

Е. М. Балдин<sup>1</sup>, В. Е. Блинов<sup>1</sup>, А. С. Зайцев<sup>1</sup>, А. А. Король<sup>1</sup>, И. Б. Логашенко<sup>1,2</sup>,  
А. М. Сухарев<sup>1</sup>, А. А. Талышев<sup>1</sup>, Ю. А. Тихонов<sup>1</sup>

<sup>1</sup> Институт ядерной физики им. Г. И. Будкера СО РАН  
пр. Акад. Лаврентьева, 11, Новосибирск, 630090, Россия

<sup>2</sup> Новосибирский государственный университет  
ул. Пирогова, 2, Новосибирск, 630090, Россия

E-mail: ivan.logashenko@gmail.com

## ПРОЕКТ СУПЕР *ст*-ФАБРИКИ: ТРЕБОВАНИЯ К ВЫЧИСЛИТЕЛЬНОЙ ИНФРАСТРУКТУРЕ \*

В Новосибирске на базе Института ядерной физики им. Г. И. Будкера (ИЯФ) СО РАН разрабатывается проект Супер *ст*-фабрики. В статье изложены базовые требования и основные параметры, на которые следует ориентироваться при планировании и развертывании вычислительной инфраструктуры комплекса.

*Ключевые слова:* физика высоких энергий, детектор элементарных частиц, система обработки данных, система хранения данных.

### Введение

Супер *ст*-фабрика (далее ЦТФ) [1] – это электрон-позитронный коллайдер с энергией в системе центра масс от 3 до 5–6 ГэВ, имеющий высокую светимость (около  $10^{35} \text{ см}^{-2} \text{ с}^{-1}$ ). В этом диапазоне энергии рождаются практически все состояния чармониев (мезонов, представляющих собой связанные состояния *c*-кварка и антикварка), очарованных мезонов и барионов, содержащих *c*-кварк или антикварк. Кроме того, при энергии больше  $2m_\tau \approx 3,6$  ГэВ начинают рождаться пары  $\tau$ -лептонов. Такой коллайдер при наличии высокой светимости будет интенсивным источником очарованных частиц и  $\tau$ -лептонов. Схема расположения ЦТФ на территории ИЯФ представлена на рис. 1.

Основная цель экспериментов на супер *ст*-фабрике – изучение процессов рождения очарованных кварков (*c*-кварков) и  $\tau$ -лептонов с набором данных, который на 3–4 порядка больше по объему, чем использо-

ванный до сих пор в любом другом эксперименте. Предполагается, что эти эксперименты будут чувствительны к новым физическим эффектам, не описываемым Стандартной Моделью.

Для выполнения этой задачи необходим универсальный магнитный детектор с отличным разрешением по импульсу заряженных частиц, хорошим разрешением по энергии фотонов и системой идентификации частиц с рекордными параметрами. Считывающая электроника должна обеспечивать скорость, достаточную для считывания всех изучаемых событий, а вычислительная инфраструктура должна предоставить возможность обработать эти данные со скоростью не ниже средней скорости набора.

### Общие требования

Эффективность функционирования интегральной системы обработки, хранения и распределенного анализа данных в первую

---

\* Работа выполнена при поддержке Министерства образования и науки РФ.



Рис. 1. Схема расположения ЦТФ на территории ИЯФ СО РАН

очередь определяется способностью системы обеспечить успешное проведение исследований на научной мегаустановке. Исходя из этого можно выделить следующие основные требования к системе:

- возможность сохранения всего выходного потока данных, идущего с системы сбора данных;
- возможность выполнения процесса физического моделирования экспериментальных событий в детекторе в необходимом объеме;
- сохранность и постоянная доступность данных, размещенных в системах хранения, на длительных (15–20 лет) промежутках времени;
- высокая степень отказоустойчивости системы сбора и системы обработки данных;
- доступ удаленных исследовательских групп к системе обработки данных и к массиву набранных экспериментальных данных;
- возможность привлечения к процессу обработки данных вычислительных ресурсов географически удаленных центров обработки данных и суперкомпьютерных центров.

Во время работы эксперимента данные поступают с детектора частиц на оцифровывающую электронику по сигналу специальной электронной схемы (аппаратный триггер). После оцифровки данные подаются на специализированную компьютерную ферму

для частичной или полной реконструкции экспериментальных событий и программного отбора (программный триггер). На основании решения программного триггера данные записываются в долговременную память.

