, связанные с ними файл-серверы, а также высокоскоростной доступ к центральным ресурсам, будут ориентироваться на FDDI [112].

Для ликвидации разрыва в скоростях передачи данных между участниками работ в исследованиях по физике высоких энергий в Европе начата реализация проекта CHEOPS, предложенного ESA (European Space Agency). В этом проекте предлагается использовать возможности спутника «Olympus». В течение нескольких ночных часов его каналы со скоростью до 8 Мбит/с могут быть использованы для передачи данных между партнерами по коллаборации. Реализация проекта позволит связать CERN и институты в странах Европы, для которых возможностей сухопутных связей или нет, или их стоимость слишком высока (Греция, Финляндия и Португалия) [116,135].

Широкое распространение в Европе и США получили сети EARN и BITNET. Они соединяют академические и научно-исследовательские организации выделенными телефонными линиями и позволяют обмениваться файлами и электронной почтой между ЭВМ IBM, VAX и некоторыми другими.

Комментируя сложившуюся в физике высоких энергий ситуацию с сетями, Ф.Флюкигер в 1991 г. отметил следующее.

1. Наступила эра мегабитных скоростей передачи данных. В Европе такого рода сети объединяют исследовательские институты Франции, Германии, Швейцарии и CERN.

2. Пропускная способность трансатлантических линий тоже поднялась до мегабитных скоростей.

3. Институты Центральной и Восточной Европы также подключены к сети EPNET, правда, с меньшими скоростями [116]. Исследования показывают, что участникам экспериментов требуются:

— интегрированные локальные и глобальные вычислительные системы;

— хороший доступ к данным объемом от 1 до 40 Тбайт;

— вычислительные мощности 1000 MIPS и более с хорошим доступом к данным;

— интерфейс пользователя и графические возможности на уровне, предлагаемом современными рабочими станциями.

Для обеспечения оперативной связи между участниками экспериментов также создаются системы для проведения видеотелеконференций [117].

Следует отметить, что в связи с высокой стоимостью сетевых услуг вопрос о передаче больших объемов информации на магнитных носителях не потерял актуальности.

3.5. Примеры компьютерных сред в центрах физики высоких энергий. Ведущим международным центром фундаментальных исследований в различных областях физики, включая физику высоких энергий, является CERN. В последние годы для развития его компьютерной среды были характерны следующие тенденции.

1. Нарастивание вычислительных мощностей за счет больших ЭВМ, которое продолжалось до 1992 г., в соответствии с разработанными ранее планами обеспечения LEP-экспериментов.

2. Создание распределенных вычислительных систем для удовлетворения компьютерных потребностей LEP-экспериментов.

3. Начало постепенной замены больших ЭВМ распределенными вычислительными системами.

В 1989 г. центральная вычислительная мощность CERN составляла ~88,5 «черновских» единиц за счет следующих ЭВМ: CRAY X/MP-48 (32), IBM 3090-600E (39), Siemens 7890 (13) и VAX Cluster (~4,5).

Примечание. Одна «черновская» единица вычислительной мощности — это мощность IBM 370/168 или VAX 8600.

Основные пакетные и интерактивные запросы обеспечивались IBM 3090/600E и Siemens 7890. Эти машины имели доступ к 25 лентопротяжным и 32 кассетным устройствам и дисковому пространству (205 Гбайт).

Основной интерактивный сервис обеспечивался VAX-кластером.

Локальная сеть X.25 была подключена к 12 внешним выделенным каналам, включая один спутниковый и один канал связи с США по дну

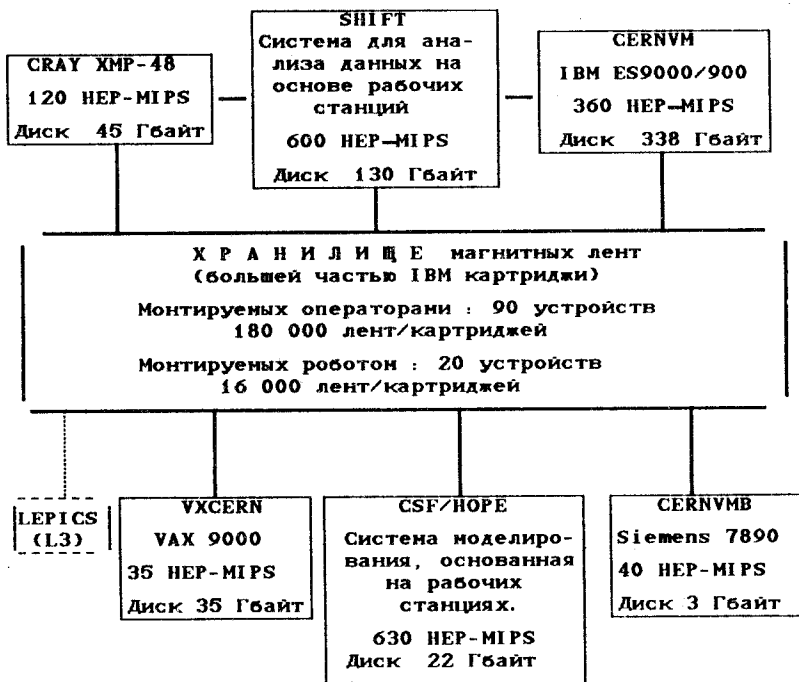


Рис. 10. Упрощенная схема вычислительного центра CERN, показывающая его основные возможности на начало 1992 г. (CERNVMB имеет доступ к CERNVM дискам)

океана. Швейцарский узел EARN был размещен в CERN и также подключен к 12 выделенным каналам.

Большинство больших экспериментов имело по крайней мере один мощный 32-битный компьютер. Обычно это была большая ЭВМ VAX, а также ряд небольших машин, обычно объединенных в кластер. Эти ЭВМ использовались для решения многих задач, включая сбор данных во время испытаний детекторов и сеансов калибровки. Имелось по меньшей мере 100 машин VAX и микроVAX и 20 компьютеров Nord Data.

В 1989 г. по оценкам в CERN имелось по крайней мере 2000 ПК фирмы Apple на базе микропроцессора Motorola 68000 (Macintosh) и не менее 800 совместимых с IBM ПК, более 120 рабочих станций Apollo и по меньшей мере 250 VAX-станций.

К 1992 г. суммарная вычислительная мощность ВЦ CERN составила около 2000 НЕР-МІРС, из которых ~ 60—70% приходилось на долю распределенных вычислительных систем (рис.10). Суммарный объем диско-

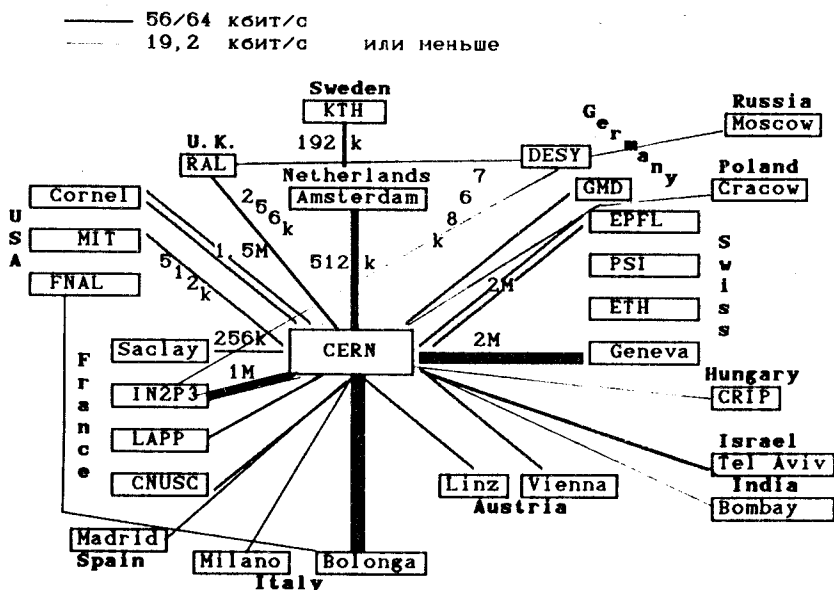


Рис. 11. Международные линии связи CERN с институтами физики высоких энергий

вой памяти превысил 550 Гбайт. В хранилище магнитных лент имелось около 200 тыс. магнитных лент и картриджей. На 3000 персонала и 4000 визитеров имелось более 5000 различных персональных компьютеров и рабочих станций [16].

Таким образом, основными компонентами вычислительного центра CERN в 1992 г. являлись центральная библиотека магнитных лент и картриджей, включая автоматические загрузчики, большие ЭВМ общего назначения и основанные на рабочих станциях системы, роль которых все более и более возрастала [125].

Наряду с развитием центрального вычислительного комплекса CERN, шло оснащение больших электронных экспериментов собственными вычислительными системами, примеры которых были приведены выше.

В 1992 г. специалистам CERN стало ясно, что эпохе больших и супер-ЭВМ в физике высоких энергий приходит конец и нужно искать варианты для их замены [143,145].

В апреле 1992 г. в CERN было объявлено о планах вывода из эксплуатации системы CRAY X-MP/48 и замены CERN VAX на более мощные машины архитектуры RISC, названные DXCERN [144].

В марте 1993 г. система CRAY X-MP/48 была заменена кластерами на базе машин IBM RS6000 и HP9000/700 [144].

На рис.11 показаны международные линии связи, соединяющие CERN с участниками мирового сообщества физиков по состоянию на начало 1992 г.

В качестве примера вычислительной среды национального центра рассмотрим национальную лабораторию Японии КЕК [131,137,146].

Исследования по физике высоких энергий в Японии проводятся на  $e^+e^-$ -коллайдере TRISTAN и протонном синхротроне (12 ГэВ/с) в основном двумя национальными и одной международной коллаборацией.

