

P10-2004-208

Д. А. Белослудцев, В. Е. Жильцов, А. И. Зинченко,  
В. Д. Кекелидзе, Д. Т. Мадигожин, Ю. К. Потребеников,  
С. В. Хабаров, С. Н. Шкаровский, Б. Г. Щинов

КОМПЬЮТИНГ ДЛЯ ДЕЙСТВУЮЩИХ  
ЭКСПЕРИМЕНТОВ ПО ФИЗИКЕ ЧАСТИЦ  
В ЛФЧ ОИЯИ

Направлено в журнал «Письма в ЭЧАЯ»

Белослудцев Д. А. и др.  
Компьютинг для действующих экспериментов  
по физике частиц в ЛФЧ ОИЯИ

P10-2004-208

Описана созданная в ЛФЧ ОИЯИ компьютерная инфраструктура, нацеленная на обеспечение эффективного решения задач проводимых с участием специалистов из ОИЯИ действующих экспериментов по физике частиц и ядерной физике. Изложены принципы построения компьютерной фермы персональных машин, описаны используемые компьютерные и информационные сервисы, обеспечивающие эффективное использование распределенных компьютерных и информационных ресурсов.

Работа выполнена в Лаборатории физики частиц ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна, 2004

Belosludtsev D. A. et. al.  
Computing for Ongoing Experiments on High Energy  
Physics in LPP, JINR

P10-2004-208

The computer infrastructure made at the Laboratory of Particle Physics, JINR, purposed for active participation of JINR experts in ongoing experiments on particle and nuclear physics is presented. The principles of design and construction of the personal computer farm have been given and the used computer and informational services for effective application of distributed computer resources have been described.

The investigation has been performed at the Laboratory of Particle Physics, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna, 2004

## ВВЕДЕНИЕ

Реализация современных экспериментальных проектов по физике частиц предполагает участие в них большого количества физиков и специалистов, получение и обработку огромных объемов экспериментальных данных, высокую стоимость используемого оборудования. Такие условия объективно способствуют появлению крупных международных коллабораций, способных общими усилиями обеспечить выполнение задуманной экспериментальной программы на самом современном технологическом и научном уровнях. Как следствие, именно здесь острее, чем на большинстве других направлений науки и техники, ощущается недостаток в информационно-вычислительных ресурсах и именно здесь находится передний край применения информационно-компьютерных технологий. В ряде случаев физика частиц даже определяет прогресс в области мирового компьютеринга в целом. Примером тому может служить всемирная паутина (www), идея создания которой родилась в Европейском центре ядерных исследований (ЦЕРН). Еще один пример - интенсивное развитие возможностей новой информационной технологии GRID [1] – мощной сети широкодоступных распределенных информационно-вычислительных ресурсов, способной в недалеком будущем значительно изменить сложившиеся в мире реалии.

Действующие эксперименты по физике частиц, такие как, например, NA-48 и COMPASS в ЦЕРН, STAR в Брукхейвенской национальной лаборатории (США), ВаВаг в Стенфордском центре линейных ускорителей (SLAC, США), достигли объемов обрабатываемых данных, измеряемых сотнями терабайт в год. Уже сейчас организован анализ полученных данных, а также моделирование условий экспериментов в распределенной информационно-вычислительной среде участвующих в этих экспериментах организаций. С запуском же в 2007 г. нового коллайдера LHC в ЦЕРН предполагается существенно, возможно на порядок, превзойти эти объемы. Все это определяет необходимость уже сейчас создавать адекватную информационно-вычислительную инфраструктуру в тех научных центрах, специалисты которых занимаются или планируют заняться анализом накопленных и будущих экспериментальных данных.

Упомянутая выше концепция GRID сочетает в себе два аспекта: вычислительный (computational grid) и информационный (data grid). Предполагается, что для ее использования будут созданы мощные информационно-вычислительные комплексы различного уровня, связанные между собой компьютерными сетями сверхвысокой по современным меркам пропускной спо-

способности. Применение таких решений позволит изменить модели вычислений, обеспечив переход от вычислений на разрозненных компьютерных комплексах к использованию географически распределенных систем. Этот технологический сдвиг способен оказать такое же сильное воздействие на общество, как и Интернет.