В табл. 1 представлены оценки сечения ( $\sigma$ ) и загрузки ( $f$ ) детектора полезными событиями при различных энергиях столкновения электрона и позитрона, а также оценки загрузки от фоновых процессов.

Из-за высокой светимости ЦТФ загрузка от космических частиц (Космика) не является существенной. Сечение нерезонансного рождения адронов (Адроны) составляет примерно  $\sim 20$  нБ, что соответствует скорости счета 20 кГц при светимости  $10^{35} \text{ см}^{-2} \text{ с}^{-1}$ . Эти события нельзя отделить от событий распада векторных мезонов. Основным фоновым физическим процессом являются события  $e^+e^-$ -рассеяния (Bhabha).

В настоящий момент сложно оценить частоту срабатывания триггера от фоновых событий, связанных с потерями пучка и взаимодействием пучка с остаточным газом. Опыт работы на В-фабриках в КЕК [2] и SLAC [3] и текущие разработки SuperB-фабрики [4] в Италии указывают, что частота срабатывания будет на уровне нескольких десятков килогерц. Эти фоновые события, как правило, являются нецентральными и могут быть подавлены на этапе обработки.

Таблица 1

## Сечение и частота записи событий

Процессы	$J/\psi$	$\psi(2S)$	$\tau^+\tau^-$	$\psi(3770)$	$\tau^+\tau^-$	$\Lambda_c\Lambda_c$
Энергия, МэВ	3 097	3 686	3 700	3 770	4 250	4 650
$\sigma$ , нБ	$\sim 1\,450$	$\sim 400$	2,5	$\sim 6$	3,5	0,5
$f$ , кГц	140	40	2,5	6	3,5	0,5
Фоновые процессы						
Космика, кГц	$\sim 2$					
Адроны, кГц	19	17	17	16	14	12
Vhabha, кГц	90	80	80	80	60	50

В пике  $J/\psi$ -мезона (принимая энергетический разброс в месте встречи  $\sim 2,1$  МэВ) загрузка полезными событиями будет максимальной. С учетом пучкового фона из табл. 1 видно, что пиковая загрузка, с которой должна справляться интегрированная система обработки, хранения и анализа данных, должна быть на уровне не менее 300 кГц.

Невозможно предсказать заранее, в каком из классов физических событий может проявиться «новая физика». Поэтому необходимо сохранять 100 % полезных событий. Однако для уменьшения объема сохраняемой на долговременных носителях информации можно пожертвовать хорошо изученными и легко выделяемыми событиями  $e^+e^-$  рассеяния. Число этих событий перед записью для длительного хранения можно подавить в  $10\div 20$  раз. Это позволит уменьшить объемы долговременного хранения в  $5\div 10$  раз для набора событий вне пиков  $J/\psi$  и  $\psi(2S)$  мезонов.

В основное время набора ( $\sim 95$  % времени вне пиков узких резонансов) частота записи на долговременные носители не будет превышать  $20\div 30$  кГц.

Необходимое количество распадов  $J/\psi$ -мезона,  $10^{12}$  событий, в силу высокого сечения этого узкого резонанса при проектной светимости ЦТФ будет получено в течение трех месяцев, что в рамках десятилетней программы составляет около 3 % от общего времени. В процессе набора статистики в пике  $J/\psi$  необходимо не только считывать, но и записывать экспериментальные события с частотой 300 кГц. В пике  $\psi(2S)$ -мезона загрузка экспериментальными событиями для записи не должна превышать  $\sim 70$  кГц.

**Обработка данных, размер события и объем вычислений.**

Обработка данных после их набора во время эксперимента (offline обработка) включает в себя калибровку (перевод оцифрованных данных с детектора в физические измерения), реконструкцию событий (построение на основе этих данных физических объектов) и анализ данных. Кроме экспериментальных данных для анализа необходим расчет параметров эксперимента и различных физических моделей. В основном, для этого используется метод стохастического моделирования (моделирование Монте Карло). Для того чтобы систематическая неопределенность физических измерений была достаточно маленькой, число событий моделирования должно быть в 2–3 раза больше, чем число записанных экспериментальных событий.