Университеты и институты Японии, включая КЕК, принимают участие в экспериментах, проводимых в других центрах: CDF (FNAL), OPAL (LEP), ZEUS (HERA), SLD (SLAC) и др. Всего в Японии в области физики высоких энергий работают более 45 институтов. Большое количество данных, получаемых в этих экспериментах, обрабатывается в вычислительном центре КЕК.

КЕК располагает основными вычислительными ресурсами в Японии, обеспечивая потребности как самой лаборатории, так и других организаций, участвующих в экспериментах в физике высоких энергий, которые имеют к ним доступ через компьютерную сеть, объединяющую все компьютеры Японии.

Основные характеристики центральной компьютерной системы КЕК приведены в табл.8.

Таблица 8. Центральная компьютерная система КЕК

Характеристики системы	M-880/210 (Hitachi)	M-680H (Hitachi)	S-820/80 (Hitachi)
Производительность CPU	93 MIPS	56 MIPS	3,0 GFLOPS
Основная память	256 Мбайт	28 Мбайт	512 Мбайт
Внешняя память	1,5 Гбайт		2 Гбайт
Операционная система	VOS3/AS	VOS3/AS, HI-OSF	VOS3/AS
Файлы (для трех компьютеров)			
Магнитный диск для пользователей		180 Гбайт	
Библиотека кассетных лент		2,2 Тбайт	
Лентопротяжные устройства		14	

Центральная вычислительная система КЕК состоит из двух универсальных ЭВМ (НИТАС M-680H и M-280H) и векторного процессора (НИТАС S280/80). Максимальная производительность M-680H —



31 MIPS, M-280H — 14 MIPS. Пиковая скорость S820/80 — 3 GFLOPS. С июня 1992 г. эти три компьютера составляют слабо связанную мультипроцессорную (LCMP) систему, обеспечивающую им возможность совместно использовать большое дисковое пространство.

Вычислительная система TRISTAN предназначена для обеспечения проводимых на коллайдере экспериментов. Она состоит из двух ЭВМ: одна — для работы в режиме разделения времени (FACOM M-780/10S), другая — для работы в пакетном режиме (FACOM M780/30). Характеристики этих систем приведены в табл.9 [146].

Таблица 9. Компьютерная система TRISTAN

Характеристики системы	Batch-система M1800/30 (Fujitsu)	TSS-система M 1800/10s (Fujitsu)
Производительность CPU (предполагаемая)	140 MIPS	33 MIPS
Основная память	128 Мбайт	256 Мбайт
Операционная система	OS IV/F4 MSP	OS IV/F4 MSP UXP/M (UNIX)
Система дисковой памяти	234 Гбайт	
скорость передачи	9 Мбайт/с	
Библиотека VHS лент	3 устройства	
полная емкость	1,8 Тбайт	
скорость передачи	3 Мбайт/с	
Кластер рабочих станций S-4/10 модель 40	11 машин 96 MIPS/машина	

В начале 1993 г. производительность системы была увеличена примерно в 1,7 раза, полная емкость дискового пространства — до 230 Гбайт, а ленточного пространства — до 3,3 Тбайт.

В дополнение к комплексу больших машин имеется кластер из 11 S-4/10 (Fujitsu) SUN-совместимых рабочих станций, производительностью 94 MIPS каждая.

Компьютерная сеть КЕК HEPnet-J (High Energy Physics network — Japan) обеспечивает связь центра с японскими и международными институтами, проводящими исследования в физике высоких энергий и ядерной физике. HEPnet-J соединена через КЕК с Америкой посредством выделенной международной линии, а через Америку с Европой, образуя международную сеть, которая функционирует с июня 1992 г. (рис.12).

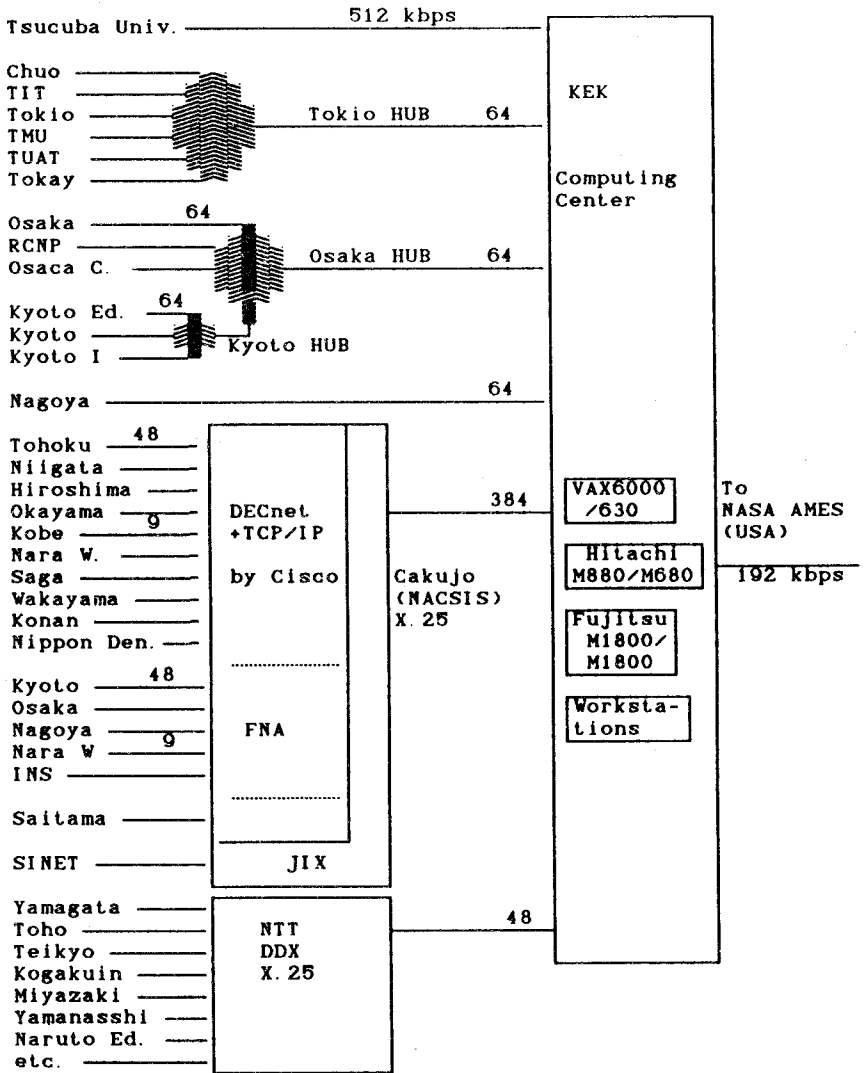


Рис. 12. Сеть, соединяющая институты физики высоких энергий Японии

Изучается возможность организации компьютерной связи с Россией и Китаем.

С декабря 1991 года в КЕК работает система телеконференций, которая часто используется коллаборацией Японии — США и имеет хорошую репутацию. В среднем она используется около 25 часов в месяц.

Коллаборанты в университетах используют свои компьютеры не только для решения своих локальных задач, но и как удаленные узлы вычислительной сети, обеспечивающие доступ к большим ЭВМ центральной вычислительной системы КЕК. Они могут посылать задачи для решения в пакетном режиме и получать результаты через сеть.

Будущая экспериментальная программа исследований в физике высоких энергий ориентируется как на развитие собственной ускорительной базы, так и на участие в международных коллаборациях.

В заключение этого раздела рассмотрим статус и перспективы развития информационно-вычислительной инфраструктуры Объединенного института ядерных исследований.

Программа научной деятельности ОИЯИ охватывает экспериментальные и теоретические исследования в следующих областях:

- физика промежуточных энергий;
- физика тяжелых ионов;
- нейтронная ядерная физика;
- физика элементарных частиц;
- релятивистская ядерная физика;
- физика конденсированных сред.

Объединенный институт ядерных исследований играет ключевую роль в координации исследований, проводимых учеными 18 стран-участниц ОИЯИ, и широкого научного сотрудничества институтов этих стран с ведущими физическими центрами мира и международными организациями. Институт также имеет специальные соглашения о проведении совместных работ с CERN и исследовательскими центрами США, Германии, Франции, Италии, Венгрии.

Все эти обстоятельства требуют поддержания информационно-вычислительной инфраструктуры ОИЯИ на соответствующем уровне. О том, насколько вычислительные средства Института удовлетворяют современным требованиям, предъявляемым к международному центру фундаментальных исследований, можно судить по следующим данным, характеризующим его компьютерную среду.

Основу информационно-вычислительной инфраструктуры ОИЯИ составляют Центральный вычислительный комплекс (ЦВК) ОИЯИ, локальные измерительно-вычислительные центры в лабораториях и подразделениях института, средства компьютерной и терминальной связи.

В состав ЦВК ОИЯИ входят (рис.13):

- распределенный UNIX-сервер на базе ЭВМ CONVEX-C120, SUN SPARC станции II и microVAX;
- кластер из трех ЭВМ типа VAX;

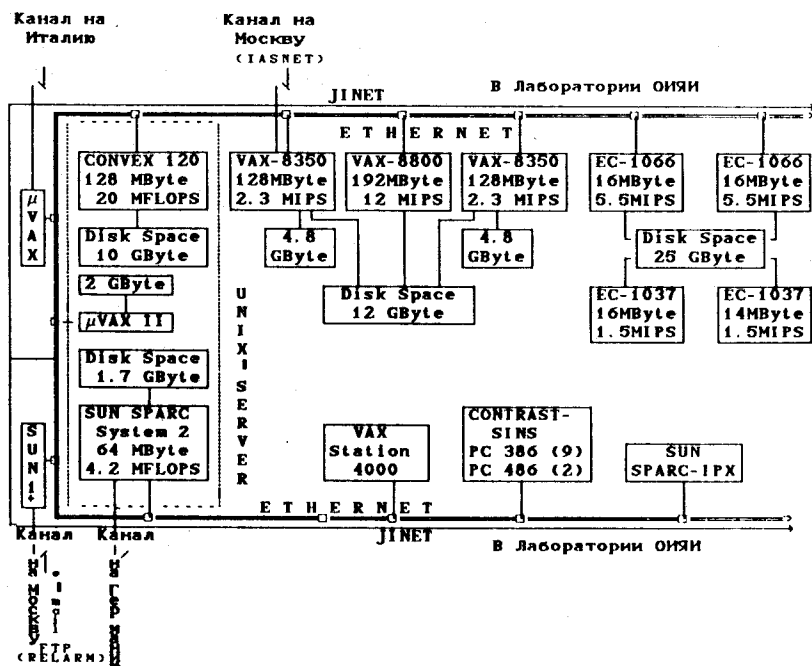


Рис. 13. Центральный вычислительный комплекс ОИЯИ

- многомашинный комплекс ЕС ЭВМ, в составе которого имеется два матричных процессора с производительностью до 20 MIPS на векторных операциях;
- VAX-станция-4000;
- SUN SPARC-IPX-станции (2 шт.);
- комплекс рабочих станций КОНТРАСТ-СИНС.

