УДК 520.27:520.88

Банк наблюдательных данных ODA–B/SS. Основная архивная база

В. К. Кононов, В. Г. Клочкова

Специальная астрофизическая обсерватория РАН, п.Нижний Архыз, 369167, Россия

Аннотация. Описывается текущая версия основной архивной базы Банка наблюдательных данных ODA–B/SS. Эта версия базы включает 3 независимых локальных архива для 3-х звездных спектрографов 6-метрового оптического телескопа БТА — NES, LYNX и PFES. Обсуждается содержание локальных архивов, процедуры их формирования и связь Банка данных с системами сбора телескопа. Приводится общая архивная статистика, а также статистика по отдельным разделам информационной системы.

The observational databank ODA–B/SS. The main archive database

V. K. Kononov, V. G. Klochkova

Special Astrophysical Observatory of the Russian AS, Nizhnij Arkhyz, 369167, Russia

Abstract. The current version of the main archive base of the observational databank ODA–B/SS is described. This version of the base includes 3 independent local archives conforming to three stellar spectrographs of the 6 m optical telescope BTA — NES, LYNX and PFES. The content of the local archives, the procedures of their forming and the connection of the databank with the data acquisition systems of the telescope are discussed. The general archive statistics and the statistics for individual divisions of the informational system are given.

Содержание

1 Введение

К концу 2003 г. мы завершили очередной этап работ по проекту создания Банка наблюдательных данных ODA–B/SS для звездных спектрографов 6-метрового оптического телескопа БТА. Проект ODA–B/SS (Observational Data Archive – BTA/Stellar Spectrographs) в окончательном виде был сформулирован в 2001 г. (Кононов, Панчук, 2001). С 2002 г. работы по этой тематике поддерживаются РФФИ (N 02–07–90245).

Главной целью проекта является создание мультибазовой распределенной информационной системы, обеспечивающей накопление в режиме "on-line" результатов всех наблюдений, проводимых на БТА с помощью звездных спектрографов, долговременное хранение полученного наблюдательного материала и автоматизированный доступ пользователей к архивной информации. Концепция Банка ODA–B/SS предусматривает создание и поддержку 2-х типов баз данных: *основной* архивной базы (ОБД) и вспомогательных *справочных* баз. ОБД содержит параметризованные числовые массивы — результаты наблюдений, представленные в самодокументированном формате. Многоуровневые проблемно-ориентированные справочные базы данных создаются как надстройки над основной базой и используются в системе для оптимизации доступа к архивной информации. Таким образом, ключевую роль в построении Банка данных играет формирование основной базы, которая несет на себе главную функциональную нагрузку, определяемую приоритетными задачами сохранения экспериментального материала.

Наиболее важными факторами при создании ОБД являются:

- охват архивными данными максимально возможного временнóго периода;
- максимальная полнота архивных данных (включение результатов всех проводимых наблюдений);
- распределение информации по отдельным независимым астрономическим приборам;
- максимальная и корректная параметризация экспериментальных данных;
- эффективная форма организации и хранения архивного материала;
- оперативное пополнение ОБД новой информацией.

Все эти моменты отражены в настоящей статье, которая содержит результаты последних разработок в области архивизации наблюдательных данных звездных спектрографов телескопа БТА.

2 Основная архивная база

Работы по формированию основной архивной базы Банка ODA–B/SS выполнялись в два этапа, в результате чего были последовательно созданы две версии ОБД.

2.1 Объединенная архивная база (первая версия)

Конечной целью первого этапа было формирование архивной базы, максимально охватывающей все наблюдения, выполненные с помощью звездных спектрографов за весь период их эксплуатации на телескопе. Исходной информацией для этого послужили как текущие (новые) данные, постоянно поступающие с приборов в результате очередных сетов наблюдений, так и старый экспериментальный материал, накопленный в предшествующие годы и сохраненный в том или ином виде на различных машинных носителях. Старые данные были представлены первичным архивом, созданным в 90-е годы системами архивизации ODA Version 1.0 и 1.1 (Кононов и др., 1996; Кононов, 1996a). Эти системы функционировали, соответственно, в среде операционных систем (ОС) NTS на машинах типа СМ и Linux на IBM-совместимых персональных компьютерах (Кононов, Евангели, 1991). Первичный архив размещался сначала на томах бобинных магнитных лент объемом 10 Mb, а в дальнейшем — на DDS-кассетах (Digital

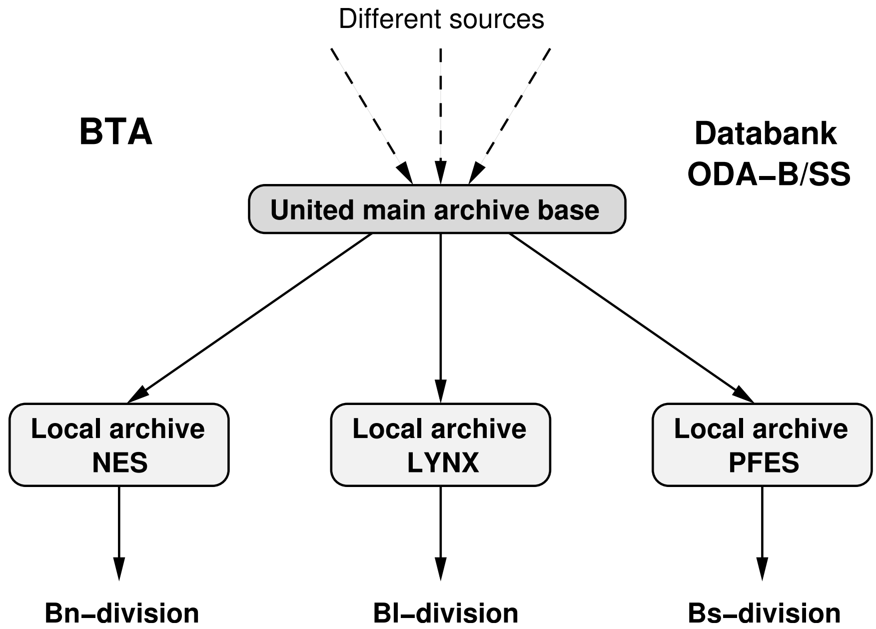

Рис. 1: Объединенная основная архивная база и локальные архивы Банка данных ODA–B/SS.

Data Storage) емкостью 1.3–2 Gb (Кононов, 1996б). Все старые данные хранились в форматах, предусмотренных соответствующими вариантами систем архивизации.

В результате сбора, анализа и систематизации большого объема исходного материала в 2002 г. была создана первая (объединенная) версия ОБД, которая включила в себя все ранее накопленные многолетние наблюдательные данные, упорядоченные в хронологии их получения, но без разделения их по типам звездных спектрографов. Эта версия основной архивной базы на 1.01.2003 г. имела следующие характеристики:

Архивный интервал	—	1992–2002 гг.
Число годов	—	11
Число наблюдательных дат	—	295
Число наблюдений	—	6590
Объем данных	—	29.0 Gb

Сформированная таким образом ОБД была размещена на 46-ти томах оптических дисков CD–R (Compact Disk – Recordable) емкостью 650–700 Mb и стала основой для дальнейшего развития Банка ODA–B/SS и формирования его стандартизованных архивных баз (рис. 1).

2.2 Локальные архивы звездных спектрографов

На втором этапе была проведена полная реорганизация объединенной архивной базы для создания *локальных* архивов, отражающих фрагментацию ОБД по типам данных. Каждый локальный архив представляет собой подмножество основной базы и ориентирован на поддержку экспериментальных данных определенного типа — данных конкретного спектрографа (Кононов, Панчук, 2000а,б). Важным моментом при этом является соблюдение принципа *однородности*, который предполагает проектирование одинаковых по своей структуре специализированных подсистем архивизации — разделов Банка ODA–B/SS, а каждый локальный архив представляет собой основную архивную базу соответствующего раздела системы (рис. 1).