Наиболее детально схема организации такого типа систем проработана в физических проектах на LHC [2], для реализации которых разработан и успешно осуществляется проект LCG [3]. Для экспериментов на LHC предложена иерархическая структура вычислительных центров, включающая вычислительные центры пяти ярусов, каждый со своим объемом ресурсов и различными возможностями сетевого доступа. Нулевой ярус — главный центр в ЦЕРН, первый ярус — крупные национальные центры ведущих стран мира, второй ярус — региональные центры, а также менее крупные центры третьего и четвертого ярусов. Хотя критерии для региональных центров второго яруса еще не зафиксированы, в их число войдут географическая распределенность, широкий доступ к высокоскоростным сетям, а также наличие вычислительных мощностей в диапазоне от 10 до 50 K SpecInt95 (SI95)\*). Практическая реализация таких распределенных центров включает в себя, как минимум, создание вычислительной структуры соответствующей производительности, адаптацию программного обеспечения промежуточного уровня, организацию работы с данными, адаптацию систем разработки и отладки программ для конкретных научных приложений. Центр такого уровня с элементами первого яруса уже создается и в ОИЯИ.

Существенные изменения намечаются и в области организации экспериментов по физике частиц. По финансовым причинам все меньше экспериментальных установок могут быть созданы локальными группами физиков и даже отдельными странами. Таким образом, эти установки, а также получаемые с их помощью данные все в большей степени становятся международной собственностью. В системы управления экспериментом все шире внедряется дистанционное управление, распространяемое даже на работу экспертов в ходе проведения сеансов экспозиции установок (удаленный контроль работы отдельных детекторов и эксперимента в целом, управление работой детекторов и систем и т.п.). Распределенные компьютерные ресурсы, как уже отмечено выше, становятся единственным средством проведения анализа полученной информации. В этой связи определяющее значение для участия в проведении физического эксперимента группой специалистов имеет наличие в их распоряжении развитой информационно-вычислительной структуры, ко-

---

\*) SpecInt95 - широко распространенная единица оценки производительности центральных процессоров компьютеров, основанная на выполнении специального набора тестирующих программ.

торая становится фактически основным инструментом экспериментального подразделения, обеспечивающим потенциальную возможность получения новых физических результатов.

Данная работа посвящена вопросам развития компьютеринга в ЛФЧ ОИЯИ, ориентированного на проведение работ в рамках действующих экспериментов по физике частиц и ядерной физике и нацеленного на построение к 2006 г. на площадке ЛВЭ достаточно мощного компьютерного кластера, интегрированного в распределенную компьютерную инфраструктуру ОИЯИ.

## ПРЕДЫСТОРИЯ ВОПРОСА

Первым шагом к развитию компьютерной инфраструктуры Лаборатории физики частиц, нацеленной на решение реальных задач действующих экспериментов по физике частиц и ядерной физике, явился перевод локальной сети лаборатории на новую технологию и новое сетевое оборудование, выполненный в 1999 г. Приобретенное и использованное оборудование фирмы Hewlett Packard (HP-2524 switch) обеспечило надежное функционирование локальной сети на скорости до 100 Мбит в секунду и создало возможность перехода на скорости до 1 Гбит в секунду.

Параллельно с этим были начаты работы по созданию централизованного вычислительного межлабораторного ресурса — компьютерной фермы персональных машин ЛФЧ–ЛВЭ. Основной задачей фермы стало обеспечение возможности проведения работ по анализу экспериментальных данных действующих экспериментов по физике частиц и ядерной физике, по моделированию существующих и подготавливаемых экспериментов в этих областях исследований. Компьютерная ферма, заработавшая уже в 1999 г. [4], имела следующие параметры: производительность — 1,2 К SI95 (40 CPU); дисковая память — 3 Тбайт; ленточная подсистема — 0,5 Тбайт (три устройства типа DLT, в том числе — один робот на девять лент). Базовыми компьютерами фермы явились двухпроцессорные персональные машины Pentium-II, частота процессоров которых росла в процессе их приобретения от 400 МГц до 1 ГГц. Управление всеми машинами фермы осуществлялось с одной операторской консоли (монитор + клавиатура + «мышь») за счет использования специализированных переключателей CPU. Программное обеспечение фермы в целом обеспечивало работу пользователей в принятой в ведущих физических центрах мира среде. В частности, на ферме впервые в ОИЯИ была адаптирована и внедрена распределенная файловая система AFS (Andrew File System), используемая всеми крупными физическими центрами. С помощью этой системы сейчас, например, ежедневно автоматически обновляется адаптированное к особенностям фермы программное обеспечение эксперимента NA-48. Были инсталлированы и поддерживались на современном уровне про-

граммные комплексы, системы и пакеты как общего назначения (SGE batch system, CERNLIB, ROOT и т. п.), так и ориентированные на задачи каждого отдельного эксперимента (ROOT4STAR и STAF — для эксперимента STAR в BNL, CORAL, PHAST и COMGEANT — для эксперимента COMPASS в ЦЕРН, NASIM, СМС и COMPACT — для экспериментов NA-48 в ЦЕРН, SIMEX — для эксперимента ЭКСЧАРМ на серпуховском ускорителе У-70 и т. п.).