*Примечание:* В июле 1994 г. производительность ЦВК была существенно увеличена за счет включения в его состав ЭВМ CONVEX 220 с производительностью 100 MFLOPS.

Для архивирования, хранения и обработки больших объемов данных используются магнитные ленты различной плотности (от 800 до 6250 bpi), картриджи, EXABYTE-8500, CD-ROM.

Основная задача UNIX-сервера заключается в обеспечении в среде UNIX пользователей локальной сети ОИЯИ вычислительным, программным, файловым и информационным сервисом. Для ее решения планируется приобретение интегрированной многопроцессорной системы на базе CONVEX-220 с двумя RISC-процессорами HP/735, многопроцессорного сервера SUN SPARC 1000, расширение дискового пространства и создание системы массовой памяти на базе устройств типа EXABYTE-8500.

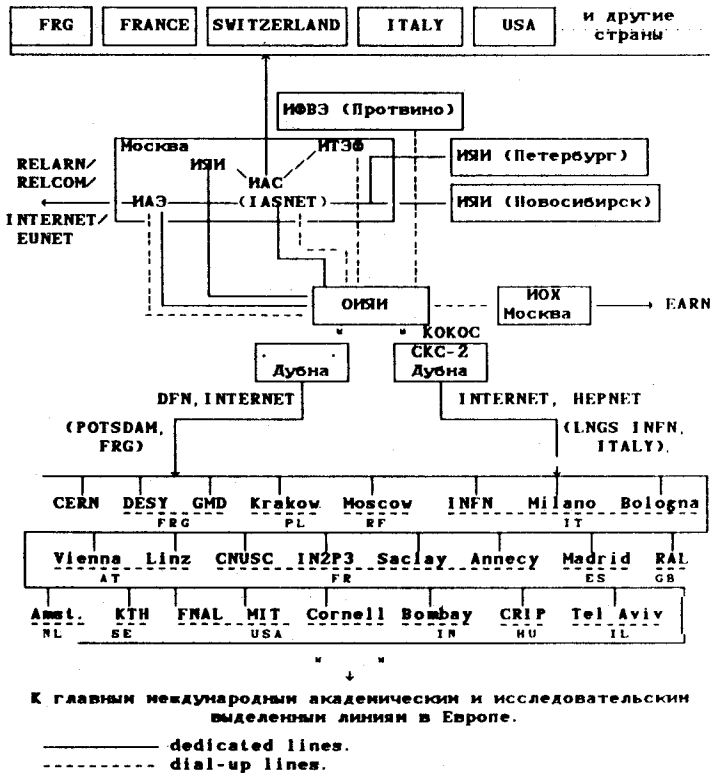


Рис. 14. Схема связей ОИЯИ с международными компьютерными сетями Европы, Америки, а также с ядерными центрами России

Средства вычислительной техники подразделений института (ЭВМ измерительно-вычислительных центров лабораторий, SUN- и VAX-станции, microVAX, персональные ЭВМ класса IBM PC XT/AT/286/386/486) являются абонентами сетей JINET (Joint Institute NETwork) и ETHERNET, что обеспечивает пользователям доступ к ресурсам базовых ЭВМ центрального вычислительного комплекса Института. Кроме того, они могут обмениваться письмами, заказывать препринты в научно-технической библиотеке, читать информационные бюллетени, новости, касающиеся сетей и ЭВМ, объявления о семинарах, приказах дирекции и т.п.

Программное обеспечение комплекса включает сетевые пакеты, традиционные системы программирования, СУБД, библиотеки программ CERNLIB, NAGLIB и другое прикладное программное обеспечение,

используемое в физических центрах для проведения экспериментальных и теоретических исследований, в том числе пакеты программ для анализа и управления данными, моделирования электронных экспериментов, поддержки больших программных комплексов, представления данных и др. На ЭВМ CONVEX создан и поддерживается архив свободного программного обеспечения.

Сотрудники Института имеют доступ к базе данных по физике частиц PPDS, информационной подсистеме INIS и международной системе WWW.

Локальная сетевая инфраструктура ОИЯИ подключена к международным компьютерным сетям через выделенные и коммутируемые наземные и спутниковые каналы связи (рис. 14). Функционируют два выделенных канала компьютерной связи на Москву. Один с пропускной способностью 9600 бит/с, второй — 19,2 кбит/с. Каналы работают по международному протоколу TCP/IP и предоставляют средства удаленного входа в ЭВМ (в том числе и зарубежные), обмена электронной почтой, материалами телеконференций, файлами данных и программами. Эти системы подключения позволяют организовать обмен по электронной почте и сеансы связи пользователей локальных сетей JINET и ETHERNET ОИЯИ с абонентами или ЭВМ в локальных сетях основных физических центров России (ИФВЭ, Протвино; ИТЭФ, ИЯИ, ИАЭ, ФИАН, Москва; ИЯФ, Гатчина; ИЯФ СО РАН, Новосибирск и т.д.) и западных стран (ФРГ, Франции, Италии, Швейцарии, США).

Для обеспечения участников совместных международных экспериментов более эффективной оперативной связью на базе ОИЯИ, станции космической связи в Дубне (СКС-2) и завода «Тензор» реализованы два канала спутниковой связи.

Один обеспечивает прямой выход в германскую компьютерную сеть DFN, второй — в Италию и через них доступ к другим международным компьютерным сетям.

Эти каналы предоставляют следующие возможности:

- оперативный обмен информацией по электронной почте;
- доступ к базам данных;
- использование общей вычислительной мощности ЭВМ в сети;
- передача массивов экспериментальных данных для обработки на местах в off-line режиме;
- передача факсимильной информации, текстов статей, проведения телеконференций.

Скорость передачи данных в сети по этим каналам составляет 64 кбит/с с возможностью дальнейшего расширения.

Эти каналы открывают сотрудникам ОИЯИ доступ в интерактивном режиме к мощным компьютерам и информационным базам данных науч-

ных центров Европы и мира. Рассматривается возможность использования спутникового канала на Потсдам для проведения телеконференций между ОИЯИ и научными центрами Германии.

Таким образом, хотя компьютерная среда ОИЯИ и обеспечивает участие Института в различных международных коллаборациях, ее возможности заметно ниже тех, которыми располагают ведущие ядерные центры. Это, в первую очередь, относится и к имеющемуся на центральных ЭВМ дисковому пространству, и к скоростям передачи данных по имеющимся каналам связи, а также к оснащению рабочих мест современными персональными компьютерами и рабочими станциями. Следует, однако, отметить, что в последние годы в решении этих вопросов, а также в увеличении производительности центрального вычислительного комплекса, происходят существенные сдвиги.

#### 4. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ АНАЛИЗА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Программное обеспечение экспериментов в физике высоких энергий — это большой и сложный комплекс разнообразных программ, пакетов и систем, обеспечивающих, в частности, решение следующих задач в ходе обработки и анализа данных: моделирование; реконструкция событий; организация структур данных; интерактивный анализ и представление данных; организация баз данных и работа с ними; статистический анализ и вычисление характеристик изучаемых физических процессов.

4.1. Некоторые специфические особенности программного обеспечения. Проблемы, возникающие при разработке программного обеспечения для больших детекторов современных и будущих экспериментов в физике высоких энергий, были рассмотрены в докладе Р.Бека [85] на конференции по компьютингу в 1987 г. [3]. В частности, отмечается необходимость более согласованных действий в этой области между «экспериментами» и большими институтами; возможное устранение в программном обеспечении трудно поддающихся контролю ситуаций, благодаря внедрению в практику программирования современных технологий и методов, а также необходимость управления небольшим числом программистских лидеров деятельностью добровольцев-физиков (traditional physicist volunteers) и больших контактов как внутри сообщества физиков, так и с представителями других наук, включая компьютерную, а также промышленности.

Оценивая понятие термина «большой эксперимент» применительно к экспериментам в физике высоких энергий, Р.Бек отмечает, что его можно характеризовать следующими данными:

— Типичная коллаборация насчитывает от 100 до 400 физиков. Половина из них обычно вносит вклад в развитие программного обеспечения, а дюжина или более — в решение ответственных или координационных проблем.

— Число компьютеров в типичных экспериментах  $\geq 100$ .

— Объем программного обеспечения коллаборации — от 100 до 500 тысяч линий. Из них  $2/3$  составляют выполняемые коды,  $1/10$  — декларативные операторы и одна четверть — комментарии. Естественно, что это относится к программам 80-х гг., написанным на Фортране.

— Анализирующие данные программы обычно состоят из нескольких сотен подпрограмм, их которых от 10 до 20% пишутся непосредственно физиками. Остальная часть содержит программные элементы из поддерживающих эксперимент библиотек общего назначения данной коллаборации и даже коммерческих.

Текущее состояние дел в этой области в 1991—1992 гг. было рассмотрено в обзорах Р.Маунта [16] и П.Кунца [38].