В процессе реорганизации особое внимание было уделено единой форме представления экспериментальных массивов. Поскольку в качестве стандартного для данных всех звездных спектрографов БТА принят международный Basic FITS-формат (Ханиш и др., 2001), были выполнены необходимые операции по приведению части материала к стандартному архивному формату. Эта процедура затронула около 1300 наблюдений ($\sim 20\%$ от общего числа). Одновременно была удалена часть информации, которая носила вспомогательный характер или являлась некорректной с точки зрения астрономов-наблюдателей (в основном за 1992 г.). Сортировка данных по типам осуществлялась по параметрам заголовков FITS-файлов с дополнительным анализом многолетних журналов наблюдений во избежание возможных ошибок. Параллельно с реорганизацией ОБД проводилось накопление нового экспериментального материала и его архивизация.

В результате в 2003 г. была создана новая версия основной архивной базы с перераспределением информации по 3-м локальным архивам для 3-х звездных спектрографов и соответствующим им типам данных:

Bn (NES)	—	эшелле-спектрограф с коллимированным пучком большого диаметра в фокусе Нэсмит–2;
Bl (LYNX)	—	эшелле-спектрометр высокого разрешения в фокусе Нэсмит–2;
Bs (PFES)	—	эшелле-спектрометр в первичном фокусе.

В скобках указаны краткие названия приборов. В целом состояние новой версии ОБД на 1.01.2004 г. характеризуется следующими параметрами:

Архивный интервал	—	1993–2003 гг.
Число годов	—	11
Число наблюдательных дат	—	329
Число наблюдений	—	8234
Объем данных	—	42.5 Gb
Число типов данных	—	3

2.3 Структуры архивных томов

Идеология Банка данных ODA–B/SS, в зависимости от особенностей задач последующей обработки экспериментального материала, предусматривает различные варианты организации и хранения локальных архивов, включая их помещение на тот или иной машинный носитель и реализацию определенных методов доступа к архивной информации.

2.3.1 Постоянные и резидентные архивы

Эталонной формой для долговременного хранения является размещение локальных архивов на сменных томах оптических дисков CD–R в виде *постоянных* (invariable) архивов. При этом удовлетворение запросов на доступ к данным постоянных архивов реализуется с помощью администратора Банка, поскольку требует монтировки отдельных томов, то есть выполнения ручных операций. Одновременно возможен вариант поддержки *резидентных* архивов — архивов, размещенных непосредственно на жестких дисках HD (Hard Disk) в виде полных копий постоянных архивов или их отдельных частей. Такие резидентные архивы предоставляют пользователям возможность прямого доступа к данным без какого-либо посредничества. С этой точки зрения использование резидентных архивов весьма эффективно, однако их объемы зависят от ресурсов архивного сервера.

В любом случае с формой хранения локальных архивов связано понятие архивного тома и его структуры. Отправным пунктом здесь является том прямого доступа постоянного архива — CD–R. В резидентных архивах все данные аналогичным образом разбиты на тома в виде копий

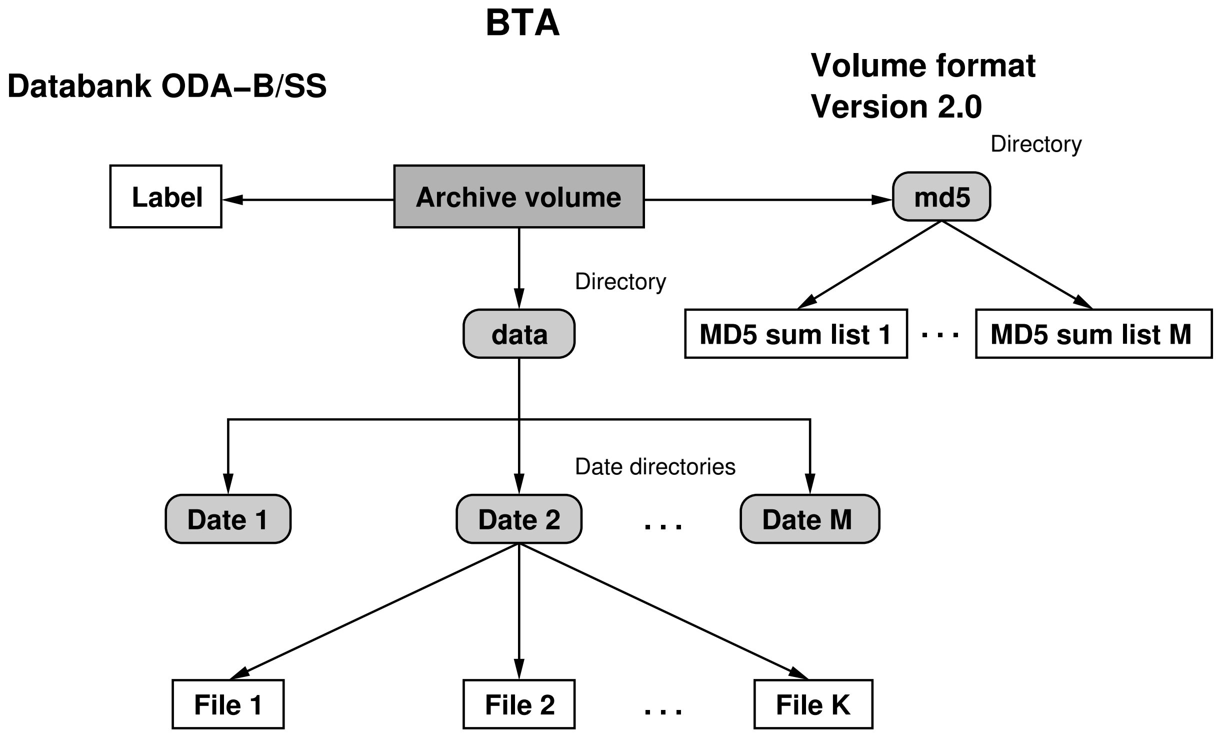

Рис. 2: Формат архивного тома. Версия 2.0.

соответствующих томов постоянных архивов, размещенных по отдельным каталогам файловой системы на HD. Все это обеспечивает гибкость поддержки резидентных архивов как с точки зрения динамичного изменения их содержимого, так и для реализации унифицированного доступа к информации.

Формирование и сопровождение постоянных архивов на оптических дисках CD–R в первую очередь предполагает определенную устойчивость структур архивных данных, поддержанную соответствующими стандартами. Это является еще одним проявлением принципа *однородности* в проектировании интегрированной информационной системы. Поскольку созданные локальные архивы для 3-х звездных спектрографов БТА требуют в этом смысле дополнительной реорганизации, все они пока представлены только в виде резидентных архивов (плюс резидентные копии) с параллельным хранением постоянного архива первой объединенной версии ОБД.

2.3.2 Формат архивного тома. Версия 2.0

В Банке ODA–B/SS используется унифицированный формат архивного тома версии 2.0, который был заимствован у Банка данных ODA–R радиотелескопа РАТАН–600 (Кононов, 2003). Этот формат, разработанный в 2001 г., не зависит от типа данных и форматов внутреннего представления файлов и позволяет по сравнению с предыдущей версией:

- упростить схему размещения данных по подкаталогам;
- развить элементы самоконтролируемой структуры тома;
- ввести стандарт идентификации томов.

Все файлы с наблюдательными данными размещаются в стандартном каталоге *data* с 1-уровневой структурой подкаталогов — по подкаталогам-датам (рис. 2). Для каталогов-дат введены стандартные имена в виде полных идентификаторов дат вида

$$yyyymmdd \quad ,$$

где: $yyyy$ — год (4 цифры);
mm — месяц (2 цифры);
dd — день (2 цифры).