Появление такой фермы, на то время первой в ОИЯИ и одной из самых мощных в России, позволило начать решать реальные задачи действующих и подготавливаемых экспериментов. Так, например, в 2000 г. на ферме ЛФЧ–ЛВЭ в течение трех месяцев была решена задача специализированного моделирования с наложением экспериментальных событий от случайного срабатывания триггера установки, результаты которой были использованы для оценки влияния случайных совпадений во времени на основной результат эксперимента NA-48. Для этого потребовалась практически непрерывная более чем 12-часовая работа двух устройств DLT-7000 (чтение 190 магнитных лент, 10,8 Тбайт, с экспериментальной информацией, записанной по случайному срабатыванию триггера эксперимента), работа восьми параллельных задач на пакетных машинах кластера и 500 Гбайт дискового пространства для хранения исходных данных и результатов моделирования. В это же время были начаты работы по реальному моделированию адрон-адронных и ядро-ядерных взаимодействий в эксперименте STAR, а несколько позже — и в эксперименте COMPASS. К 2004 г. в качестве пользователей компьютерной фермы было зарегистрировано 57 человек из четырех лабораторий ОИЯИ — ЛФЧ, ЛВЭ, ЛИТ и ЛЯП. Интерактивные машины для работы на кластере имели исполнители проектов ATLAS, CMS, COMPASS, EXCHARM, HERA-B, H1, NA-48, STAR.

Еще одним направлением развития компьютерной инфраструктуры лаборатории в то время стало обеспечение возможности использования современных сервисов для общения ученых в процессе проведения исследовательских работ. В качестве первого шага в этом направлении был выбран сертифицированный сервис в области физики частиц — система Virtual Room Video System (VRVS) [5] — под использование которого был оборудован специальный конференц-зал, приобретена современная интегрированная аппаратура фирмы Polycom — система Polycom H323 и необходимый набор сервисных устройств для управления, видео- и звукового обеспечения видеоконференций. Уже в 2001 г. созданный видеоконференционный комплекс был введен в опытную эксплуатацию [6]. Это позволило экспериментаторам в режиме видеообщения участвовать в работе основных совещаний международных коллабораций не только в качестве наблюдателей, но и в качестве докладчиков с удаленным представлением своих материалов в виде компьютерных презентаций приемлемого на то время качества.

Несмотря на то, что появилась принципиальная возможность проведения отдельных экспериментальных работ, имеющиеся в то время ресурсы были крайне ограничены. Например, решение одной задачи моделирования событий исследуемой конфигурации с их последующей реконструкцией в эксперименте COMPASS требовало около 15% всех годовых ресурсов CPU-фермы, а проведение моделирования 10% событий для эксперимента NA-48 — 25% этих ресурсов. Работа же аппаратуры для проведения видеоконференций на начальном этапе ее эксплуатации не могла считаться эффективной из-за малой пропускной способности внешних каналов связи Института и недоработок в сервисном обеспечении этого нового вида коммуникаций.

### ПРОЕКТ «Ф-КЛАСТЕР»

Накопленный положительный опыт работ по развитию компьютерной инфраструктуры лаборатории, с одной стороны, и резкое несоответствие растущим требованиям действующих экспериментов, с другой, потребовали более системного подхода к дальнейшим работам в этой области. Сделанные оценки потребностей работающих экспериментальных групп позволили сформулировать следующие задачи развития компьютеринга в лаборатории на три года начиная с 2004-го:

1) увеличить вычислительные возможности компьютерного кластера ЛФЧ–ЛВЭ за три года не менее чем в четыре раза – до 5 K SI95, объем дискового пространства, доступного пользователям, – не менее чем в пять раз – до 15 Тбайт, обеспечить переход на новую технологию обмена данными между компьютерами кластера, позволяющую достигнуть скоростей передачи данных не менее 1 Гбит в секунду;

2) поддерживать на современном уровне, достаточном для обеспечения действующих экспериментов, компьютерное, программное (в том числе — ориентированное на конкретный эксперимент) и информационное обеспечение кластера и рабочих мест пользователей для выполнения физических и методических программ лаборатории;

3) обеспечить на компьютерном кластере ЛФЧ–ЛВЭ возможность внедрения технологии GRID в экспериментах по физике частиц и ядерной физике;

4) создать систему обеспечения сотрудников лаборатории высокоскоростным доступом в локальную сеть ОИЯИ, в том числе и беспроводным;

5) реализовать в лаборатории высококачественную систему поддержки видеоконференционной (VRVS, H323) связи.