Характеризуя состояние дел с программным обеспечением экспериментов в 1992 г., Р.Маунт отмечал, что, несмотря на очень большие объемы программного обеспечения (сотни тысяч и миллионы линий программных кодов), которое приходится писать участникам коллабораций, в них обычно очень мало или даже вообще нет профессиональных программистов.

Большинство физиков, работающих в этой области исследований, являются программистами, и все они отлично сознают, что ни их собственные коды, ни те, что имеются в базовом обеспечении, не могут не содержать ошибок.

Имеется еще одно важное обстоятельство, на которое нельзя не обратить внимание. Если еще в не очень отдаленном прошлом достоинства программиста определялись его способностями включать максимальное число кодов в минимально возможный объем машинной памяти, а «читабельность» текста другими людьми, вообще говоря, не имела особого значения, то в настоящее время ситуация существенно изменилась. И вопрос — «ясность или эффективность», в основном, решается в пользу первой, естественно, при разумном их сочетании. Более того, в большинстве коллабораций делаются формальные или неформальные попытки заставить программистов писать программы так, чтобы в их текстах могли разбираться и другие лица, например, выпускники университетов.

В качестве основного алгоритмического языка используется Фортран, и попытки заменить его до настоящего времени особых успехов не имели, хотя некоторые сдвиги в этом направлении имеются, а альтернативы Фортрану обсуждаются на конференциях по применению компь-



ютинга в физике высоких энергий [34—38]. Фортран является удобным языком программирования для научных вычислений, но, с точки зрения физиков, он имеет ряд недостатков. Поэтому для их устранения были созданы специальные средства, в частности, для управления структурами данных и машинно-независимого ввода/вывода, которые, собственно говоря, в значительной степени и обеспечили его «живучесть» [16].

**4.2. Программные средства.** Программные средства в физике высоких энергий делятся на следующие три категории [16]:

1. Системы, написанные для того, чтобы преодолеть ограничения Фортрана. К ним относятся пакет ZEBRA [39], а также его предшественники HYDRA [40], ZBOOK [41] и BOS [42], ZEBRA (The Data Structure Management System) была разработана в CERN, чтобы преодолеть недостатки Фортрана в организации динамических структур данных.

2. Средства, обеспечивающие машинно-независимые функции, такие, как управление кодами, базами данных и библиотеками программ общего назначения, и являющиеся основой для создания конкретных прикладных программ.

Первым примером средств управления кодами является система PATCHY [43]. Эта ориентированная на перфокарты система более двадцати лет обеспечивала развитие и сопровождение программ и адаптировалась на каждый тип ЭВМ, включая БЭСМ-6 [44].

К настоящему времени для решения этой проблемы создана новая система управления кодами CMZ (Code Management System using ZEBRA). Это интерактивная, быстрая, машинно-независимая и совместимая с PATCHY система управления и сопровождения исходных текстов на языках Фортран-77 и Си. Она предоставляет разработчикам эффективные средства поддержки для создания программного продукта, начиная с разработки и кончая его сопровождением [45].

Другим примером, относящимся к этой категории средств, является система управления базой данных HEPDB, которая предназначена для хранения калибровочных констант и обеспечения системы учета. Она создается на основе двух систем DBL3 и OPCAL [46—48], разработанных для отдельных экспериментов. HEPDB использует возможности прямого доступа системы ZEBRA, обеспечивая машинную независимость и поддержку доступа к данным с помощью средств Фортрана [42, 73—80].

Третьим примером этой категории средств является CERN KERNLIB — компактное ядро библиотеки часто используемых сервисных и вычислительных программ и подпрограмм. Эта библиотека внедрена практически на всех типах компьютеров и является важнейшим элементом библиотеки программ CERN [86]. На ее основе создается большинство пакетов прикладных программ CERN и других лабораторий.

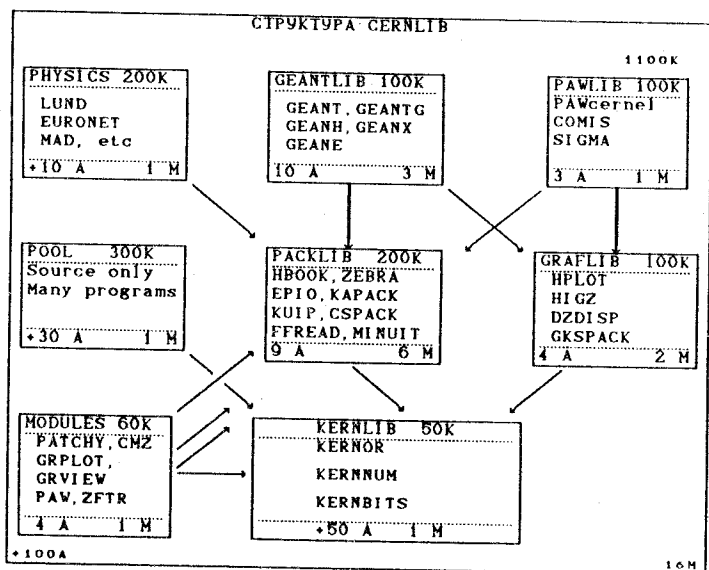


Рис. 15. Структура библиотеки программ CERN (1100 K)

В настоящее время библиотека программ CERN (CERNLIB) распространена среди почти 600 лабораторий и институтов 55 стран мира [86,88,89].

Структура библиотеки показана на рис.15. Ее различные компоненты в сумме составляют более 1 млн. линий с текстами программ, большинство которых написано на Фортране-77. Часть программ в KERNLIB написана на ассемблере (~ 5000 линий). PACKLIB, GRAFLIB и KERNLIB содержат все более возрастающее число подпрограмм, написанных на Си (~ 10000 линий). Оценки числа линий, содержащихся в каждой из компонент библиотеки, приведены в правом верхнем углу соответствующего прямоугольника. Полное число авторов библиотеки — более ста. Внизу слева приведены приближенные оценки числа авторов каждого раздела, а справа — число людей, занимающихся развитием и сопровождением соответствующего раздела.

Для получения документации по программам библиотеки разработана специальная система автоматической генерации программной документации SIM (A Software Information Manager) [87].

4.3. Специфические программные средства. В эту большую категорию включены программы и утилиты, которые специально написаны для многих или отдельных экспериментов.

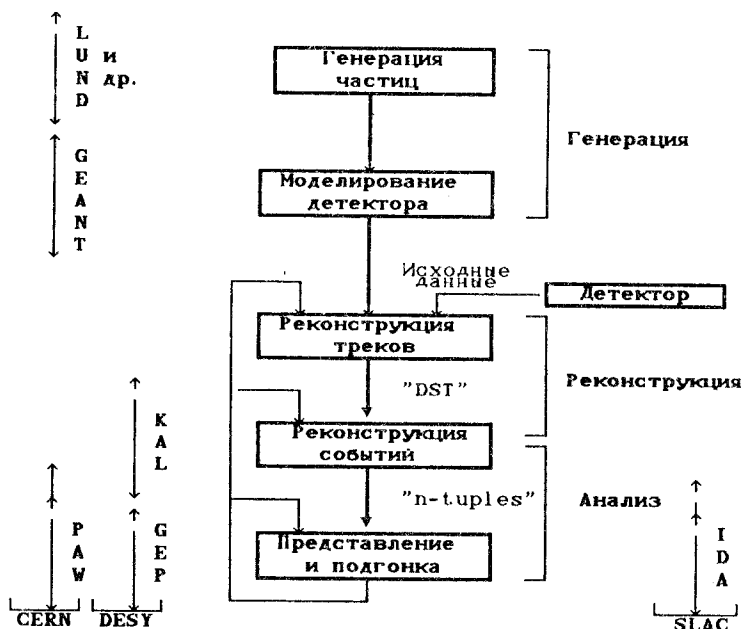


Рис. 16. Схема анализирующей цепочки с программами, которые широко используются на ее различных стадиях

На рис.16 приведена принципиальная схема анализирующей цепочки данных и названия некоторых широко используемых программных средств с указанием стадий, на которых они используются, и лабораторий, в которых они были разработаны.

На стадии генерации частиц моделируется первичное взаимодействие, приводящее к образованию частиц, и распады короткоживущих частиц. Ее результат — 4-векторы долгоживущих частиц. Для этого используется программа LUND [81] или другие программы моделирования подобного типа, которые обеспечивают генерацию частиц.

Многие компоненты для моделирования событий являются общими для некоторых экспериментов. Генераторы событий объединяются в пакеты, которые моделируют известные или гипотетические случаи образования частиц в индивидуальных  $e^+e^-$ ,  $e^-p$ ,  $pp$ -столкновениях.

Примерами широко используемых генераторов событий являются LUND [69], KORALS [49] и KORALB [50], JETSET [51,52], HERWIG [53—57], PYTHIA [58—60], ISAJET [61], MARS [71,72], EGS [64] GHEISHA [65].

На следующей стадии моделируется поведение события в детекторе. Здесь частицы «проводятся» через детектор, генерируются пространственные точки и «отклики» детектора, которые записываются на файл результатов в той же форме, что и для реальных детекторов. Для решения этих задач широко используется пакет GEANT [62].

На стадии реконструкции треков решается задача распознавания зарегистрированных установкой элементов события, восстанавливаются сегменты его треков и производится предварительная идентификация, в ходе которой составляется список вероятностей принадлежности трека к тому или иному типу частиц.

На стадии реконструкции события восстановленным трекам на основе данных об идентификации приписываются значения масс и формируются 4-векторы. Из их комбинаций составляются гипотезы о вершинах и топологии события. Результаты этой стадии выдаются в форме гистограмм и  $n$ -tuple. Под последними понимаются двумерные массивы, характеризуемые двумя числами. Первое определяет число записей на один элемент массива, а второе — полное число элементов. Каждый элемент  $n$ -tuple может рассматриваться как физическое событие, характеризуемое небольшим набором физических величин (записей).