В каждом каталоге-дате находятся только файлы с наблюдательными данными.

В стандартном каталоге *md* размещаются файлы-описатели, содержащие списки контрольных сумм файлов с данными соответствующих дат, подсчитанных по алгоритму утилиты *md5sum* ОС Linux. Имена файлов-описателей соответствуют версии 2 общего стандарта идентификации FLEX–N в части описателей дат (Кононов, 2002a) и имеют следующую структуру:

$$kkyyyymmdd.\mathrm{D\_md5} \quad ,$$

где: kk — тип наблюдательных данных (2 символа);
$yyyymmdd$ — 8-символьный идентификатор даты;
.D_md5 — стандартный постфикс.

Архивный том идентифицируется файлом-меткой, который имеет имя вида

$$\mathrm{Volume}{=}\mathit{identifier} \quad ,$$

где *identifier* — обобщенный идентификатор архивного тома, имеющий стандартную структуру:

$$kkgnnnmT\ldots \quad ,$$

где: kk — тип наблюдательных данных;
g — односимвольный идентификатор группы томов.
Для архивных томов с экспериментальными данными — символ “_”;
nnn — порядковый номер тома (3 цифры);
m — односимвольный модификатор (по умолчанию — символ “_”);
$T\ldots$ — тип архивного носителя (строка символов произвольной длины).

Примеры идентификаторов: Bl_002aCDR74, Bn_001_CDR80.

2.3.3 Формирование архивных томов

Постоянное расширение локальных архивов за счет поступления новых наблюдательных данных требует решения вопросов разбиения исходного потока информации по архивным томам. Это относится к любой форме хранения локальных архивов, в том числе и к резидентным архивам, если они являются первичными. Общие правила проведения границы в исходных данных при переключении томов, то есть при переходе от одного тома к следующему, описаны в работе Кононова (2003). Для 3-х локальных архивов звездных спектрографов эти правила одинаковы и формально относятся к j-уровню — уровню наблюдения (рис. 3).

Особенностью j-уровня является то, что данные одного наблюдения не разделяются архивными томами, в то время как суточные данные могут разделяться. Другими словами, данные конкретной даты могут находиться на разных томах, а данные одного и того же наблюдения — только на одном томе. Такой вариант размещения информации используется тогда, когда максимальный объем суточных данных соизмерим с емкостью архивного носителя, а максимальный объем одного наблюдения составляет небольшой процент от емкости тома (до 10–15%). Действительно, в случае звездных спектрографов максимальный объем суточных данных может достигать сотен Mb при стандартной емкости архивного тома CD–R — 700 Mb, а объем

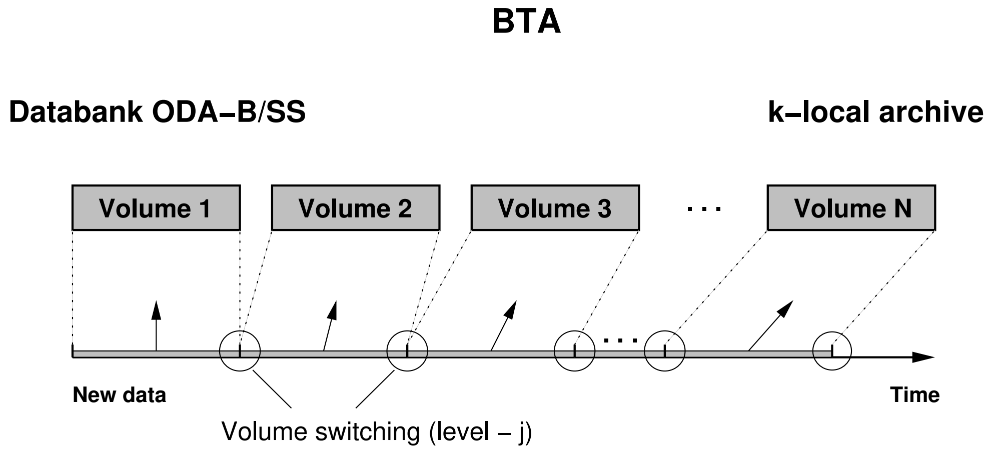

Рис. 3: Формирование томов основной архивной базы для k-го локального архива.

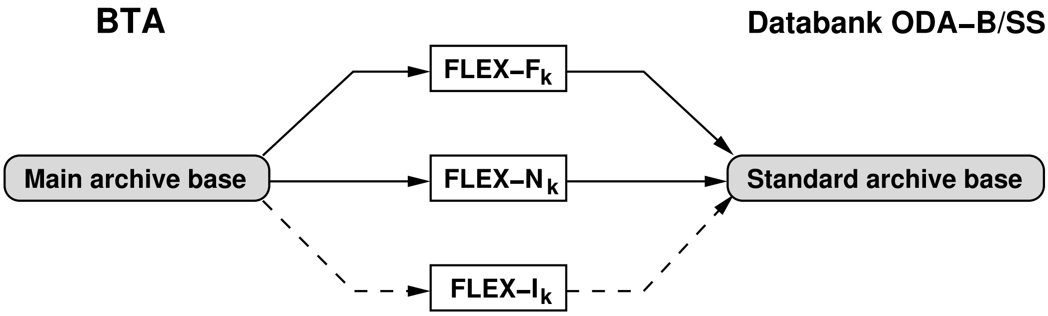

Рис. 4: Стандартизация локальных архивов.

одного наблюдения (например, для спектрографа NES) составляет 8 Mb — всего 1% от емкости тома. Таким образом, использование j-уровня разбиения информации по томам с точки зрения заполняемости архивных носителей очень эффективно.

Следует отметить, что для каждого локального архива создается свой отдельный набор архивных томов при соблюдении единых правил их идентификации.

2.3.4 Вопросы стандартизации и создания постоянных архивов

Как уже было отмечено, постоянные архивы должны играть роль эталонов достаточно продолжительное время, что предполагает их максимально стандартизованную организацию. Именно поэтому сравнительно недавно созданные локальные архивы звездных спектрографов пока представлены только в резидентном виде: они требуют дополнительного преобразования, чтобы полностью соответствовать принятым ранее стандартам, лежащим в основе концепции унифицированного FLEX-интерфейса (FLexible EXchange) (Кононов, 1995а). Следует также заметить, что выполнение достаточно трудоемкой процедуры переноса большого объема информации на CD–R (на 1.01.2004 г. — это 63 тома) целесообразно совместить с предварительным приведением данных к стандартному виду, а не создавать промежуточные версии постоянных архивов.

Стандартизация локальных архивов должна выполняться, по крайней мере, в 2-х областях (рис. 4). Первое направление работ связано с приведением всех файлов к единому формату — приближением к подобным друг другу FLEX–F_k-стандартам (Кононов, 1995б). Эти работы уже завершены, и в этом смысле локальные архивы стали однородными.

Второе направление работ — это внедрение FLEX–N_k-стандартов — унифицированной системы идентификации файлов (Кононов, 2002a), которая уже реализована и автоматически поддерживается новой версией системы сбора спектрографа NES, начиная с августа 2003 г. Все файлы многолетних данных приборов LYNX и PFES, а также частично и NES, должны быть соответствующим образом переименованы. Это потребует комплексного анализа FITS-заголовков с привлечением журналов наблюдений. Одновременно должны быть решены вопросы по снятию неопределенностей или некорректностей в параметризации значительной части сравнительно старых данных (например в идентификации объектов исследований или типов наблюдений). Кроме того, предполагается решить проблемы приведения к FLEX–N_k-стандартам информационных выходов новых версий систем сбора спектрографов LYNX и PFES (по мере их внедрения в практику наблюдений).