Компьютерный кластер с такими возможностями мог бы позволить лаборатории считаться достойным партнером в проведении ряда действующих экспериментов, прежде всего — экспериментов COMPASS и NA-48 в ЦЕРН.

Для придания ведущимся работам должного статуса и обеспечения их хотя бы минимальным финансированием в лаборатории был подготовлен специальный проект развития компьютерной инфраструктуры «Кластер распределенной компьютерной инфраструктуры ОИЯИ для действующих экспериментов по физике частиц (Ф-кластер)», получивший одобрение ПКК по физике частиц и Ученого совета ОИЯИ в 2004 г. и принятый в Институте для реализации с первым приоритетом.

Важно отметить, что ресурсы создаваемого в лаборатории кластера должны были быть интегрированы в общую компьютерную инфраструктуру ОИЯИ и доступны не только сотрудникам ОИЯИ, но и всем участникам действующих экспериментов из стран-участниц ОИЯИ независимо от их территориального расположения.

### **РАЗВИТИЕ КОМПЬЮТЕРНОГО КЛАСТЕРА ЛФЧ В 2003–2004 гг.**

За полтора последних года в развитии компьютерного кластера лаборатории достигнуты значительные успехи. В результате развития компьютерной фермы ЛФЧ–ЛВЭ число используемых в ней компьютеров возросло с 20 до 32. Более чем втрое, до 3,8 К SI95, повышена суммарная производительность их процессоров, в том числе с 0,7 до 2,4 К SI95 — для batch-компьютеров — компьютеров, предназначенных только для пакетной обработки заданий пользователей. Более чем вдвое, с 3 до 6,5 Тбайт, увеличена емкость дисковой памяти, доступной для пользователей. Высвобожденные двухпроцессорные компьютеры низкой производительности (с частотой процессоров 400 МГц) переданы для использования в ЛФЧ на рабочих местах. На компьютерах вычислительной фермы ЛФЧ–ЛВЭ установлена система для работы с большими объемами данных CASTOR [7]; предложен способ и проведено тестирование системы как средства интеграции вычислительных ресурсов ОИЯИ при решении задач обработки и анализа экспериментальных данных.

Осуществлен перевод локальной компьютерной сети лаборатории на использование оборудования Gigabit Ethernet. С использованием устройства Cisco Catalyst 3550 локальной компьютерной сети ОИЯИ организованы четыре локальные подсети лаборатории. Увеличено количество конечных пользователей, получивших выход в Интернет через локальные подсети ЛФЧ, в том числе в корпусах лаборатории, не имевших до этого подключения к сети.

Установлено и протестировано оборудование, обеспечивающее беспроводный выход отдельных мобильных пользователей в локальную сеть ОИЯИ. Дооборудованы системой беспроводного выхода в Интернет залы для проведения видеоконференций и семинаров, которые, кроме того, оснащены полным набором необходимых проекторов. Эти изменения привели к тому, что видеоконференции в лаборатории стали нормой при проведении совещаний



таких коллабораций, как NA-48 и COMPASS. Например, в 2003 г. из 67 докладов сотрудников ЛФЧ, сделанных на совещаниях коллаборации NA-48, 13 были представлены из Дубны с использованием этого сервиса, а за первую половину 2004 г. таких докладов уже сделано более десяти.

Схематически создаваемый кластер изображен на рисунке. Основные составляющие кластера, расположенные в различных корпусах Института, объединены в лабораторные подсети оптическими линиями связи и соединены с центральным узлом опорной сети ОИЯИ через коммутатор Cisco Catalyst 3550

