

93 - 131



сообщения  
объединенного  
института  
ядерных  
исследований  
Дубна

P10-93-131

В.В.Первушов

ОСОБЕННОСТИ  
ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ  
МАТЕМАТИЧЕСКОЙ ОБРАБОТКИ  
 $\uparrow d-p$ -ЭКСПЕРИМЕНТА В РАМКАХ СВМ ЕС ЭВМ

1993

По исторически сложившимся причинам математическая обработка данных со 100-сантиметровой водородной пузырьковой камеры, экспонированной в пучке поляризованных ядер дейтронов ( $\alpha$ - $d$ - $p$ -эксперимент), проводилась на ЭВМ CDC-6500 с помощью специализированной высокоавтоматизированной системы программ<sup>/1/</sup>, которая позволяла в достаточно сжатые сроки обрабатывать экспериментальный материал.

Однако по причинам, связанным с ростом объемов информации (до 100 тысяч событий в год), физическим и моральным старением ЭВМ CDC-6500 (ненадежная работа магнитофонов, ограниченность ресурсов дисковой памяти и быстродействия), изменением в методике обработки, вызванным уточнением и коррекцией исходных данных, а также накопленным опытом работы, возникла необходимость переноса программного обеспечения  $\alpha$ - $d$ - $p$ -эксперимента на ЕС ЭВМ и его модернизации.

Кроме того, опыт, накопленный в ходе обработки  $^{16}O$ - $p$ -эксперимента на ЕС ЭВМ в рамках операционной системы СВМ с использованием различных возможностей этой системы, заставил пересмотреть традиционную "идеологию" обработки, ориентированную на пакетный (batch) режим, обусловленный операционной системой CDC-6500<sup>/2/</sup>. Поэтому встала задача не только по переносу и адаптации математического обеспечения  $\alpha$ - $d$ - $p$ -эксперимента применительно к новой операционной системе, но и адаптации всех программ комплекса к иной технологии и интерфейсу<sup>/3/</sup>, практически полностью ориентированных на интерактивный режим работы. Поставленная задача была успешно реализована. Особенности, связанным с данной реализацией, и посвящено содержание работы.

1  
ЭВМ

## Основные технологические этапы и программы обработки

Стандартная схема движения потока экспериментальных данных, характерная для реализации на любых ЭВМ /1,4,5/, показана на рис. 1.