Третье направление работ для спектрографов БТА является только перспективным, так как оно связано с дополнительным структурированием информационных выходов систем сбора на основе системы FLEX–I_k-стандартов для поддержки журнализации астрофизических экспериментов (Кононов, 1995в; 1996в). Эти вопросы могут быть решены только лишь в случае стабильной и максимально автоматизированной работы систем сбора и потребуют дополнительных обсуждений. Поэтому на данном этапе для стандартизации постоянных архивов это направление является несущественным.

3 Взаимодействие с системами сбора

Одной из главных задач Банка ODA–B/SS является оперативное наполнение локальных архивов новыми данными при взаимодействии с системами сбора звездных спектрографов в режиме “on-line”. Такое взаимодействие должно осуществляться с помощью специализированного программного обеспечения путем формализации межсистемных связей с использованием оптимальных многомашинных конфигураций на базе локальной вычислительной сети обсерватории.

3.1 Сетевая система буферизации ODA/BS–2L

В процессе разработки многофункционального Банка ODA–B/SS принципиальным моментом является создание его специализированных программных подсистем, каждая из которых предназначена для решения своего комплекса задач. Разделение Банка на отдельные функционально-ориентированные подсистемы позволяет более эффективно реализовывать различные внутренние процедуры, обеспечивающие непосредственную архивизацию данных, поддержку логических структур архивных баз и интерфейсов, контроль корректности данных и диагностику, организацию поиска и выборки информации, а также установление внешних связей с другими системами.

В рамках проекта ODA–B/SS проблема взаимодействия с системами сбора звездных спектрографов решена с помощью 2-уровневой сетевой подсистемы буферизации данных ODA/BS–2L (Oservational Data Archive/Buffering System – 2-Layered). При построении этой специализированной подсистемы были учтены следующие наиболее важные требования:

- Банк данных как информационная система включается в непрерывный наблюдательный цикл БТА и функционирует круглосуточно.

- Взаимодействие между системами осуществляется в максимально автоматизированном режиме, с вмешательством со стороны администратора Банка данных только лишь в экстренных ситуациях.

- Межсистемные связи поддерживаются в режиме “on-line” с минимальной задержкой по времени между моментами регистрации новых данных и их окончательного оформления в Банке ODA–B/SS и предоставления доступа к информации с помощью стандартных средств системы.

- Интерфейсы с различными системами сбора максимальным образом унифицируются и поддерживаются единым программным комплексом Банка данных.

- Процесс получения и накопления новых экспериментальных данных подробным образом журнализируется для предоставления администратору Банка полной информации о передаче данных и их преобразовании на каждом этапе.

- Конфигурации межсистемных связей должны быть достаточно гибкими, а программные средства поддержки должны предоставлять возможность их глобальной настройки по параметрам для оптимального соответствия реальным условиям функционирования систем.

По существу ODA/BS–2L является унифицированной входной интерфейсной оболочкой Банка данных и представляет собой специализированный комплекс программ и командных файлов, разработанных в операционной среде ОС Linux Red Hat. Ее апробированная идеология целиком заимствована у соответствующей подсистемы Банка наблюдательных данных ODA–R радиотелескопа РАТАН–600, которая была введена в рабочую эксплуатацию в 2000 г. для данных радиометров континуума, а в дальнейшем стала поддерживать также солнечные и спектральные радиоданные (Кононов, 2002б). Подсистема реализует межсистемные связи посредством непрерывных автоматизированных конвейеров по передаче и приему экспериментальных данных, асинхронно поступающих от разных систем сбора звездных спектрографов. Архитектура ODA/BS–2L отражает особенности общей организации Банка ODA–B/SS: она также состоит из отдельных подсистем, каждая из которых отвечает за обработку соответствующего типа данных — Bn, Bl и Bs, то есть настраивается на автономное взаимодействие с конкретной системой сбора. По своей структуре и набору функций эти подсистемы подобны друг другу.

На рис. 5 показана каноническая схема 2-уровневой буферизации наблюдательных данных k-го типа. Для обработки каждого типа данных ODA/BS–2L использует 2 специализированных буфера, через которые последовательно проходит вся экспериментальная информация:

1. *Буфер связи*, заполняемый непосредственно системой сбора в режиме “on-line” и содержащий результаты последнего наблюдения;

2. *Оперативный буфер*, в котором накапливаются результаты наблюдений текущих суток (или более при необходимости).

Оба буфера являются разделяемыми информационными ресурсами: буфер связи разделяется системой сбора и оболочкой ODA/BS–2L, оперативный буфер — оболочкой ODA/BS–2L и ядром Банка данных. Очистка буферов после регулярной передачи данных из одного буфера в другой, а затем собственно в архив, с одновременным контролем данных и их сортировкой выполняется согласованными между собой процедурами Банка, запускаемыми демоном *cron* ОС Linux. Текущее содержание системных буферов и возможные экстремальные варианты их заполнения зависят от интенсивности межбуферного обмена, которая в свою очередь определяется параметрами настройки в централизованных описаниях.

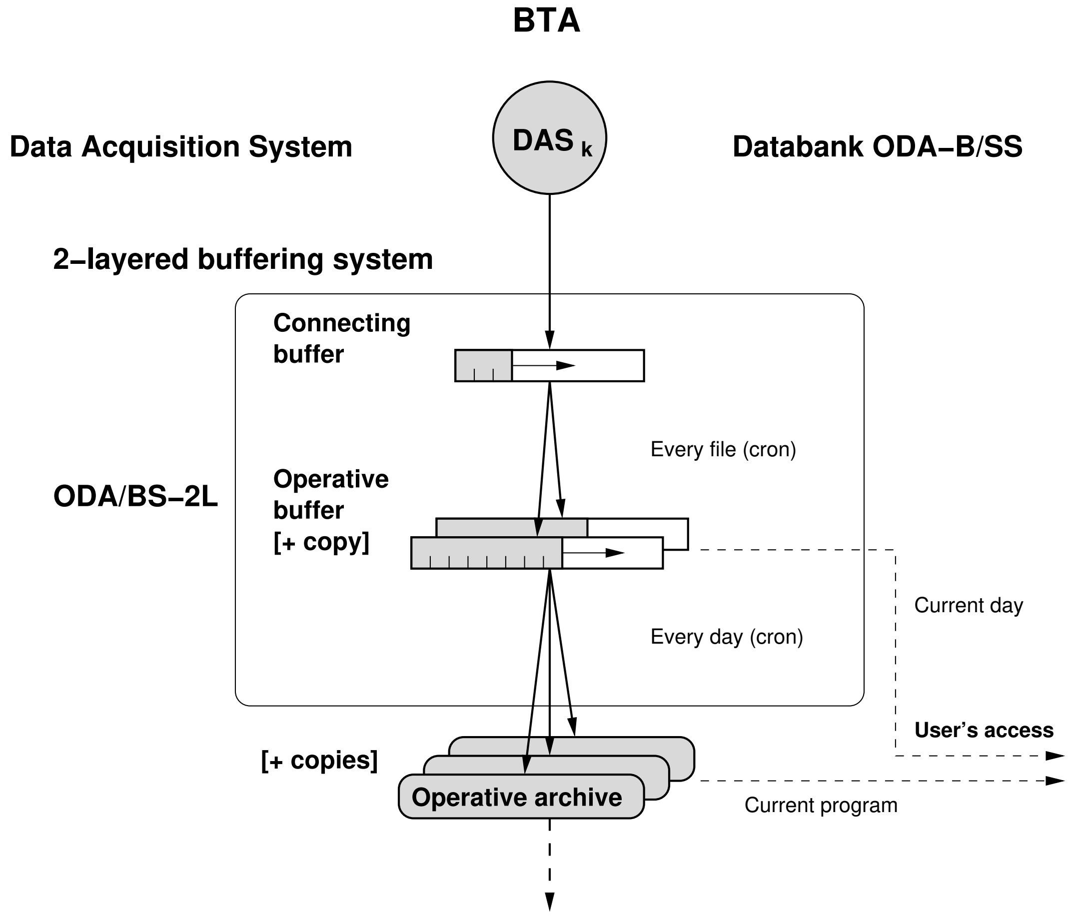

Рис. 5: Схема 2-уровневой буферизации наблюдательных данных k-го типа.