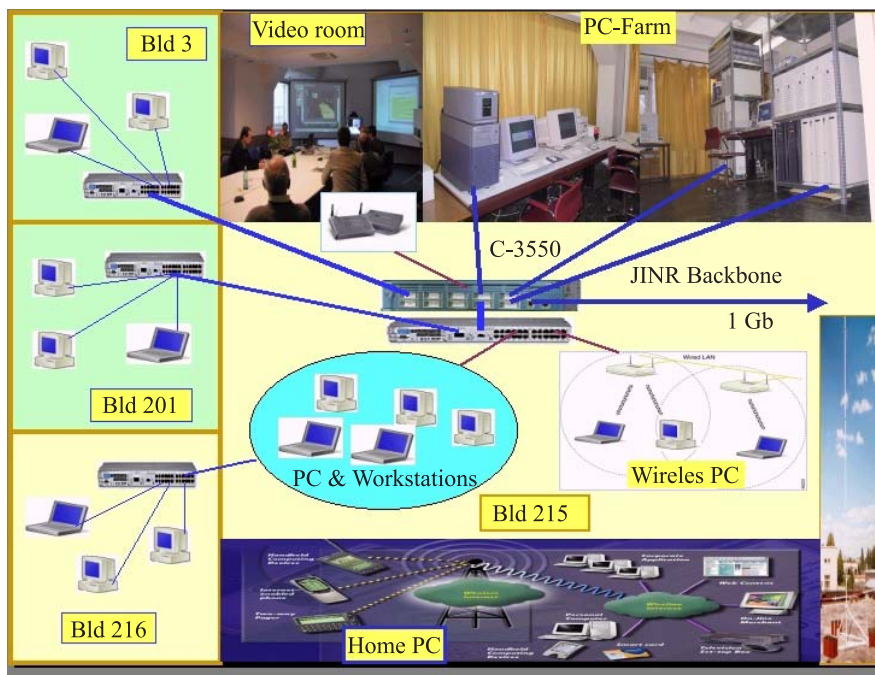


Схема Ф-кластера

каналом с пропускной способностью 1 Гбит в секунду. Доступ к центральным ресурсам кластера и внешним ресурсам с рабочих мест пользователей обеспечен как по проводным, так и по беспроводным каналам связи. Специализированная компьютерная ферма, конференц-зал и зал для проведения видеоконференций, поддерживаемые современным программным обеспечением, позволяют уже сейчас вести исследовательские работы на уровне ведущих физических центров мира.

## **ИСПОЛЬЗОВАНИЕ ВОЗМОЖНОСТЕЙ Ф-КЛАСТЕРА ПРИ РЕШЕНИИ ЗАДАЧ ДЕЙСТВУЮЩИХ ЭКСПЕРИМЕНТОВ**

Рост мощности вычислительной фермы персональных компьютеров Ф-кластера и появление в нем современных сервисов обеспечили его конкурентоспособность в отдельных задачах уже в 2000 г. Так, например, именно в этом году на компьютерной комплексе кластера была разработана, протестирована и запущена в эксплуатацию программа специализированного моделирования условий эксперимента NA-48, о которой уже было упомянуто выше (см. раздел «Предыстория вопроса» настоящей работы). С помощью этой программы на компьютерной ферме Ф-кластера было проведено и реальное моделирование, в котором использованы экспериментальные данные объемом 10,8 Тбайт. Выделение для решения этой задачи ~50% имеющейся на тот момент мощности фермы позволило завершить указанные работы в разумные сроки — примерно за три месяца. Эти данные были единственными, которые коллаборация NA-48 использовала в процессе вычисления систематических погрешностей основного физического результата эксперимента. Оценка же влияния эффектов наложения случайной активности, регистрируемой в эксперименте независимо от условий триггера, на результаты реконструкции событий была впервые получена именно группой сотрудников ОИЯИ, участвующих в эксперименте.

Еще одной важной задачей, решенной с использованием возможностей Ф-кластера в эксперименте NA-48, явилось участие ЛФЧ ОИЯИ в проведении полномасштабного моделирования условий эксперимента NA-48/2 для исследования систематических погрешностей при измерении асимметрии в распадах заряженных каонов на три заряженных пиона. Координация работ по созданию соответствующей программы моделирования, а затем и координация работ по проведению этого моделирования была возложена коллаборацией на дубненскую группу участников эксперимента. Наравне с другими пятью институтами ОИЯИ принял участие в моделировании (с последующей реконструкцией событий) и обеспечил получение 15% от общего его объема в более чем три миллиарда событий распада заряженных каонов. Для решения этой задачи было использовано около 50% всех ресурсов компьютерной фермы. В отдельные моменты в работе находились до 60 запущенных параллельно задач. В настоящее время все полученные данные переданы в ЦЕРН и доступны участникам коллаборации для использования. А данные в наиболее упакованном виде, называемом в коллаборации Nuprep-Compact, объемом более 200 Гбайт, переданы по компьютерным сетям из всех участвующих в проведении моделирования центров в Дубну и записаны на диски компьютерной фермы Ф-кластера для ведения необходимого их анализа «дома».