Примером программного средства реконструкции событий является разработанный в DESY пакет KAL [83].

На завершающей стадии (представления и подгонки) собранные на предшествующих стадиях статистические данные представляются в виде гистограмм, диаграмм рассеяния и т.п.

Закрывают цепочку пакеты PAW, GEP и IDA.

PAW — набор пакетов для интерактивного анализа данных, включая генерацию и представление гистограмм и графиков. Он требует от физика написания программы для извлечения из сложной структуры данных, описывающих каждое событие, небольшого числа переменных на событие, которые затем могут быть использованы им для анализа [68].

GEP — интерактивная анализирующая программа для анализа данных в физике высоких энергий, созданная в DESY. Она предоставляет пользователям много возможностей для операций с данными, получения гистограмм, диаграмм рассеяния, графиков, интерактивной подгонки данных и генерации новых гистограмм или диаграмм рассеяния из многомерных распределений [82].

IDA — интерактивная система для анализа данных в физике высоких энергий, которая широко используется в SLAC для анализа DST [84].

Пакет GEANT является в настоящее время компонентой большинства монте-карловских программ для текущих и будущих экспериментов. Подпрограммы GEANT «проводят» частицы через детектор с учетом сечений взаимодействия в чувствительных и нечувствительных элементах

детектора. Затем эти данные преобразуются в отклики элементов детектора [62,63].

Типичная коллаборация включает сотни физиков из десятков институтов. Вычисления для этих экспериментов производятся как в основном центре, так и в институтах — участниках коллабораций, создавая распределенную базу данных, которая требует управления и доступа к большим количествам данных, находящихся на различных устройствах, распространенных по большой географической площади. Для ее решения была разработана специальная система FATMEN (Distributed File and Tape Management System).

FATMEN — это каталог наборов данных, который является новым инструментом, сконструированным для того, чтобы решать проблему управления и адресации к сотням тысяч наборов данных, распределенных (и, вероятно, дублированных) по многим компьютерам. FATMEN обеспечивает пользователям доступ к данным независимо от системы, на которой они работают, или от размещения самих данных. Данные могут храниться на локальной или удаленной машине и находиться на диске, ленте или другом носителе. Во всех случаях интерфейс пользователя идентичный. FATMEN обеспечивает три основные функции: идентификацию, доступ и управление данными [66,67].

**4.4. Будущие программные средства.** Рассматривая возможные пути развития программных средств в физике высоких энергий, П.Кунц [38] отмечает, что никто не может с уверенностью предсказать их будущее, т.к. в сфере компьютерной технологии наблюдается очень быстрое развитие как технических, так и программных средств. Ключевым вопросом, по его мнению, является путь, по которому пойдет развитие программного обеспечения. Эволюционный подход будет означать постепенный переход к Фортрану-90 с его новыми возможностями, которые облегчат задачу написания программ для физического анализа и анализирующих средств [90]. Революционный — изменение языка программирования и использование объектно-ориентированной парадигмы. В том и другом случае подразумевается добавление к имеющимся средствам графического интерфейса пользователя CUI (Graphic User Interface) [138].

Вплоть до недавнего времени почти все средства для физического анализа создавались на основе Фортрана и средств, расширяющих его возможности, о которых уже ранее шла речь.

Многие исследователи, работающие вне физики высоких энергий, нашли, что объектно-ориентированное программирование (ООП) является лучшим способом для создания средств моделирования и анализа. Это есть реальная революция в создании программ. Объектно-ориентиро-

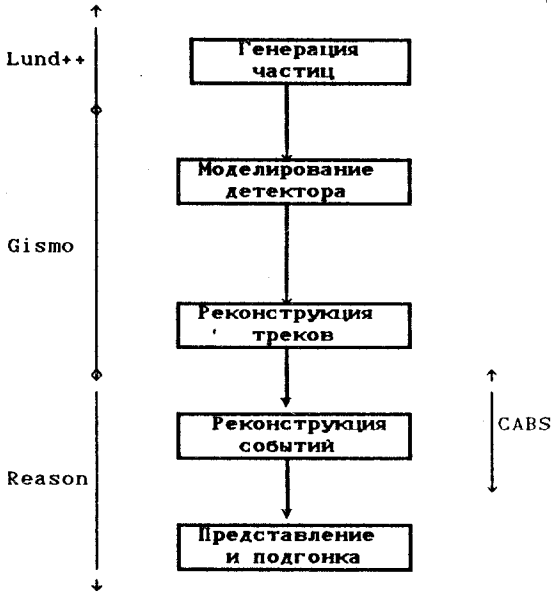


Рис. 17. Новые средства анализа данных в 90-х годах

ванное программирование не связано с определенным языком. Для его реализации используются языки Objective-C, Си++, SmallTalk и Eiffel [16].

В связи с тенденцией к большему использованию UNIX [106—111,139] возрастает применение языка Си для создания систем управления окнами, операционных и сетевых систем. Он также используется многими физиками при создании систем сбора данных. Таким образом, есть большой вопрос для будущего: остановит ли Фортран-90 переход к Си [91]?

Однако применение Си не решает проблему ввода-вывода структур данных. Для ее решения была создана система СНЕЕТАН, которая обеспечивает пользователя набором Си-функций для управления этими структурами [92].

Техника объектно-ориентированного программирования позволяет создавать сложные программы с минимальными затратами [93—96].

Универсальный язык программирования Си++ является развитием языка Си в направлении объектной ориентированности и включения в него механизмов контроля типов, абстракции данных и совмещения операций [97].

Прототипы потенциально революционных новых программных средств для физического анализа рассмотрены П.Кунцем [38].

На рис.17 приведена схема этих средств, охватывающая весь цикл анализа, начиная с генерации событий и кончая представлением данных и фитированием.

Lund++ — программа для моделирования процессов генерации частиц [98,99]. Первоначальный вариант программы был написан в SLAC на языке Objective-C. В дальнейшем она была переписана на языке Си++. Эти прототипы показали, что ООП-подход является мощным средством для решения задачи моделирования.

Gismo — пакет программ для моделирования детекторов и реконструкции треков [35].

Reason — проект был начат летом 1989 г. для исследования возможности приложения техники визуального программирования [100].

CABS — пример использования объектно-ориентированной техники для физического анализа распадов *B*-мезонов [101].

На компьютерной конференции в 1992 г. примерам применения техники объектно-ориентированного программирования была посвящена специальная секция, на которой было сделано 11 докладов. Это говорит о все большем проникновении этих методов в физику высоких энергий.

Проводятся работы по автоматизации процесса генерации цепочек программ для анализа данных на Фортране, например, САВ (The Cosmos Application Builder) [102].

Таким образом, в настоящее время развитие программного обеспечения экспериментов в физике высоких энергий идет по двум направлениям в условиях «мирного сосуществования». Наряду с этим также разрабатываются методы параллельных вычислений [128,129].

**4.5. Обеспечение доступа к большим массивам данных.** В настоящее время и, вероятно, в течение последующих 20 лет большой проблемой в ФВЭ-компьютинге будет обеспечение доступа сотен физиков к большому количеству данных, которые хранятся на различных носителях информации.

Для решения этой задачи используются различные носители информации (магнитные ленты, картриджи с автоматическими загрузчиками, EXABYTES) и магнитные диски, на которые переписывается информация, требующаяся в данный момент для текущей работы. Для организации доступа к нужным данным разрабатываются подходы, основанные на использовании специализированных баз данных [103,104].

Возможно, что при снижении стоимости устройств прямого доступа эти данные будут запоминаться на устройствах с прямым доступом. В связи с этим следует, по-видимому, обратить внимание на оптические

диски, которые в настоящее время представляются наиболее перспективными носителями больших объемов информации. По сравнению с магнитными носителями они пока характеризуются большим временем доступа и более высокой стоимостью. Однако развитие операционных систем, общих баз данных, систем обработки изображений, возможно, приведет к смещению экономического баланса в пользу оптической памяти. Никакая другая технология не может сейчас конкурировать с технологиями оптической памяти в ситуациях, когда объем ежедневно записываемой информации составляет гигабайты. Внедрение перезаписываемых оптических систем, позволяющих работать фактически с безграничными объемами информации, возможно, создает хорошие предпосылки для последующего внедрения этой технологии [105].

Частное решение этой проблемы может опираться на то, что не все данные требуются физикам для анализа одинаково часто. Некоторые события требуются для изучения чаще, чем другие. Кроме того, в каждом событии имеется часть исходных или восстановленных величин, к которым обращаются чаще, чем к другим.

Одним из частных вариантов решения проблемы управления данными может быть не очень гибкое решение, которое было применено в L3-эксперименте. Его суть заключается в следующем: данные после реконструкции разделяются на 27 отдельных потоков. Одна половина потоков содержит полную информацию о событиях, другая — только часть наиболее часто используемых структур восстановленных данных. Информация, содержащаяся в каждом из потоков, объединяется в массивы по 200 Мбайт каждый, записывается на картриджи и при необходимости переписывается с них на диск для последующей пересылки на рабочие станции физиков. Такого типа системы используют специальные каталоги для обеспечения физикам доступа к данным и соответствующим устройствам [66].

Разнообразие операционных систем, с которыми приходится иметь дело физикам, создает определенные трудности в работе, и ситуация меняется в пользу операционных систем UNIX [16, 106—111, 139].

## **5. ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ЭКСПЕРИМЕНТОВ И МЕЖДУНАРОДНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ**

Вопросам информационного обеспечения исследований в физике высоких энергий уделяется большое внимание. Это проявляется, в частности, в создании специализированных баз данных и информационно-поисковых систем, например, FreeHEP, QSPIRES, PPDS, WWW.



Для сбора информации о программном обеспечении экспериментов в физике высоких энергий создана специальная организация FreeHEP, доступ к информационному фонду которой осуществляется с помощью электронной почты.