3.2 Конфигурации межсистемных связей

Подсистема ODA/BS–2L является сетевой и реализует идею "плавающих" буферов с организацией межмашинных связей под управлением архивного сервера. Суть ее заключается в том, что буферы, через которые осуществляется связь систем сбора БТА с Банком данных ODA–B/SS, могут размещаться на разных компьютерах локальной сети САО. Различные варианты размещения буферов на машинах сети определяют соответствующие конфигурации межсистемных связей.

В общем случае в технологической цепочке "сбор – архивизация" одного типа данных могут быть задействованы, в зависимости от задач, от 2-х до 4-х компьютеров. При этом в минимальной конфигурации используются только 2 компьютера: машина сбора и архивный сервер. Дополнительно в цепочку может быть включено до 2-х вспомогательных (так называемых транзитных) машин (Кононов, 2002б).

В настоящее время для организации автоматических конвейеров передачи данных спектрографов NES, LYNX и PFES в качестве машины сбора используется *uppsala.sao.ru*, расположенная на Верхней научной площадке САО, а архивным сервером является компьютер *rain.sao.ru*, находящийся в лабораторном корпусе на Нижней научной площадке обсерватории, в 3-х км от БТА. При этом машина *rain* имеет также статус архивного сервера II радиотелескопа РАТАН–600 и поддерживает зеркальную копию Банка наблюдательных данных ODA–R.

Впервые автоматический конвейер передачи данных с телескопа БТА в Банк ODA–B/SS

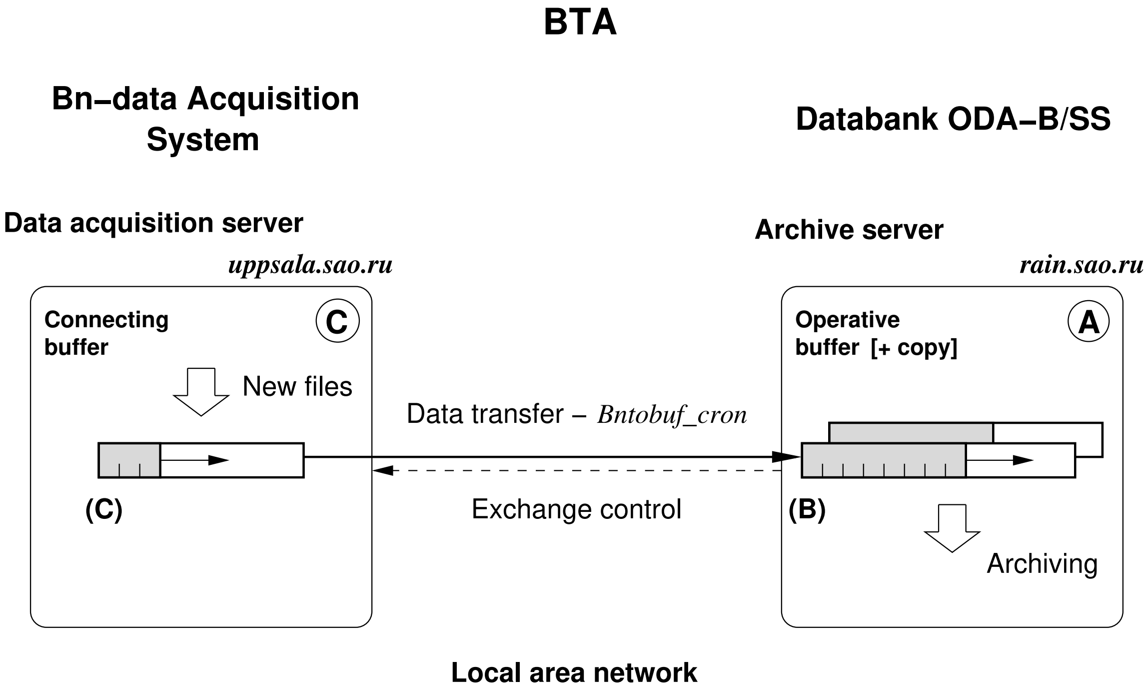

Рис. 6: 2-машинная конфигурация межсистемных связей $C_C^{uppsala} A_B^{rain}$ для данных Bn-типа (спектрограф NES).

был запущен в августе 2003 г. для спектрографа NES, после введения в эксплуатацию новой версии его системы сбора. Причем была задействована оптимальная 2-машинная конфигурация, которая, если использовать обозначения из работы Кононова (2002б), реализует самый эффективный вариант межсистемных связей — $C_C^{uppsala} A_B^{rain}$. Иначе говоря, буфер связи размещен на машине сбора *uppsala*, а оперативный буфер — на архивном сервере *rain* (рис. 6). Такой вариант позволяет минимизировать сетевые обмены и нагрузку на систему сбора и обеспечивает оптимальный режим работы с оперативным буфером при организации доступа к данным текущих суток.

Для спектрографов LYNX и PFES аналогичные автоматические конвейеры передачи данных планируется активизировать после модернизации соответствующих систем сбора и приведения их информационных выходов к FLEX-стандартам. Они будут использовать тот же вариант межсистемных связей — $C_C^{uppsala} A_B^{rain}$.

3.3 Процедуры сетевых обменов и журнализация

2-машинная конфигурация определяет следующую общую схему формирования, передачи и архивизации данных:

1. Система сбора спектрографа в результате очередного наблюдения создает файл в буфере связи на машине *uppsala*.

2. Оболочка ODA/BS–2L копирует этот файл в оперативный буфер на архивный сервер *rain*, уничтожая исходный файл в буфере связи после удачного завершения операции копирования.

3. Данные текущей даты, накопленные в оперативном буфере, раз в сутки архивизируются

```
# Date      T1      T2     Tsec  Nf    Nkb Rkbs  RC Knet Kep Klf Kmd  Note
...
20030821  000000  000212  132   14 115332  874   0   6    1   1   1    +m
20030821  010000  010455  295   30 247140  838   0   6    1   1   1    +m
20030821  020000  020500  300   29 238902  796   0   6    1   1   1    +m
20030821  030000  030045   45    5  41190  915   0   6    1   1   1    +m
20030821  040000  040020   20    2  16476  824   0   6    1   1   1    +m
20030821  050000  050011   11    1   8238  749   0   6    1   1   1    +m
20030821  060000  060217  137   15 123570  902   0   6    1   1   1    +m
20030821  070000  070307  187   20 164760  881   0   6    1   1   1    +m
20030908  130000  130010   10    1   8238  824   0   6    1   1   1    +m
20030908  140000  140015   15    1   8238  549   0   6    1   1   1    +m
20030908  150000  150044   44    2  16476  374   0   6    1   1   1    +m
20030908  160000  160021   21    2  16476  785   0   6    1   1   1    +m
20030908  170000  170019   19    2  16476  867   0   6    1   1   1    +m
20030908  180000  180029   29    3  24714  852   0   6    1   1   1    +m
20030908  190000  190019   19    2  16476  867   0   6    1   1   1    +m
...
```

Рис. 7: Фрагмент журнала процедуры *Bntobuf_cron*.