В рамках эксперимента COMPASS (NA-58 на SPS ЦЕРН) компьютерная ферма использовалась для проведения работ по определению продольной и

поперечной поляризации  $\Lambda$ - и анти- $\Lambda$ -гиперонов, для получения инклюзивной спиновой асимметрии  $A_1$ , получения полуинклюзивной азимутальной асимметрии, а также для анализа работы адронного калориметра эксперимента — HCAL1. Обработывались данные, накопленные в сеансе 2002 г. Общее количество обработанных событий составило  $\sim 6 \cdot 10^7$  при средней скорости обработки одного события на отдельном компьютере фермы 0,002 с. При решении указанных задач использовано около 800 Гбайт дискового пространства, 30% которого заняли экспериментальные данные, 70% — результаты моделирования. Полученные физические результаты доложены как на рабочих совещаниях коллаборации COMPASS, так и на ряде международных конференций (см., например, [8, 9]).

Значительный ресурс  $\Phi$ -кластера используется экспериментаторами при анализе данных и моделировании эксперимента ЭКСЧАРМ. Практически все последние результаты, полученные в этом эксперименте при исследовании рождения  $\phi$ -мезонов, изучении корреляций при рождении идентичных бозонов и исследовании процессов рождения и распада гиперонов и антигиперонов, связаны с использованием значительного вычислительного ресурса фермы и инфраструктурных возможностей кластера. Целый ряд расчетов по программе изучения  $B$ -физики в экспериментах на LHC и HERA по моделированию новых детекторов для проведения экспериментов на коллайдерах нового типа (CLIC, TESLA), не требующих экстремальных компьютерных ресурсов, был также успешно выполнен с использованием возможностей компьютерной фермы и кластера в целом.

Круг пользователей компьютерной инфраструктуры  $\Phi$ -кластера постоянно расширяется с ростом его потенциала и компьютерных мощностей, дальнейшее наращивание которых с одновременным расширением сопровождаемых новых и поддерживаемых сервисов делает этот ресурс все более привлекательным как для работающих, так и для потенциальных пользователей.

## ЛИТЕРАТУРА

1. *Foster I., Kesselman C.* The Grid. Blueprint for a New Computing Infrastructure. ISBN 1558604758. Morgan Kaufmann. August, 1998. P. 701.
2. *Segal B.* (on behalf of the Project). Grid Computing: the European Data Grid Project. IEEE Nuclear Science Symposium and Medical Imaging Conference. Lyon, 15–20 October, 2000.
3. CERN Council gives go ahead for LHC Computing Grid project. PR10.01 20.09.2001. CERN, September 2001.

4. *Kartashov S. N. et al.* LPP&LHE PC-farm at JINR. // XVIII International Symposium NEC'2001. Varna, 2001.
5. *Adamczyk D. et al.* Global Platform for Rich Media Conferencing and Collaboration. CHEP 03 La Jolla, USA. March 24–28 2003.
6. *Belosludtsev D. F., Potrebenikov Yu. K., Shchinov B. G.* Videoconferencing at LPP of JINR. // International Symposium on Nuclear Electronics & Computing (NEC'2003).
7. *Baud J. -P., Barring O., Durand J. -D.* CASTOR Project Status. CHEP2000, Teatro Antonianum, Padova, 2000.
8. *Peshehonov D.* Semi-inclusive spin asymmetry from COMPASS. // 16th International Spin Physics Symposium SPIN04. Trieste, 2004.
9. *Sapozhnikov M.* Longitudinal polarization of Lambda and anti-Lambda hyperons in deep inelastic scattering at COMPASS. // XVII International Baldin Seminar on High Energy Physics Problems «Relativistic Nuclear Physics and Quantum Chromodynamics». JINR, Dubna, 2004.

Получено 24 декабря 2004 г.

Редактор *М. И. Зарубина*

Подписано в печать 17.02.2005.

Формат 60 × 90/16. Бумага офсетная. Печать офсетная.

Усл. печ. л. 0,62. Уч.-изд. л. 0,76. Тираж 300 экз. Заказ № 54791.

Издательский отдел Объединенного института ядерных исследований  
141980, г. Дубна, Московская обл., ул. Жолио-Кюри, 6.

E-mail: [publish@pds.jinr.ru](mailto:publish@pds.jinr.ru)

[www.jinr.ru/publish/](http://www.jinr.ru/publish/)