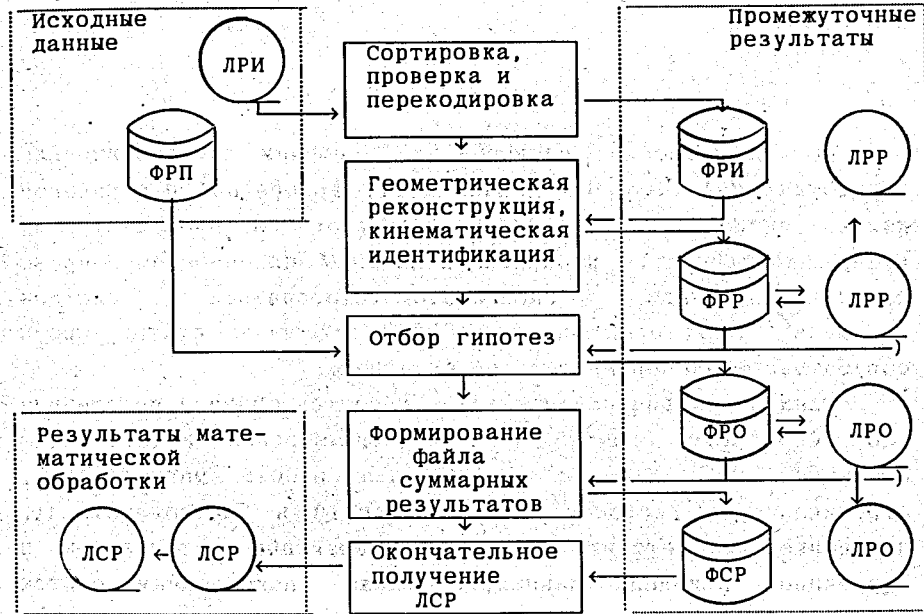


Рис. 1

Исходными данными для всего комплекса программ являются магнитные ленты с результатами измерений (ЛРИ) и информация предварительного просмотра фотоплёнок, записанная на файл результатов просмотра (ФРП), который кроме информации, необходимой для отождествления результатов просмотра и измерений, содержит визуальную оценку плотностей почернения изображений вторичных треков событий по трехбалльной шкале /6/. В ходе обработки информация на ФРП уточняется, проверяется и дополняется. На первом этапе производится сортировка, проверка и преобразование результатов измерений в стандартный формат системы "Гидра" /7,8/, данные в котором

записываются на файл результатов измерений (ФРИ). На втором этапе производится восстановление пространственной картины измеренных событий и кинематическая идентификация результатов реконструкции. Результаты реконструкции накапливаются на файле результатов реконструкции (ФРР). На следующем этапе из результатов реконструкции выбираются и переписываются на файл результатов отбора (ФРО) только те гипотезы, которые удовлетворяют заданным критериям и визуальным оценкам ионизации, определённым на этапе предварительного просмотра.

При исчерпании лимитов дискового пространства, а также для обеспечения надёжности хранения данных файлы ФРР и ФРО "сбрасываются" соответственно на магнитные ленты ЛРР и ЛРО с помощью специальной процедуры.

На основании ФРО (ЛРО) формируется файл суммарных результатов ФСР. На завершающем этапе после подведения итогов и итоговых проверок формируются ленты суммарных результатов ЛСР данного эксперимента.

Особенностями работы в рамках СВМ являются штатные средства операционной системы, обеспечивающие интерактивное взаимодействие человека с вычислительной машиной, взаимодействия между этапами обработки и автоматического протоколирования всех этапов процесса обработки. Основные компоненты этого математического обеспечения и их взаимосвязь /3/ показаны на рис. 2.

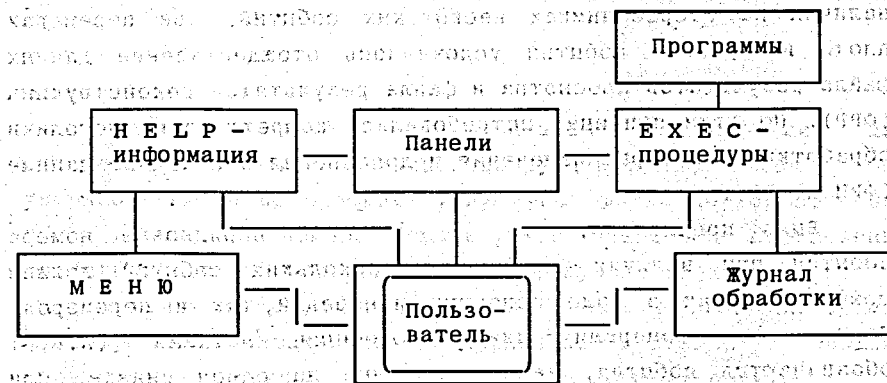


Рис. 2

Конкретно это означает, что для каждого этапа обработки разработана на специальном языке REXX [9] соответствующая ЕХЕС-процедура, которая обеспечивает следующие функции:

- формирование титульных данных для соответствующих программ;
- организацию панели для передачи параметров, определяемых пользователем;
- анализ и проверку заданных параметров, полный контроль вводимых данных;
- организацию диалога с пользователем при запросах, в случае каких-либо затруднений, с помощью организации режима меню для обращения к HELP - информации;
- запуск соответствующих программ на счёт;
- регистрацию протокола прохождения задачи на специальном файле.