ядром Банка. После успешного завершения архивизации вся исходная информация из оперативного буфера удаляется.

Непосредственный межбуферный сетевой обмен осуществляется оболочкой ODA/BS–2L путем периодического запуска с помощью демона *cron* процедур типа *tobuf*. Для каждого типа данных (спектрографа) выполняется своя такая процедура. Например, для данных Bn-типа это — *Bntobuf_cron* (рис. 6). Все подобные процедуры запускаются асинхронно по отношению друг к другу, обеспечивая независимую “подкачку” новых данных в различные локальные архивы, используя независимые пары системных буферов.

Особенностью процедуры типа *tobuf* является то, что за один вызов она забирает все имеющиеся данные в буфере связи (один или несколько файлов). Поэтому степень средней и пиковой заполняемости буфера связи находится в зависимости от соотношения частот создания новых файлов системой сбора и активизации процедуры оболочки. В настоящее время задано так, что процедура *Bntobuf_cron* для данных спектрографа NES запускается один раз в час. Процедуры *Bltobuf_cron* и *Bstobuf_cron* для LYNX и PFES пока деактивированы.

Все процедуры типа *tobuf* в процессе работы с системными буферами автоматически ведут свои журналы, фиксируя там основные параметры, которые могут быть использованы администратором Банка для сбора статистики или для анализа нештатных ситуаций. Структуры журналов унифицированы. На рис. 7 в качестве примера приведен фрагмент журнала процедуры *Bntobuf_cron*, который содержит следующие параметры (по колонкам):

$Date$ — дата вызова процедуры (*yyyymmdd*);
T_1 — время инициирования процедуры (*hhmmss*);
T_2 — время завершения процедуры (*hhmmss*);
T_{sec} — длительность работы процедуры, сек;
N_f — число переданных файлов;
N_{kb} — суммарный объем переданных данных, kb;
R_{kbs} — оценочное значение скорости передачи данных — N_{kb}/T_{sec}, kb/сек;

RC	—	код возврата процедуры;
K_{net}	—	значение счетчика сетевых обращений внутри процедуры;
K_{ep}	—	значение счетчика использования буферов процедурами низкого приоритета;
K_{lf}	—	значение счетчика контроля завершения формирования нового файла в буфере связи;
K_{md}	—	значение счетчика контроля совпадения контрольных сумм файлов при удаленном копировании;
$Note$	—	комментарий (в данном случае "$+m$" означает, что используется поддержка копии оперативного буфера).

4 Архивная статистика

Ниже приводится статистика, собранная по локальным архивам Банка ODA–B/SS. Она отражает состояние второй версии ОБД на 1 января 2004 г. Используются следующие обозначения:

N_y	—	число годов;
N_d	—	число наблюдательных дат;
N_o	—	число наблюдений;
N_{Mb}	—	объем данных, Mb;
N_{Gb}	—	объем данных, Gb;
$N_{Gb}, \%$	—	относительный объем данных, %.

Наблюдательной датой считается дата, в течение которой было выполнено, по крайней мере, одно наблюдение.

4.1 Общая архивная статистика

В табл. 1 приведено распределение по всем годам суммарных значений параметров, включая N_d, для всех локальных архивов Банка (учтен тот факт, что смена приборов в течение одной ночи не проводилась).

Рис. 8, 9 и 10 иллюстрируют распределения по годам суммарных значений объема данных, числа наблюдательных дат и числа наблюдений для всех типов данных. Резкий рост объема спектральной информации в последние 3 года (до 14 Gb/год) обусловлен преимущественным использованием в наблюдениях спектрографа NES с ПЗС-светоприемником размером 2000×2000 элементов.

4.2 Статистика по локальным архивам

Табл. 2 содержит общие параметры 3-х локальных архивов Банка ODA–B/SS.

На рис. 11 дано сравнение в форме диаграмм 3-х созданных локальных архивов системы по объемам данных, числу наблюдательных дат и числу наблюдений.

В табл. 3, 4 и 5 приведены значения основных параметров по годам для каждого локального архива в отдельности.

Рис. 12, 13 и 14 иллюстрируют, соответственно, распределения по годам объемов полученной информации, числа наблюдательных дат и числа наблюдений для каждого типа данных.

5 Заключение

В результате выполнения целого комплекса работ, включая развертывание необходимой технической базы, разработку программного обеспечения, анализ и реорганизацию большого объема

Таблица 1: Распределение по годам суммарных значений параметров для 3-х локальных архивов

$Year$	N_d	N_o	N_{Mb}
1993	2	3	2
1994	10	92	55
1995	21	272	546
1996	22	353	835
1997	32	757	1790
1998	28	496	1053
1999	30	723	1707
2000	36	667	1574
2001	49	1488	9055
2002	52	1615	12821
2003	47	1768	14248
$Total$	329	8234	43686

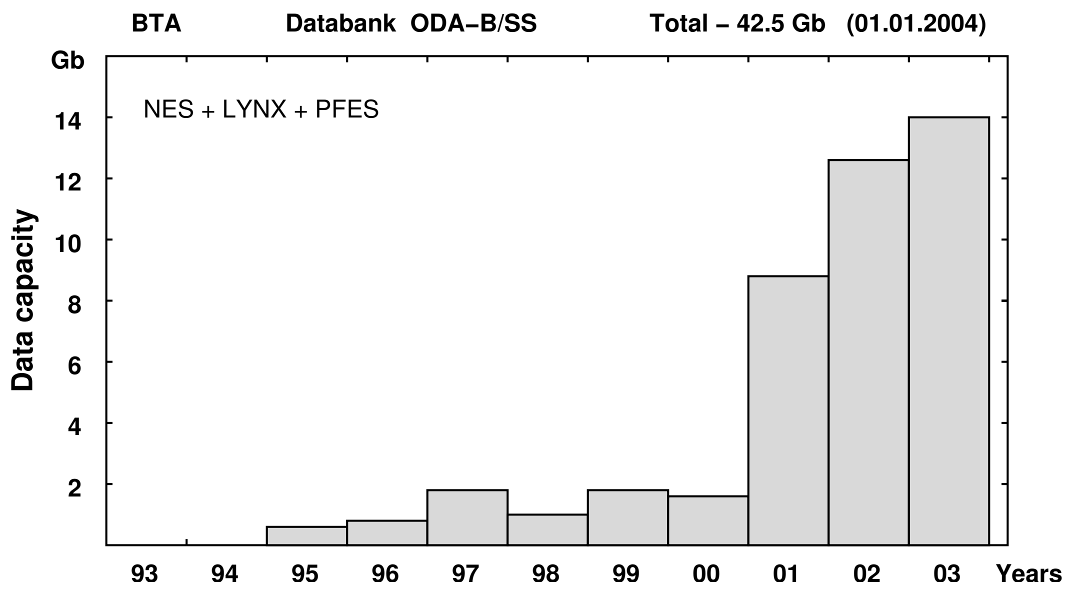

Рис. 8: Распределение по годам суммарных значений объемов архивных данных для 3-х локальных архивов.