На вход системы обработки данных поступают так называемые сырые данные (RAW), т. е. информация, содержащая коды оцифровывающей электроники – номера каналов, коды амплитуд и т. д. Сырые данные представляют собой целые числа и могут иметь различную разрядность. Исходя из возможностей современной электроники и структуры детектора [1] размер этой части можно оценить в среднем как 30 кБ/событие.

После калибровки и реконструкции к событию добавляются дополнительные данные (ESD, DST)<sup>1</sup>, а именно измерения в фи-

<sup>1</sup> См.: ATLAS Collaboration. ATLAS Computing Technical Design Report // CERN-LHCC-2005-022. ISBN 92-9083-250-9.

зических единицах (см, МэВ, нс), результаты работы алгоритмов распознавания образов (параметры треков, кластеров, колец RICH, вершин), физические объекты (4-импульсы и типы частиц, потерянный импульс и т. д.). Зная структуру детектора и условия проведения эксперимента (средняя адронная множественность примерно 2 нейтральных и 5 заряженных  $\pi$  мезонов), можно ожидать, что средний размер этих данных будет около 13 кБ/событие.

Для физического анализа необходима только часть данных ESD, включающая физические объекты и информацию о событии в целом (AOD, mDST). Объем этих данных составляет примерно 1,2 кБ/событие. Эти данные обычно находятся в оперативном доступе.

Данные моделирования в основном аналогичны экспериментальным (RAW, ESD, AOD), но с добавлением информации об условиях проведения моделирования (SIM), т. е. 4-импульсы начальных частиц и т. д. Средний размер дополнительной части события составляет примерно 0,3 кБ/событие.

Значительную часть работы процедур реконструкции составляют распознавание образов и получение параметров полученных объектов. Основной вклад в потребление процессорного времени вносят алгоритмы реконструкции трековой системы и системы идентификации (см., например: [5; 6]). Опыт проведения экспериментов на аналогичных по сложности детекторах BABAR [7], KEDR [8] дает (в единицах Linpack флопс [9]) оценку около  $0,15 \text{ Гфлопс} \frac{\text{с}}{\text{событие}}$ .

Ожидаемое число запусков процедуры реконструкции одних и тех же событий составит 2–3 раза в год. Таким образом, потребность в вычислительной мощности для реконструкции экспериментальных данных и моделирования (для средней скорости набора  $\sim 100$  кГц) можно оценить как

$$(2 + 2 \cdot 2) \cdot 0,15 \cdot \text{Гфлопс} \frac{\text{с}}{\text{событие}} \cdot 10^5 \text{ Гц} = 0,09 \text{ Пфлопс}.$$

Для физической программы планируемого эксперимента средняя вычислительная емкость моделирования событий (без реконструкции) составит примерно

$$2 \text{ Гфлопс} \frac{\text{с}}{\text{событие}}$$

(см., например, [10]). Таким образом, вычислительную емкость моделирования можно оценить как

$$2 \cdot 2 \cdot \text{Гфлопс} \frac{\text{с}}{\text{событие}} \cdot 10^5 \text{ Гц} = 0,4 \text{ Пфлопс}.$$

По данным реконструкции и моделирования проводится физический анализ. Потребность в вычислительных ресурсах оценить достаточно сложно из-за уникальности обработки и наличия множества параллельных исследовательских групп. Можно, однако, предполагать, что вычислительные потребности собственно анализа не превышают потребностей реконструкции. Таким образом, общие вычислительные потребности проекта составят примерно 0,6 Пфлопс. Оценка может измениться при внедрении новых вычислительных технологий (графические процессоры, программируемые схемы и др.), но вряд ли значительно.

#### Оценка потоков и объемов

Пиковая частота регистрации полезных событий при наборе статистики при энергиях в районе пика  $J/\psi$ -мезона составляет 140 кГц. С учетом наличия событий фона и необходимости наличия определенного запаса производительности системы требуется, чтобы система обработки данных успевала воспринимать события, поступающие с частотой 300 кГц. Учитывая, что средняя длина события оценивается в 30 кБ, максимальный поток данных на входе системы обработки равен примерно 10 ГБ/с. Средний поток данных на входе системы при наборе статистики при других энергиях на порядок меньше и не превышает 1 ГБ/с. Общая схема потоков данных представлена на рис. 2.