#### Некоторые изменения методики обработки $d-p$ -эксперимента

Опыт, полученный при массовой обработке данных эксперимента с пучком поляризованных дейтронов по стандартной методике, показал, что очень часто автоматический отбор физических гипотез по ионизационным критериям сильно затруднён субъективными факторами при оценке ионизации, связанными с разбросом в качестве фотографий. Кроме того, при наличии на стереоснимках нескольких событий, при перемерах плохо измеренных событий усложнялось отождествление данных файла результатов просмотра и файла результатов реконструкции (ФРР). По этим причинам потребовалась корректировка методики обработки, некоторые изменения и дополнения в исходные данные (ФРИ и ФРР).

Было предложено, что порядковые номера событий при наличии на снимке нескольких событий должны сохраняться как в ходе основных измерений, так и перемеров. Порядковыми номерами или различными метками (А, В, С) обозначаются события, зафиксированные на одном снимке. При существовавшей ранее практике измерений при перемерах плохо

измеренных событий повторно измеренному событию мог быть приписан номер, отличавшийся от номера предыдущих измерений, что существенно затрудняло отождествление событий. Кроме того, события с одинаковыми номерами (т.е. на снимке в ФРП и измерение в ФРИ) считаются тождественными, если топология (число вторичных треков) восстановленного события совпадает с топологией, заданной в соответствующей записи ФРП, а X и Y - координаты вершины на первой проекции - находятся в пределах зоны, заданной в ФРП. Границы зон определяются как линии, соединяющие соответствующие пары крестов на плоскости первого снимка, с учётом погрешностей вблизи границы зон, как, например, это показано на рис. 3.

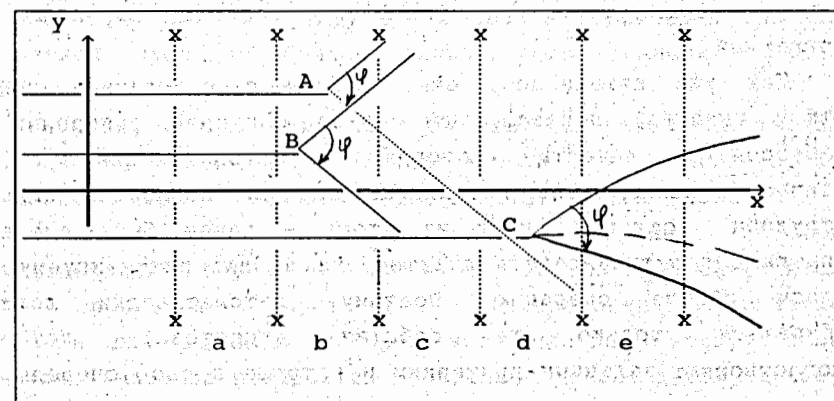


Рис. 3

Вторичные треки перед началом реконструкции для отождествления результатов реконструкции с информацией в ФРП упорядочиваются по плоскому углу (угол  $\varphi$ ) на первой проекции, и их порядковые номера используются затем для определения визуальной оценки ионизации соответствующих треков в ФРР.

В связи с тем, что порядковый номер события на снимке фиксируется и сохраняется при перемерах (в отличие от прошлой практики измерений), то это позволяет идентифицировать события, вершины которых находятся в одной зоне, не прибегая

к их записи в отдельный файл для последующего упорядочивания по Y-й координате. То есть отбор гипотез для всех событий производится стандартным способом в ходе одного прогона программы. Следует подчеркнуть, что идентификация событий по их порядковым номерам на снимке требует абсолютной надёжности, поскольку у программы нет никакой возможности обнаружить ошибку, а тем более её исправить.

Благодаря упорядочиванию вторичных треков по углу  $\varphi$  на плоскости первого снимка нет необходимости проверять порядок измерений вторичных треков, что позволило исключить из программы соответствующие проверки. Здесь также следует подчеркнуть, что какие-либо перепутывания номеров вторичных треков из-за ошибок или вследствие влияния геометрических факторов должны решаться с помощью ФРП-файла, но никак не программой.