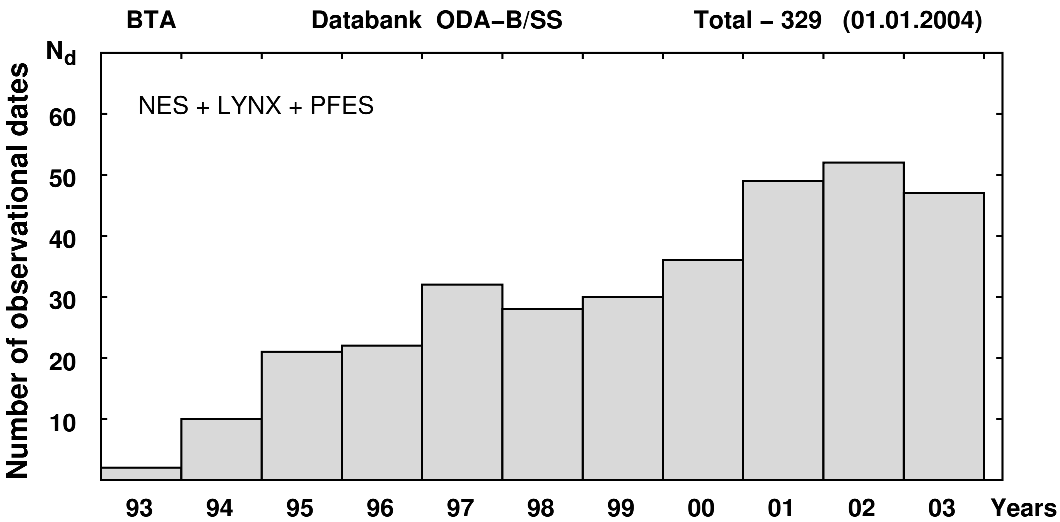

Рис. 9: Распределение по годам суммарных значений числа наблюдательных дат для 3-х локальных архивов.

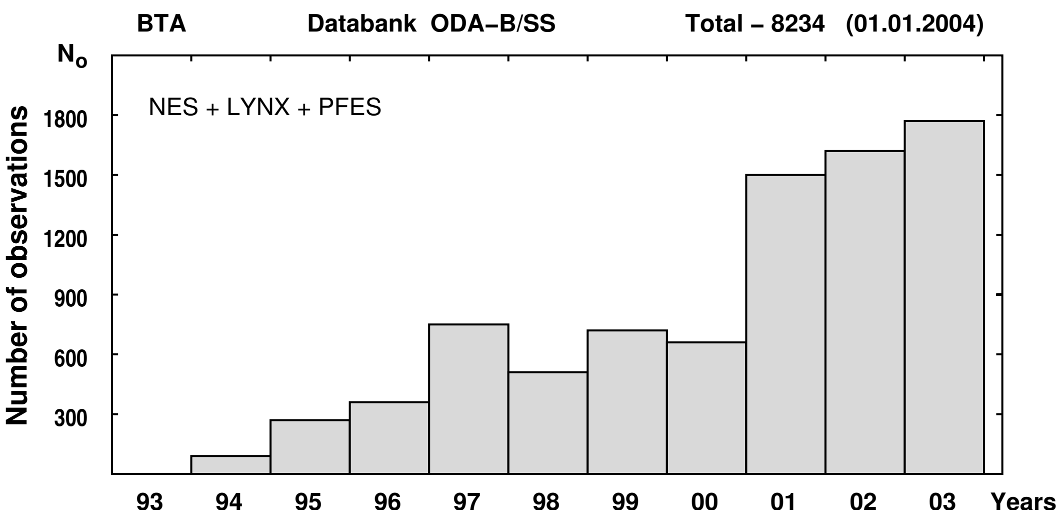

Рис. 10: Распределение по годам суммарных значений числа наблюдений для 3-х локальных архивов.

Таблица 2: Общие параметры 3-х локальных архивов на 1.01.2004 г.

Data type	*Spectrograph*	*Years*	N_y	N_d	N_o	N_{Gb}	$N_{Gb}, \%$
Bn	NES	1998–2003	6	145	4716	34.7	81.6
Bl	LYNX	1993–2003	11	94	1476	3.1	7.3
Bs	PFES	1996–2003	8	90	2042	4.7	11.1
Total		1993–2003	11	329	8234	42.5	100.0

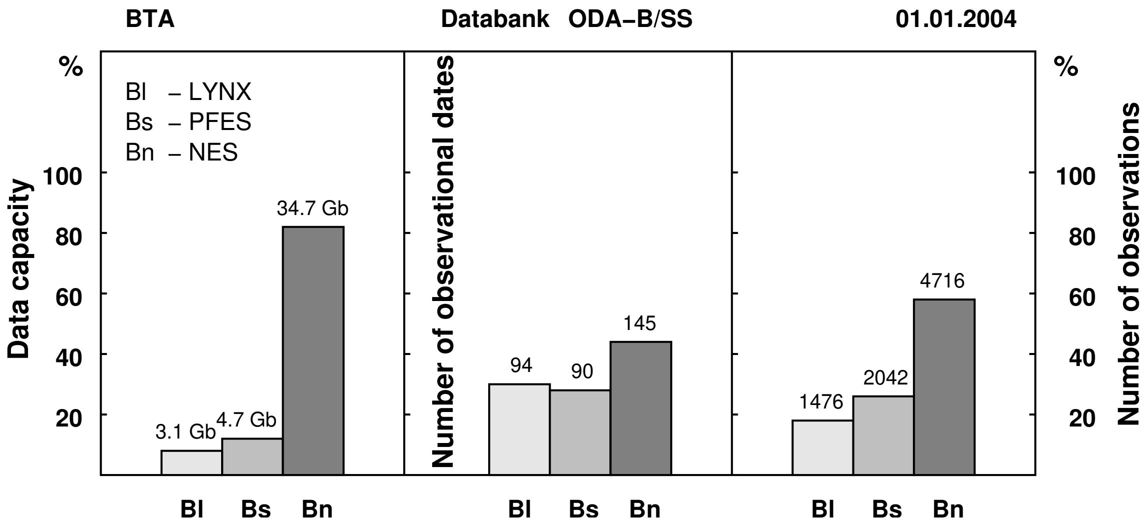

Рис. 11: Сравнение локальных архивов по объемам данных, числу наблюдательных дат и числу наблюдений.

многолетнего наблюдательного материала, впервые сформирована и сопровождается максимально систематизированная и доступная астрономам архивная база звездных спектрографов БТА.

Главным итогом за последние два года стало создание в значительной степени стандартизованных локальных архивов Банка ODA–B/SS и инициирование работ по организации автоматизированных конвейеров для оперативного наполнения этих архивов новыми экспериментальными данными. Включение Банка ODA–B/SS в наблюдательный цикл телескопа БТА существенно расширяет возможности всей системы автоматизации астрофизических наблюдений, поскольку обеспечивает надежное централизованное сохранение уникальной информации в режиме “on-line”.

Благодарности. Авторы признательны Н.С. Таволжанской и М.В. Юшкину из Лаборатории астроспектроскопии за работу по реорганизации архивной базы и внедрение FLEX-стандартов в системы сбора звездных спектрографов, а также Л.В. Минаковой за помощь в подготовке текста данной статьи.

Авторы выражают благодарность РФФИ за поддержку работы грантом N 02–07–90245.

Таблица 3: Общая статистика по Bn-локальному архиву (NES) на 1.01.2004 г.

Year	N_d	N_o	N_{Mb}
1998	18	309	611
1999	2	46	109
2000	0	0	0
2001	27	1012	7932
2002	51	1581	12741
2003	47	1768	14248
Total	145	4716	35641

Таблица 4: Общая статистика по Bl-локальному архиву (LYNX) на 1.01.2004 г.

Year	N_d	N_o	N_{Mb}
1993	2	3	2
1994	10	92	55
1995	21	272	546
1996	18	275	651
1997	7	99	234
1998	0	0	0
1999	10	252	595
2000	16	279	658
2001	9	170	401
2002	1	34	80
2003	0	0	0
Total	94	1476	3222

Таблица 5: Общая статистика по Bs-локальному архиву (PFES) на 1.01.2004 г.