Физическая программа экспериментов рассчитана на 10 лет. За это время предполагается набрать порядка  $10^{12}$  событий распадов  $J/\psi$ -мезона (в течение нескольких месяцев) и порядка  $4 \cdot 10^{12}$  событий при энергиях за пределами пика  $J/\psi$ -мезона. Общий объем оцифрованных сырых данных за время работы комплекса составит около 150 ПБ.

Общий объем реконструированных данных (AOD), используемых при анализе, за время работы комплекса оценивается в 10 ПБ. Приблизительно столько же места требуется для данных моделирования.

Полный объем реконструированных данных (ESD) оценивается в 65 ПБ, однако

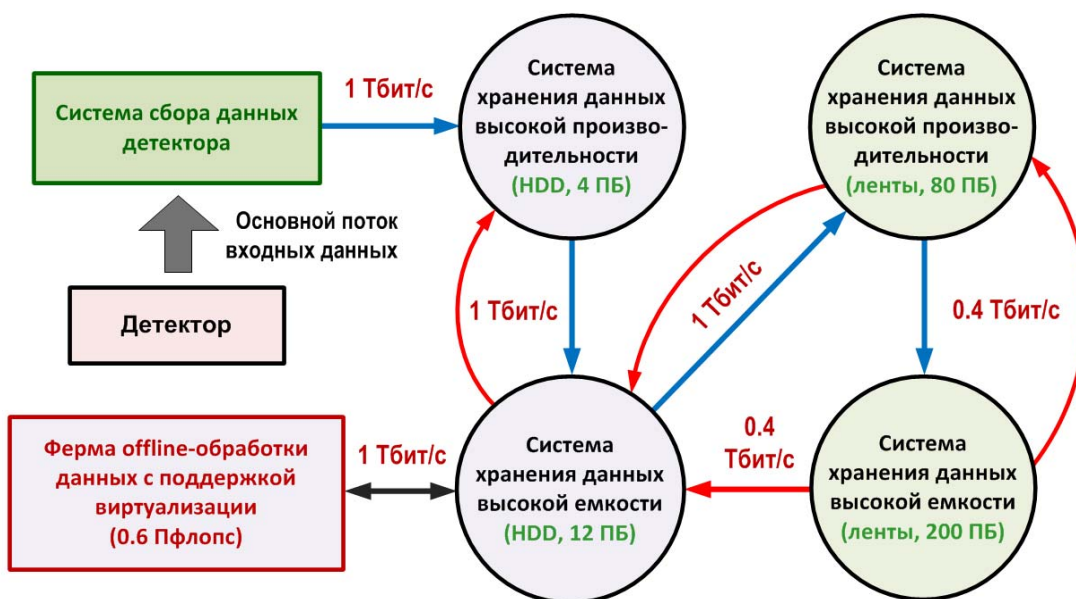


Рис. 2. Схема потоков данных

Таблица 2

Сравнение потоков данных установок с высокой светимостью

Эксперимент	Длина события (кБ)	После аппаратного триггера (ГБ/с)	После программного триггера (ГБ/с)
BABAR	35	0,02	0,01
Belle	40	0,02	0,01
SuperB	35	~ 4	~ 0,9
Super Belle	200÷300	> 2	< 0,25
ЦТФ (кроме $J/\psi$ )	30	~ 3	0,6÷0,9
ЦТФ (пик $J/\psi$ )	30	~ 10	~ 7

не требуется хранить эти данные за все время работы комплекса. Как правило, оперативный доступ к такой информации необходим только для недавно набранных данных, на время проведения калибровок и отладки программного обеспечения и реконструкции. В дальнейшем эта часть данных может быть удалена из хранилища и восстановлена по мере надобности с использованием ПО реконструкции.

Таким образом, для хранения и анализа данных экспериментов на Супер-ст-фабрике необходимо централизованное хранилище данных общим объемом порядка 200 ПБ, способное воспринимать входной поток данных до 10 ГБ/с. Такое хранилище должно располагаться в географической близости к самому комплексу.

В случае присоединения к обработке сторонних организаций и международных коллабораций со своими вычислительными мощностями возможно хранение копии части или всего объема реконструированных данных и данных моделирования непосредственно на местах обработки. Интегрированная система хранения и обработки данных должна обеспечивать сервисы синхронизации данных с такими удаленными центрами.