Как уже отмечалось, опыт работы с программой CLERN (отбор гипотез) показал, что из-за факторов, связанных с разбросом в качестве фотографий, автоматический отбор физических гипотез по визуальным оценкам ионизации сильно затруднён. Реальный вывод из этого - ранее была сделана попытка алгоритмизировать практически не алгоритмизируемую в полном объеме операцию. Поэтому автоматически должны отбираться только те события (гипотезы), которые удовлетворяют заданным критериям и которые, в свою очередь, не вызывают сомнений. Программы не должны делать никаких попыток "спасения" забракованных событий, это всё должны делать сами физики. Для этого имеется возможность отбора гипотез по их номерам, которые заносятся в соответствующие SCAN-карты (записи в ФРП) /8/ для последующего программного отбора физических гипотез.

Исходя из этого измеренное и обчисленное по программе "Геокин" /5/ (геометрическая реконструкция и кинематическая идентификация) событие отбирается для программной идентификации, если выполняются следующие условия:

1) Все треки события восстановлены.

2) Для анализируемого события в ФРП найдена соответствующая запись (SCAN-карта). Запись ФРП считается соответствующей событию, если у них (события и записи) одинаковые номера фотоплёнок, номера стереокадров на фотопленках и порядковые номера событий на стереоснимках. Кроме того, X и Y-ые координаты вершины события должны находиться в пределах зоны, указанной в записи ФРП, а топологии восстановленного и найденного событий совпадают.

3) Визуальные оценки ионизации треков, занесённые в SCAN-карту, соответствуют значениям, определённым в таблице 1.

4) В событии нет треков с углами погружения больше задаваемого в титульных данных.

5) Плоский угол между двумя любыми вторичными треками на первой проекции больше минимального значения, также задаваемого в титульных данных.

События, не прошедшие хотя бы один из указанных тестов, выдаются на печать вместе с информацией о всех SCAN-картах для данного кадра, включая номера гипотез и "BAD"-события.

Автоматический отбор гипотез среди событий, удовлетворяющих перечисленным выше критериям, производится путём сравнения теоретических (вычисленных) значений ионизации для заданных наборов масс частиц с визуальными оценками. Гипотеза отбирается программой только в том случае, если теоретические значения ионизации соответствуют их визуальным оценкам в соответствии с таблицей 1.

Таблица 1

| Теоретическая ионизация | Визуальная оценка |
|-------------------------|-------------------|
| от 1.0 до M1            | 0                 |
| от M2 до M3             | 1                 |
| больше M4               | 2                 |

Если хотя бы для одного трека (частицы) нет такого соответствия, то гипотеза не отбирается и рассматривается следующая. Всегда должно выполняться условие  $1.0 < M1 < M2 < M3 < M4$ . равенство  $M1=1.0$  ( $M2=M3$ ) означает, что ионизация с визуальными оценками "0" ("1") не может быть программно идентифицирована и события, содержащие все гипотезы, соответствующие визуальной оценке ионизации "0" ("1") в SCAN-картах, будут распечатаны для идентификации физиками. Событие считается программно идентифицированным, если отбирается только одна гипотеза. Во всех остальных случаях события считаются не идентифицируемыми программой и выдаются на печать для анализа физиками. Если программа не идентифицировала событие или, наоборот, отобрала подходящую гипотезу, в обоих случаях перед обращением к печати следует вызов подпрограммы физика (CALL PHISIC). Эта подпрограмма разрабатывается и отлаживается физиками самостоятельно. В стандартном варианте подпрограмма PHISIC "пустая".

Существуют следующие режимы работы программы CLERN:

- 1) полной распечатки всех событий без отбора физических гипотез;
- 2) с программным отбором физических гипотез; с формированием файла для DST и без создания такого файла.