FreeHEP — это организация, предназначенная для предоставления информации об имеющемся программном обеспечении, которое может представлять интерес для исследований по физике высоких энергий и организации простого доступа к ней пользователей и разработчиков.

Так как накапливаемая в FreeHEP информация должна быть максимально полной, то она принимает все программное обеспечение, которое может быть полезно для физики высоких энергий.

По этой причине не устанавливаются требования на методы распространения, постановки, документацию, языки, машинную поддержку и т.д. Все это оставляется авторам. Развитие и поддержка программного обеспечения не является задачей FreeHEP. Ее функции ограничены сбором информации, обеспечением простого доступа к ней заинтересованных лиц и оказанием помощи авторам разработок и пользователей в установлении контактов.

Основными компонентами FreeHEP являются:

- а) глобальная компиляция программного обеспечения, представляющего интерес для физиков;
- б) удобные пути доступа к информации;
- в) пользовательская сеть группы новостей для того, чтобы пользователи и авторы могли обсуждать проблемы, предложения, новые идеи и т.д.;
- г) набор руководств по различным общим направлениям, например: «как использовать», «как работать с FreeHEP» и т.д.

Для того чтобы внести в информационный фонд FreeHEP информацию или получить ее из фонда, нужно войти в контакт с одним из редакторов с помощью электронной почты.

В настоящее время в FreeHEP имеются следующие разделы: Graphics, Visualization, GUI; Detector Simulation; Data Acquisition; Detector Simulation; Analysis and Data Reduction; Event Generators & Software Engineering; General Libraries; Parallelism and Distributed Computing; Software Engineering; Software compilation; Events Generators; Data Bases; CAD/CAE systems.

Предложение о создании этой системы было высказано на совещании NEPLIV, проводившемся в лаборатории SSC в январе 1991 г. Инициаторами ее создания и авторами первоначального варианта являются Тони Джонсон, Андреа Палонек и Саул Юсеф.

Информационно-поисковой системой, ориентированной на физические исследования, является система SPIRES (Stanford Physics

Information Retrieval System) [141]. Эта система состоит из нескольких баз данных, поддерживаемых SLAC в коллаборации с DESY, LBL и рядом других институтов. Доступ к ней осуществляется через удаленный сервер, установленный на BITNET узле SLACVM и обеспечивающий поиск информации для пользователей SLAC в интерактивном режиме. ОИЯИ расположен достаточно далеко от этого узла, и доступ к нему производится в настоящее время через электронную почту. Запросы выполняются в течение двух-трех суток, и объемы пересылаемых данных ограничены. Но тем не менее новейшая информация, поступающая в систему, может быть получена в ОИЯИ намного быстрее, чем это может быть сделано любым другим способом.

В Институте физики высоких энергий (Протвино) реализуется проект КОМПАС, цель которого — развитие первоначальной идеи групп систематизации LBL (США) и RAL (Англия) о создании системы баз данных физики частиц для сбора и оценивания экспериментальных данных и обеспечения их доступности для физического сообщества [142].

Создание международных компьютерных сетей, охватывающих научный и деловой мир, и базирующихся на них распределенных информационных систем привели к созданию специализированных систем поиска информации (GOPHER, WAIS и WWW).

Поясним их назначение на примере системы INTERNET [148]. Система INTERNET — это не просто электронная почта. Компании во всем мире используют ее для поиска и получения информации разных видов. Пользуясь этой всемирной информационной сверхскоростной магистралью, можно найти все — начиная с последних статей, посвященных исследованиям в области вселенского разума, и кончая текстами популярных песен.

Однако отыскать нужную информацию не всегда просто. Если принять во внимание тысячи баз данных и десятки навигационных систем, не исключено, что поиски в разветвленной сети INTERNET займут часы или даже дни. К счастью, некоторые удобные вспомогательные программы для INTERNET помогают быстро найти всю требуемую информацию.

INTERNET способна работать с тремя основными системами поиска и получения информации: GOPHER, WAIS (Wide-Area Information Servers) и WWW (World-Wide Web).

GOPHER — наиболее широко распространенная система поиска в сети INTERNET, позволяющая находить информацию с помощью ключевых слов и фраз. Работа с системой GOPHER начинается с просмотра оглавления, при этом пользователю предлагается пройти сквозь ряд вложенных меню и выбрать нужную тему. В INTERNET в настоящее время имеется свыше 2000 систем GOPHER, часть из которых является узко-

специализированными, а часть содержит более разностороннюю информацию.

Gopher позволяет получать информацию без указания имен и адресов авторов, благодаря чему пользователь не тратит много времени и нервов. Он просто сообщает системе Gopher, что именно ему нужно, и система находит соответствующие данные. Программное обеспечение для создания Gopher-сервера бесплатно распространяется в системе INTERNET, однако если пользователь устанавливает свой собственный Gopher, он должен купить лицензию у Университета штата Миннесота.

WAIS — еще более мощное средство получения информации, чем Gopher, поскольку оно осуществляет поиск ключевых слов во всем тексте документа. Запросы посылаются в WAIS на упрощенном английском языке. Это значительно легче, чем формулировать их на языке алгебры логики, и это делает WAIS более привлекательной для случайных пользователей. Имеется более 200 WAIS-библиотек, информация в которые предоставляется преимущественно сотрудниками академических организаций на добровольных началах. Поэтому большая часть материалов относится к области исследований и компьютерных наук.

WWW — одна из распределенных информационных систем, разработка которой была инициирована CERN в 1992 г. WWW представляет собой распределенный гипертекст, т.е. набор (очень большой) серверов во всем мире, общающихся с клиентами с помощью сетевого протокола TCP. Серверы, как правило, имеют на своем компьютере некий набор документов, к которым данная организация предоставляет доступ остальному миру. Отличительной чертой таких документов является возможность иметь в теле ссылки на другие (расположенные, быть может, в другой стране) документы или их части. Эта особенность и приводит к «паутине», взаимопереплетению документов, которые могут содержать inline графику, звук и т.д.

WWW очень активно используется в исследовательском сообществе, в частности, в CERN, для организации взаимодействия и общей информационной основы для рабочих групп, например, ATLAS. В ОИЯИ имеется возможность доступа к этой системе [147].

Эти информационные системы хорошо знают друг друга и могут эффективно взаимодействовать. Как правило, они «сидят» на общем поле документов. Рядом находится архивная и справочные службы (ARCHIE, FTP), базы данных, словом, некий информационный конгломерат, который должен помочь пользователю быстро и эффективно извлечь все, что он только захочет.

## 6. НОВОЕ ПОКОЛЕНИЕ ЭКСПЕРИМЕНТОВ В ФИЗИКЕ ВЫСОКИХ ЭНЕРГИЙ, ИХ ТРЕБОВАНИЯ К КОМПЬЮТИНГУ И НАЗНАЧЕНИЕ ЦЕНТРАЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Проблемы, которые поставила перед компьютерингом разработка новых ускорителей (LHC и SSC) и планируемые на них эксперименты, были рассмотрены в 1990 г. на симпозиуме, посвященном детекторам для SSC [140]. Несмотря на прекращение сооружения этого ускорителя в 1993 г. актуальность поднятых на симпозиуме проблем сохранилась, так как речь, в основном, шла о компьютерных средах нового «поколения» экспериментов и путях решения проблемы компьютеринга в физических исследованиях в условиях его интенсивного развития. Кроме того, весь комплекс поднятых на симпозиуме проблем актуален не только для экспериментов, планируемых на LHC, но также и для других больших электронных экспериментов, так как рассмотренный путь решения проблемы компьютеринга в новых условиях имеет достаточно общий характер.

При рассмотрении основных характеристик компьютерной проблемы для нового поколения коллайдеров отмечалось следующее.

1. Каждый эксперимент будет генерировать ежегодно от 10 до 100 Тбайт необработанных данных. Кроме того, при обработке типичных событий объем информации на одно событие будет удваиваться. К большинству этих данных должны иметь доступ физики основной лаборатории и сотрудничающих с ней удаленных институтов.

Система управления данными должна удовлетворять следующим требованиям. При решении задач, требующих большой статистики, потребуется просматривать большие массивы данных событие за событием, выбирая лишь несколько цифр, относящихся к каждому из них. С другой стороны, при исследовании редких процессов потребуется поиск отдельных событий в огромных массивах данных с тем, чтобы иметь возможность анализировать все относящиеся к ним данные. Организация данных должна обеспечивать эффективный поиск этих двух типов информации.

2. Физикам потребуется хороший интерактивный компьютеринг и эффективный доступ к большим базам данных. Программное обеспечение должно включать статистические пакеты, мощную графику и интерфейсы пользователя для построения сложных программ из простых модулей. Важной целью будущих программных систем, которые необходимо реализовать, должна быть такая организация работы программ, при которой физикам кажется, что они работают на своих локальных рабочих станциях. Они не должны заботиться о том, где выполняются их программы и где хранятся требующиеся им данные.

3. Текст программного обеспечения современных экспериментов достигает миллиона строк на Фортране. Разработка таких программных систем является сложной задачей для любой группы. Для обеспечения надежности создаваемого программного обеспечения и сопровождения его не только авторами необходимо использовать при конструировании, документировании и тестировании программ современные средства разработки, а также внедрять стандарты и фирменные разработки.

Поскольку создание распределенных систем на базе RISC-процессоров и мощных рабочих станций обеспечивает вычислительный ресурс экспериментов, то решение компьютерной проблемы сводится к решению следующих задач:

- организация хранения и управления данными;
- создание соответствующей компьютерной среды;
- внедрение современных методов разработки программного обеспечения.

Далее было отмечено, что разработка систем программного обеспечения для новых экспериментов потребует также изменения стиля и методов работы физиков. Решение этой задачи должно производиться совместными усилиями трех групп: самих экспериментаторов, программистов основной лаборатории и компьютерной индустрии.