### Сравнение требований

Прямых аналогов работающих на такой же энергии установок со сравнимой светимостью в мире нет. Имеет смысл сравнивать требования, выдвигаемые к интегрирован-

ным системам обработки и хранения данных на ЦТФ с *B*-фабриками предыдущего поколения – Belle (Япония) и BABAR [3] (США), а также с *B*-фабриками, которые планируют построить – Super*B* (Италия) [4] и Super Belle [11] (Япония). Для сравнения в табл. 2 приведены данные по объемам потоков обсуждаемых ускорительных комплексов. Из приведенных данных видно, что при наборе статистики вне узких резонансов (особенно, вне пика  $J/\psi$ ) потоки данных на ЦФП приблизительно соответствуют потокам данных на *B*-фабриках нового поколения. Однако во время работы на энергии  $J/\psi$  мезона потоки данных на ЦФТ вырастают на порядок, поэтому вычислительная инфраструктура ЦФТ должна быть способна обрабатывать значительно большие потоки данных, чем аналогичные системы *B*-фабрик нового поколения.

### Заключение

Физическая программа Супер ст-фабрики рассчитана на 10 лет работы. Общий объем данных, накопленных за время работы комплекса, оценивается в 200 ПБ. Объем реконструированных данных (AOD), используемых при анализе и находящихся в оперативном доступе, оценивается в 10 ПБ. Максимальный ожидаемый входной поток данных составляет 10 ГБ/с. Для реконструкции событий в реальном времени необходимо развернуть вычислительную компьютерную ферму мощностью не менее 0,6 Пфлопс.

### Список литературы

1. Анашин В. В. и др. Супер Charm-Tau фабрика. Концептуальный проект. Новосибирск, 2011.

2. Kurokawa S. Overview of the KEKB Accelerators // Nucl. Instrum. Meth. A499. 2003. P. 1–7.

3. Aubert B. et al. The BaBar Detector // Nucl. Instrum. Meth. A479. 2002. P. 1–116.

4. Bona M. et al. SuperB: A High-Luminosity Asymmetric  $e^+e^-$  Super Flavor Factory. Conceptual Design Report. 2007.

5. Mankel R. Pattern Recognition and Event Reconstruction in Particle Physics Experiments // Rep. Prog. Phys. 2004. Vol. 67. P. 553.

6. Buszello C. P. LHCb RICH Pattern Recognition and Particle Identification Performance // Nucl. Instr. Meth. A 595. 2008. P. 245–247.

7. Pavel T. J. Choosing CPUs in an Open Market: System Performance Testing for the BABAR Online Farm, International Conference on Computing in High-Energy Physics (CHEP98). Chicago, 1998.

8. Anashin V. V. et al. Status of the KEDR Detector // Nucl. Instr. Meth. A478. 2002. P. 420–425.

9. Dongarra J., Luszczek P., Petitet A. The LINPACK Benchmark: Past, Present and Future // Concurrency and Computation: Practice and Experience. 2003. Vol. 15. P. 803–820.

10. Bozzi C. Simulation in Babar, Desy / ECFA workshop. Cern, 2001.

11. Abe T., Adachi I., Adamczyk K., Ahn S., Aihara H., Akai K., Alois M., Andricsek L., Aoki K., Arai Y. et al. Belle II Technical Design Report. ArXiv e-prints, November 2010.

Материал поступил в редколлегию 31.10.2011

E. M. Baldin, V. E. Blinov, A. S. Zaitsev, A. A. Korol, I. B. Logashenko,  
A. M. Sukharev, A. A. Talyshev, Yu. A. Tikhonov

### SUPER CHARM-TAU FACTORY: COMPUTING REQUIREMENTS

Super Charm-Tau Factory (CTF), the project of  $e^+e^-$  collider with luminosity of  $10^{35} \text{ 1 cm}^{-2}/\text{sec}^{-1}$ , is being developed in the Budker Institute of Nuclear Physics (Novosibirsk, Russia). The main requirements and parameters for the computing infrastructure of the future facility are discussed.

*Keywords:* high energy physics, particle detector, data analysis system, data storage system.