#### Особенности обработки $\alpha$ -d-p-эксперимента в рамках СВМ

Стандартная методика обработки  $\alpha$ -d-p-эксперимента во многом определялась вычислительными ресурсами CDC-6500. Так, например, самая ресурсоёмкая программа "Геокин" имеет оверлейную структуру и обрабатывает до 10 событий в минуту. Поэтому любое внесение изменений в методику, которое потребовало бы пересчёта большого массива событий, было связано с существенными затратами машинного времени. Это заставляло требовать усовершенствования методики привязывать к накопленному файлу результатов реконструкции (ФРР). Пакетный режим обработки затруднял оперативное уточнение исходных констант, задаваемых в блоке титульной информации.

Учитывая, что процессор ЕС-1066 примерно в пять раз быстрее, чем процессор CDC-6500, и при большей оперативной памяти не нужна оверлейная структура программы, ожидалось существенное увеличение производительности всего комплекса программ. На практике производительность возросла примерно на порядок; так, программа "Геокин" обрабатывает до 100 событий в минуту, это означает, что стандартный массив измеренных событий в среднем около 1000 событий (средний объём, поступающий с измерительных устройств) обрабатывается за 10 минут. Поэтому любые ошибки или несоответствия, допущенные при определении титульных констант, можно обнаружить и исправить в интерактивном режиме. За несколько обчётов всего стандартного массива событий можно более тщательно подобрать константы.

Увеличение производительности позволяет все этапы обработки проводить в интерактивном режиме и при необходимости оперативно вносить изменения не только в титульные данные, но и в методику обработки, пересчитывая по скорректированной методике накопленные массивы данных.

Программы обработки  $\alpha$ -d-p-эксперимента в рамках СВМ используют развитый интерфейс общения с пользователем, разработанный для  $^{16}O$ -p-эксперимента, включая автоматическое ведение журнала обработки. Меню, HELP-информация, ЕХЕС-процедуры, панели и средства программного сопровождения процесса обработки все вместе создают "дружественный" интерфейс, обеспечивающий пользователю-физику удобство при обращении к ЭВМ <sup>13/</sup>. Адаптация комплекса программ  $\alpha$ -d-p-эксперимента к данным средствам общения с ЭВМ потребовала минимальных переделок и привела к стандартизации и унификации всего процесса обработки данных со 100-сантиметровой водородной пузырьковой камеры.

Первые результаты работы по изменённой методике на ЕС ЭВМ показали, что возможность интерактивной работы с большими массивами данных, оперативное изменение параметров и констант дают не только преимущества, но порождают некоторые проблемы. Возникает опасность многих неконтролируемых и не-