Задача экспериментаторов — планировать, проектировать и финансировать компьютерные системы, так же, как и детекторы. В то же время им потребуются совместно с руководством лаборатории определять требования к компьютерным системам и программному обеспечению, которые должны поддерживаться центральным вычислительным комплексом.

Задача лаборатории — обеспечивать централизованную поддержку экспериментов путем создания соответствующих вычислительных и программных систем, разработки и внедрения методов их создания на основе новейших технологий и промышленных разработок.

Лаборатория должна также разработать механизм, с помощью которого группы экспериментаторов смогут управлять выбором систем, поддерживаемых на уровне лаборатории, а также способствовать поддержке подходящих стандартов и мониторингованию соответствующих разработок в компьютерной промышленности.

Опыт CERN и других ведущих физических центров показывает, что даже оснащение крупных электронных экспериментов собственными вычислительными ресурсами, включая достаточно мощные системы, не приводит к снижению темпов развития центральных вычислительных систем. Более того, усложнение характера экспериментов порождает новые требования к центральным вычислительным ресурсам. Поэтому увеличение центрального вычислительного ресурса по-прежнему остается одной из актуальных задач, при решении которой необходимо на-

учиться оценивать реальные потребности экспериментальных групп и ситуацию динамично развивающегося компьютерного рынка [33].

Другой аспект деятельности ЦВК связан с хранением и обеспечением быстрого поиска информации среди огромных объемов данных.

Естественной задачей ЦВК является организация связи центра с мировым научным сообществом.

Наличие в институтах большого парка разнообразных ЭВМ также требует централизованной службы поддержки их аппаратных и программных средств.

Интенсивное развитие компьютерной науки, включая разработку качественно новых технологий для развития программных систем, требует их внедрения в практику. Это и разработка графических интерфейсов пользователя, которые являются решающим фактором улучшения качества программного обеспечения для научных исследований, и использование стандартов, и поиск наиболее подходящих коммерческих и свободно распространяемых продуктов.

Так, например, популярность графических интерфейсов на рынке ПК демонстрирует необходимость следовать этому примеру при разработке программ для исследований в области физики высоких энергий.

Другим новым продуктом, который, вероятно, сильно повлияет на научный компьютеринг, является AVS (Application Visualization System from Stardent Computers). Используя AVS, физик может развивать сложные графические приложения без написания кодов. Этот продукт использовался в D0-эксперименте для представления событий в пространстве (3-мерное изображение). Этот подход подает большие надежды и используется для разработки более общих средств.

Современная генерация экспериментов уже перекрывает по времени более чем один цикл компьютерной технологии. Это обстоятельство создает новые проблемы для разработчиков программного обеспечения. Чтобы обеспечить «выживаемость» систем при изменениях в технологии, новое программное обеспечение должно в максимальной степени использовать коммерческие продукты. Все программное обеспечение, разрабатываемое и покупаемое, должно базироваться на стандартах.

В новых условиях центральные вычислительные комплексы, наряду с решением традиционных задач по обеспечению вычислительными ресурсами, хранением и распространением данных, сетевой поддержкой и информационным обеспечением экспериментов, организацией и координацией работ физиков и промышленности, будут играть роль «мозговых» центров по разработке и внедрению новых технологий в системы обработки и анализа экспериментальных данных, базирующихся в значительной степени на перспективных разработках аппаратных и программных средств.

## ЗАКЛЮЧЕНИЕ

Для эволюции компьютеринга в физике высоких энергий в последние годы было характерно следующее.

1. Прогресс в развитии распределенных вычислительных систем, эксплуатации быстро развивающихся рабочих станций и сетей передачи данных.

2. Рост вычислительной мощности для пакетной обработки данных, включая обычные ЭВМ общего назначения, разработка и эксплуатация в широких масштабах дешевых распределенных вычислительных систем, что в конечном итоге привело к постановке вопроса о замене больших ЭВМ распределенными вычислительными системами на базе RISC-процессоров и мощных рабочих станций.

3. Становление персональных рабочих станций как основного орудия для интерактивного анализа данных.

4. Внедрение локальных и глобальных сетей передачи данных.

Большое внимание центры физики высоких энергий уделяют привлечению к подготовке и проведению исследований ученых из других центров и высших учебных заведений, обеспечивая им возможность участия в обработке данных с рабочих мест.

Для активного участия в экспериментах необходимы хорошие средства компьютерной связи, соответствующие вычислительные ресурсы, дисковое пространство, системы массовой памяти и графические рабочие станции.

Создание в экспериментальных группах собственных мощных вычислительных систем, обеспечивающих реконструкцию и физический анализ данных, не привело к уменьшению темпов развития центральных вычислительных комплексов, но несколько изменило их функции.

В заключение автор выражает глубокую благодарность В.В.Кухтину за полезные советы и рекомендации, А.Г.Заикиной и О.И.Попковой за помощь в оформлении работы.

## СПИСОК ЛИТЕРАТУРЫ

1. CHER 1984. Proc. Symp. Recent Developments in Computing, Processor and Software Research for High Energy Physics. Ed. R. Donaldson and M. Kreisler. Mexico City: Univ. Nac. Auto. Mexico.
2. CHER 1985. Proc. Conf. on Computing in High Energy Physics. Ed. L.O. Hertzberger and W. Hoogland, Amsterdam: Elsevier.
3. CHER 1987. Proc. Int. Conf. Computing in High Energy Physics. Ed. W. Ash, Comput. Phys. Commun., 1987, vol. 45.
4. Carpenter B., Jones C., Kellner G. et al. — Report CERN DD/88/1, Geneva, 1988.
5. CHER 1988. Proc. Int. Conf. on Impact of Digital Microelectronics and Microprocessors on Particle Physics. Ed. M. Budinich, E. Calstelli and A. Colavatia. Singapore: World Scientific.

6. CHEP 1989. Proc. Int. Conf. on Computing in High Energy Physics. Ed. R.C.E. Devenish and T. Daniels. Comput. Phys. Commun., 1989, vol.57.
7. CHEP 1990. Computing for High Luminosity and High Intensity Facilities (AIP Conf. Proc., vol.209). Ed. J.Lillberg and M.Oothoud.
8. Proc. Symp. on Detector Research and Development for the Superconducting Super Collider. Ed. T.Dombeck, V.Kelly and G.P.Yost, Fort Worth, Texas.
9. CHEP 1991. Proc. Int. Conf. Computing in High Energy Physics'91. Ed. Y.Watase and F.Abe. Tokyo: Universal Academy Press.
10. Proc. Int. Workshop on Software Engineering, Artificial Intelligence and Expert Systems in High Energy and Nuclear Physics. Lyon, France, 19—24 March, 1990. Singapore: World Scientific Publishing, 1991.
11. Computing at CERN in the 1990s, CERN Int. Report, 1989.
12. CHEP 1992. Proc. Int. Conf. Computing in High Energy Physics'92. Ed. C.Werkerk and W.Wojcik. 1992, CERN 92-07.
13. Delfino M., Pacheco A. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.177.
14. Garnto K.A., Ikeda M., Levinthal D. et al. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.171.
15. Grosdidier G. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.459.
16. Mount R.P. — Phys. Rep. Prog. Phys., 1992, vol.55, p.1385.
17. Mount R.P. — CERN School of Computing. CERN, Geneva, 1989, 89-06, p.306.
18. Абрамов А.И., Казанский Ю.А., Матусевич Е.С. — Основы экспериментальных методов ядерной физики. М.: Энергоатомиздат, 1985.
19. Мохов Н.В. — ЭЧАЯ, 1987, т.18, вып.5, с.960.
20. Newman H.B. — Comput. Phys. Commun., 1987, vol.45, p.27.
21. Letter of Intent by the Solenoidal Detector Collaboration, 1990, SDS-90-00151.
22. CERN/LEPC/84-6, LEPC/PR4/L3, 23 January, 1984.
23. Dahl-Jensen E., Dress J.K., Grant A. et al. — Report DELPHI Collaboration. DELPHI 88-65 PROG., 116 CERN, 1988.
24. Смирнов Н.Н. — Программные средства персональных ЭВМ. Л.: Машиностроение, 1990.
25. Иванов В.Г., Нодарсе Фернандес Ф. — Сообщение ОИЯИ P12-92-386, Дубна, 1992.
26. Eisenhandler E. — RAL-88-026, Manchester, 1988.
27. Le Du P. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.45.
28. Карлов А.А., Ломов А.К., Смолякова Т.Ф. — В сб.: Труды Межд. школы по вопросам применения ЭВМ в физических исследованиях. ОИЯИ, Д10-89-70, Дубна, 1989, с.73.
29. Kindel В. — How Fast is Fast. BYTE, 1989.  
Липкин И. — Компьютер Пресс, 1990, т.5, с.7.
30. The Computer Journal. 1975, vol.19, No.1, p.43.
31. Communication of the ACM. 1984, vol.27, No.106, p.1013.
32. Cormell L.R. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.77.
33. Williams D.O. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.67.
34. Werner C.M.L., Souza J.M. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.425.
35. Atwood W.B., Burnett T.H., Cailliau R. et al. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.433.
36. Katayama H. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.439.
37. Oleynik G.A. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.445.
38. Kunz P.F. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.303.
39. Brun R., Zoll J. — CERN Program Library Q100, 1987.
40. Zoll J. — CERN Program Library Q101, 1989.
41. Brun R., Hansroul M., Lassale J.C. — CERN Program Library Q210, 1984.
42. Blobel V. — DESY Int. Report R1-88-91, 1988.