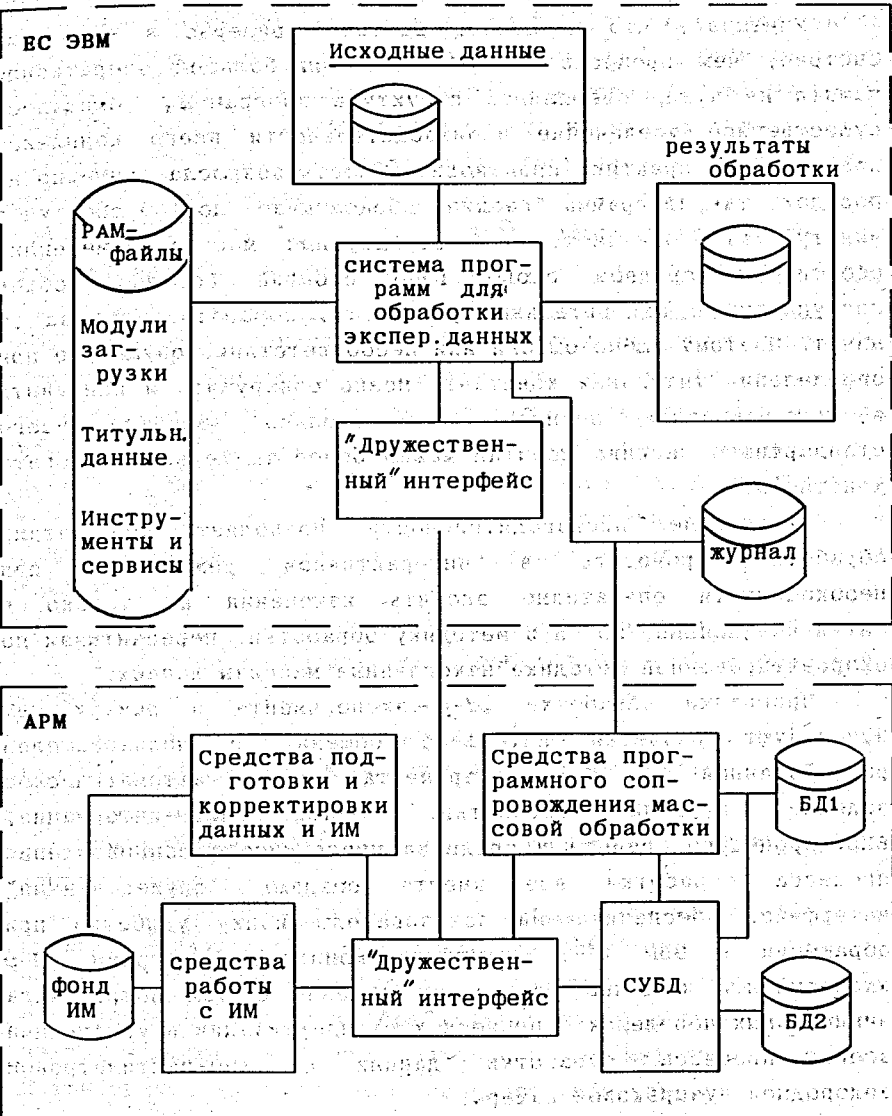


Рис. 4

регистрируемых пересчётов с изменением титульных данных, критериев отбора и т.п., которые запутывают ситуацию и могут сделать её трудно управляемой.

Чтобы этого не происходило, необходим жесткий технологический порядок при обработке и программные средства поддержки, обеспечивающие регистрацию всех изменений и их хронологии в специальной базе данных. Всё это возможно реализовать в рамках общей концепции автоматизированного рабочего места физика [11]. Пример частичной реализации данной концепции на ЕС ЭВМ показан на рис.4.

Возникает также другая проблема, связанная с уже накопленным и пополняемым статистическим материалом на лентах суммарных результатов (ЛСР). Необходимо решить задачу выборки, сравнения, классификации и объединения накопленных результатов различных экспозиций и экспериментов на 100-сантиметровой водородной пузырьковой камере. Для этого необходимы программные средства, включая базы данных, обеспечивающие регистрацию, поиск и выборку событий по определяемым физиками критериям.

В заключение автор выражает искреннюю благодарность В.Г.Иванову, В.В.Глаголеву, С.С.Шиманскому, И.Э.Шокирову за содействие в выполнении данной работы, полезные обсуждения и ценные замечания.

Литература

1. Балгансурен Я. и др. ОИЯИ, P10-86-706, Дубна, 1986.
2. CDC NOS/BE 1 Reference Manual, Pub.N:860493800. Control Data Corporation, St.Paul, Minnesota, 3-15-77.
3. Буздавина Н.А. и др. ОИЯИ, P10-91-81, Дубна, 1991.
4. Глаголев В.В. и др. ОИЯИ, P10-91-80, Дубна, 1991.
5. Бадалян С.Г. и др. ОИЯИ, P10-12474, Дубна, 1979.
6. Балгансурен Я. и др. ОИЯИ, P10-86-412, Дубна, 1986.
7. Балгансурен Я. и др. ОИЯИ, P10-86-806, Дубна, 1986.
8. HYDRA System Manual, CERN, Geneva, 1979.
9. VM/SP : System Product Interpreter Reference, SC24-5239.
10. Балгансурен Я. и др. ОИЯИ, P10-87-242, Дубна, 1987.
11. Иванов В.Г. и др. ОИЯИ, P10-91-105, Дубна, 1991.

Рукопись поступила в издательский отдел  
14 апреля 1993 года.