43. Klein H., Zoll J. — CERN Program Library Q400, 1980.
44. Говорун Н.Н., Дорж Л., Иванов В.Г. и др. — ЭЧАЯ, 1975, т.6, вып.3, с.743.
45. Brun M., Brun R., Rademakers A. — Comput. Phys. Commun., 1989, vol.57, p.235.
46. Adeve B., Bagnaia P., Banerjee S. et al. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.323.
47. Mount R.P. — Comput. Phys. Commun., 1987, vol.45, p.299.
48. Cranfield R., Holl B., Jones R.W.L. — 1991 OPAL User Guide OPAL Collaboration, CERN OC504/OPAL/OFFL/36/0003.
49. Jadach S., Ward B.F., Was Z. — In Z Physics at LEP 1. Ed. Altarelli G., Kleiss R., Verzegnassi C. CERN Yellow Report 89-03, 1989.
50. Jadach S., Was Z. — Comput. Phys. Commun., 1985, vol.36, p.191.
51. Sjostrand T. — Comput. Phys. Commun., 1982, vol.27, p.243.
52. Sjostrand T., Bengtsson M. — Comput. Phys. Commun., 1987, vol.43, p.367.
53. Marchesini G., Webber B. — Nucl. Phys., 1988, vol.B310, p.461.
54. Knowles I.G. — Nucl. Phys., 1988, vol.B310, p.571.
55. Catani S., Marchesini G., Webber B. — Nucl. Phys., 1991, vol.B349, p.635.
56. Abbiendi G., Stanco L. — 1991 Comput. Phys. Commun., 1991, vol.66, p.16.
57. Marchesini G., Webber B., Abbiendi G. et al. — Comput. Phys. Commun., 1992, vol.67, p.465.
58. Bengtsson H.-U. — Comput. Phys. Commun., 1984, vol.31, p.323.
59. Bengtsson H.-U., Sjostrand T. — Comput. Phys. Commun., 1987, vol.46, p.43.
60. Sjostrand T., van Zijl M. — Phys. Rev., 1987, vol.D36, p.2019.
61. Paige F., Protopopescu S. — Report BNL, 1986, BNL 38774.
62. Brun R., Bruyant F., Maire M. et al. — CERN Program Library W5013, 1986.
63. Brun R., Carminati F. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.451.
64. Nelson W.W., Hirayama H., Rogers D.W.O. — Report SLAC, 1985, SLAC-265.
65. Fesefeldt H. — Phys. Inst., RWTH Aachen PITHA 85/2, 1985.
66. Shiers J., Goossens M. — CERN Program Library Q123, 1991.
67. Shiers J. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.329.
68. Brun R., Couet O., Vandoni C. et al. — CERN Program Library Q121, 1989.
69. Bengtsson H.U. — Comput. Phys. Commun., 1984, vol.31, p.323.
70. Anderson V. et al. — Phys. Rev., 1983, vol.97, p.33.
71. Калиновский А.Н., Мохов Н.В., Никитин Ю.П. — Прохождение частиц высоких энергий через вещество. М.: Энергоатомиздат, 1985.
72. Mokhov N.B., Gossairt J.D. — Nucl. Instr. Meth., 1986, vol.A244, p.349.
73. Matthens R. — CERN Program Library Z303, 1987.
74. Putzer A. — Comput. Phys. Commun., 1989, vol.57, p.156.
75. Putzer A. — Proc. CERN School of Computing Troio. Sept. 1987; Heidelberg Univer. Rep. HD-IHEP-88-02, 1988.
76. Blobel V. et al. — Databases and Bookkeeping for HEP Experiments, RL-83-085, 1983.
77. Gopal G.P. et al. — Report CERN, DELPHI 86-28 PROG-46, 1986.
78. Mount R.P. — Comput. Phys. Commun., 1987, vol.45, p.299.
79. Nagy E. — L3 Report, No.486, 1987.
80. Fisher S.M., Palazzi P. — Comput. Phys. Commun., 1989, vol.57, p.169.
81. Palounek A.P.T., Youssef S. — LBL-2915-mc, May 1990.
82. Bassler E. — Comput. Phys. Commun., 1987, vol.45, p.201.
83. Hartwic Albrecht (DESY) — Private Communication.
84. Burnet T.H. — Comput. Phys. Commun., 1987, vol.45, p.195.
85. Bock R.K. — Comput. Phys. Commun., 1987, vol.45, p.15.
86. Brun R., Carminati F., Marquina M. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.315.

87. Maidantchik C., da Rocha A.R.C., de Souza J.M. et al. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.335.
88. Brun R., Lienart D. — CERN Program Library Y250, 1984.
89. James F. — CERN Program Library Q121, 1989.
90. Metcalf M. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.411.
91. Керниган Б., Ритчи Д. — Язык программирования Си: Пер. с англ. М.: Финансы и статистика, 1992.
92. Kunz P., Word G. — Proc. of the Workshop on Data Structures for Particle Physics Experiments. Erice, Nov. 1990.
93. Федоров А.Г. — Мир ПК, 1991, 3, с.20.
94. Kunz P. — Proc. Computing for High Luminosity and High Intensity Facilities. Santa Fe, April 1990.
95. Рассохин Д.Н. — Мир ПК, 1992, 6, с.120.
96. Рассохин Д.Н. — Мир ПК, 1992, 7, с.116.
97. Страуструп Б. — Язык программирования Си++: Пер. с англ. М.: Радио и связь, 1991.
98. Atwood W.B. et al. — Proc. Symposium on Detector Research and Development for the Supercollider, Fort Worth, 1990.
99. Lonnblad L. — University of Lund. Priv. commun.
100. Atwood W.B. et al. — The Reason Project. Proc. Computing for High Luminosity and High Intensity Facilities. Santa Fe, April 1990.
101. Katayama N. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.339.
102. Хехео G., de Souza J. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.359.
103. Baden A., Grossman R. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.59.
104. Varone L.M. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.299.
105. Липкин И., Николаев А. — Компьютер Пресс (обозрение зарубежной прессы), 1990, вып.5.
106. Butler J.N. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.555.
107. Eades C.A. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.583.
108. Lauer R. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.591.
109. Pabrai U. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.597.
110. Nichols J. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.605.
111. Nichols J. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.611.
112. Sambel J.N. — Comput. Phys. Commun., 1989, vol.57, p.129.
113. Мячев А.А. — Персональные ЭВМ: краткий энциклопедический справочник. М.: Финансы и статистика, 1992.
114. Смирнов А.Д. — Архитектура вычислительных систем. М.: Наука, 1990.
115. Dolbilov A., Dorokhin A., Fariseev V. et al. — Proc. 4th Joint European Networking Conference «European Research Networking in a Global Context», 1993, Rare, p.214.
116. Fluckiger F. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.709.
117. Chartrand G. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.703.
118. Kunz P.F. — Nucl. Instr. Meth., 1976, vol.135, p.435.
119. Nalatsis C. et al. — Comput. Architecture News, 1980, vol.8, p.278.
120. Kunz P.F. — CHEP 1984 (Computing in High Energy Physics Conf. Series), p.197.
121. Brafman H., Notz D. — CHEP 1984 (Computing in High Energy Physics Conf. Series), p.211.
122. Ferran P.M. et al. — CHEP 1985 (Computing in High Energy Physics Conf. Series), p.322.
123. Metcalf R., Boggs D. — Commun. ACM, 1976, vol.19, p.395.
124. Mount R.P. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.691.
125. Baund J.-P., Bunn J., Cane F. et al. — CHEP 1991 (Computing in High Energy Physics Conf. Series), p.571.

126. Biel J., Areti H., Atac R. et al. — *Comp. Phys. Comm.*, 1987, vol.45, p.331.
127. May E. — *Comp. Phys. Comm.*, 1989, vol.57, p.278.
128. Dittus F. — *Comp. Phys. Comm.*, 1989, vol.57, p.395.
129. Cooper P. — *CHEP 1990 (Computing in High Energy Physics Conf. Series)*, p.3.
130. Mount R. — *CHEP 1990 (Computing in High Energy Physics Conf. Series)*, p.44.
131. Amako K. — *CHEP 1991 (Computing in High Energy Physics Conf. Series)*, p.11.
132. Hash T. et al. — *Fermilab's Advanced Computer Research and Development Program (Fermilab FN 83)*, 1983.  
Gaines I., Areti H., Atac R. et al. — *Comp. Phys. Comm.*, 1987, vol.45, p.323.  
Hash T. — *Comp. Phys. Comm.*, 1989, vol.57, p.47.
133. Delfino M. — *CHEP 1991 (Computing in High Energy Physics Conf. Series)*, p.23.
134. Butler J.N. — *CHEP 1992 (Computing in High Energy Physics Conf. Series)*, p.6.
135. Altaber J., Cannon S., Carpenter B. et al. — *CHEP 1992 (Computing in High Energy Physics Conf. Series)*, p.403.
136. Wechselberger G. — *Computer World Moscow*, 1993, No.39, p.10.
137. Watase Y. — *Comp. Phys. Comm.*, 1989, vol.57, p.198.
138. Etienne F. — *CHEP 1991 (Computing in High Energy Physics Conf. Series)*, p.515.
139. Doepfner T.W.Jr. — *CHEP 1992 (Computing in High Energy Physics Conf. Series)*, p.123.
140. Loken S.C. — *Proc. Symp. on Detector Research and Development for the SSC. Fort Worth, Texas, Oct. 15—18, 1990*, p.604.
141. Galic H. — *SLAC-Report-393*, June 1992.
142. Алехин С.И., Базеева В.В., Грудцин С.С. и др. — В сб.: *Тр. Межд. школы по вопросам применения ЭВМ в физических исследованиях. ОИЯИ Д10-89-70, Дубна, 1989*, с.208.
143. *CERN Annual Report 1992, vol.II*, p.7.
144. *CERN Computer Newsletter, March—April 1992, No.206*.
145. *CERN Computer Newsletter, January—February 1993, No.210*.
146. *KEK Annual Report April 1992 — March 1993*.
147. Bernes-Lee T., Cailliau R. — *CHEP 1992 (Computing in High Energy Physics Conf. Series)*, p.75.
148. Levin J. — *Computer World Moscow*, 1993, No.45, p.23.