Year	N_d	N_o	N_{Mb}
1996	4	78	184
1997	25	658	1556
1998	10	187	442
1999	18	425	1003
2000	20	388	916
2001	13	306	722
2002	0	0	0
2003	0	0	0
Total	90	2042	4823

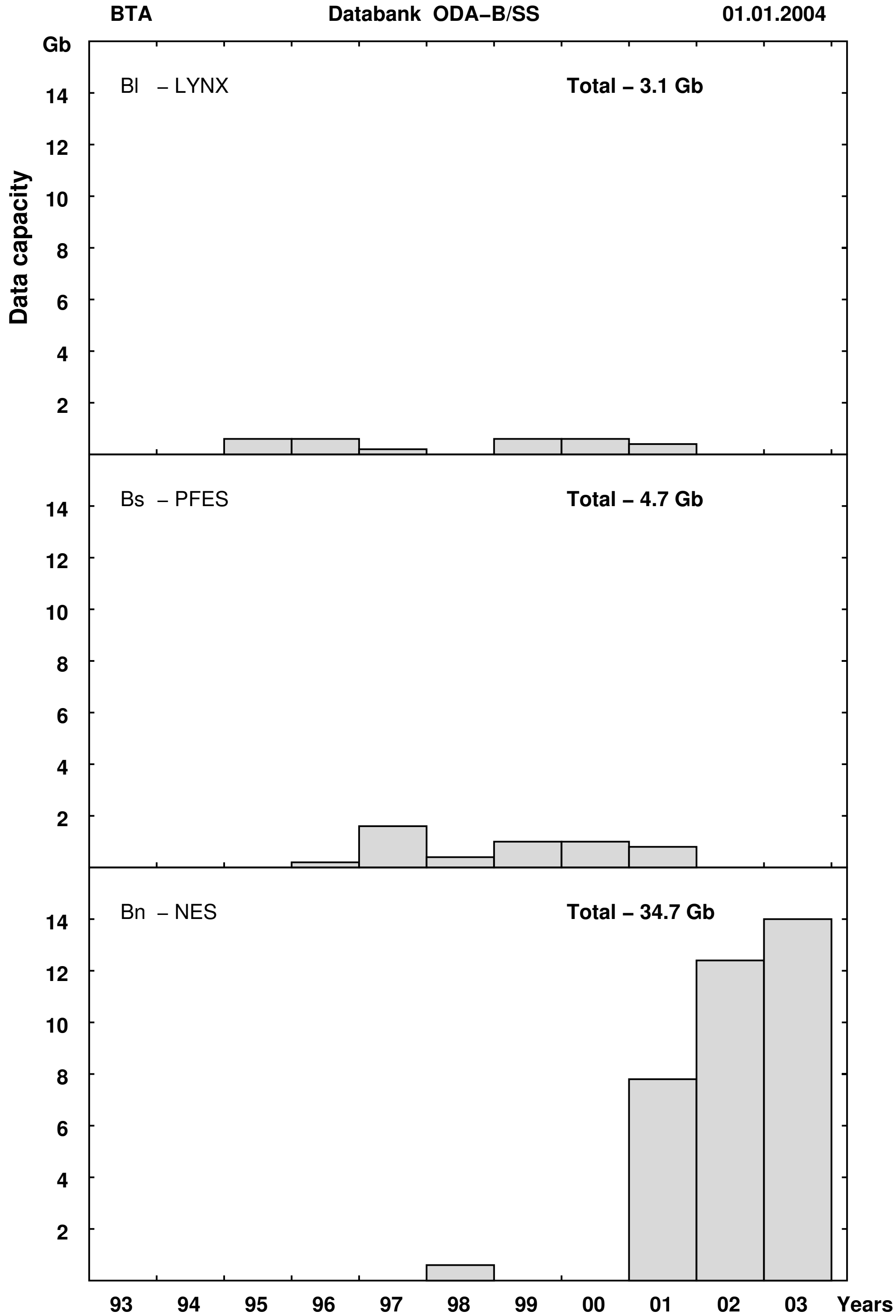

Рис. 12: Распределение по годам объемов данных для 3-х локальных архивов.

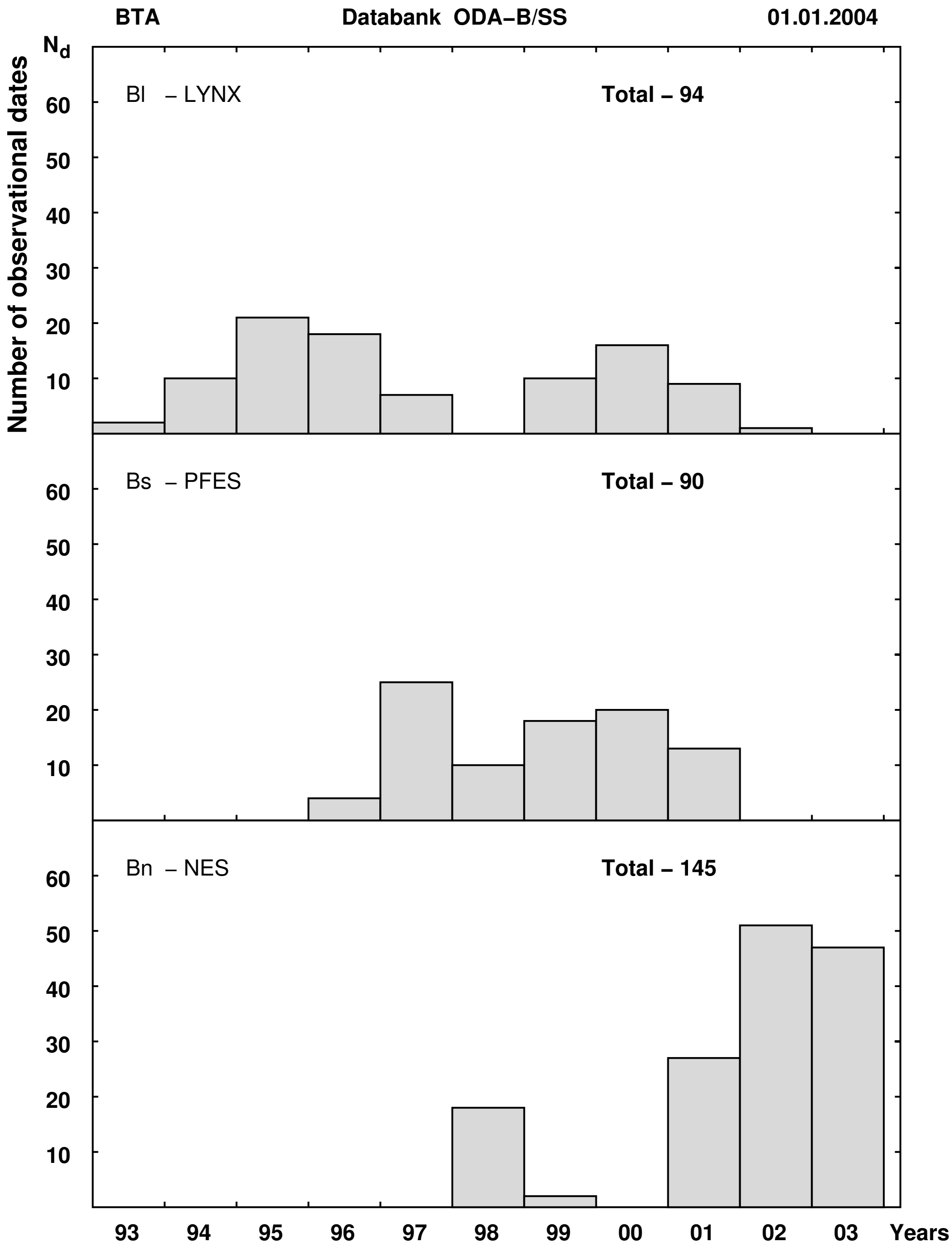

Рис. 13: Распределение по годам числа наблюдательных дат для 3-х локальных архивов.

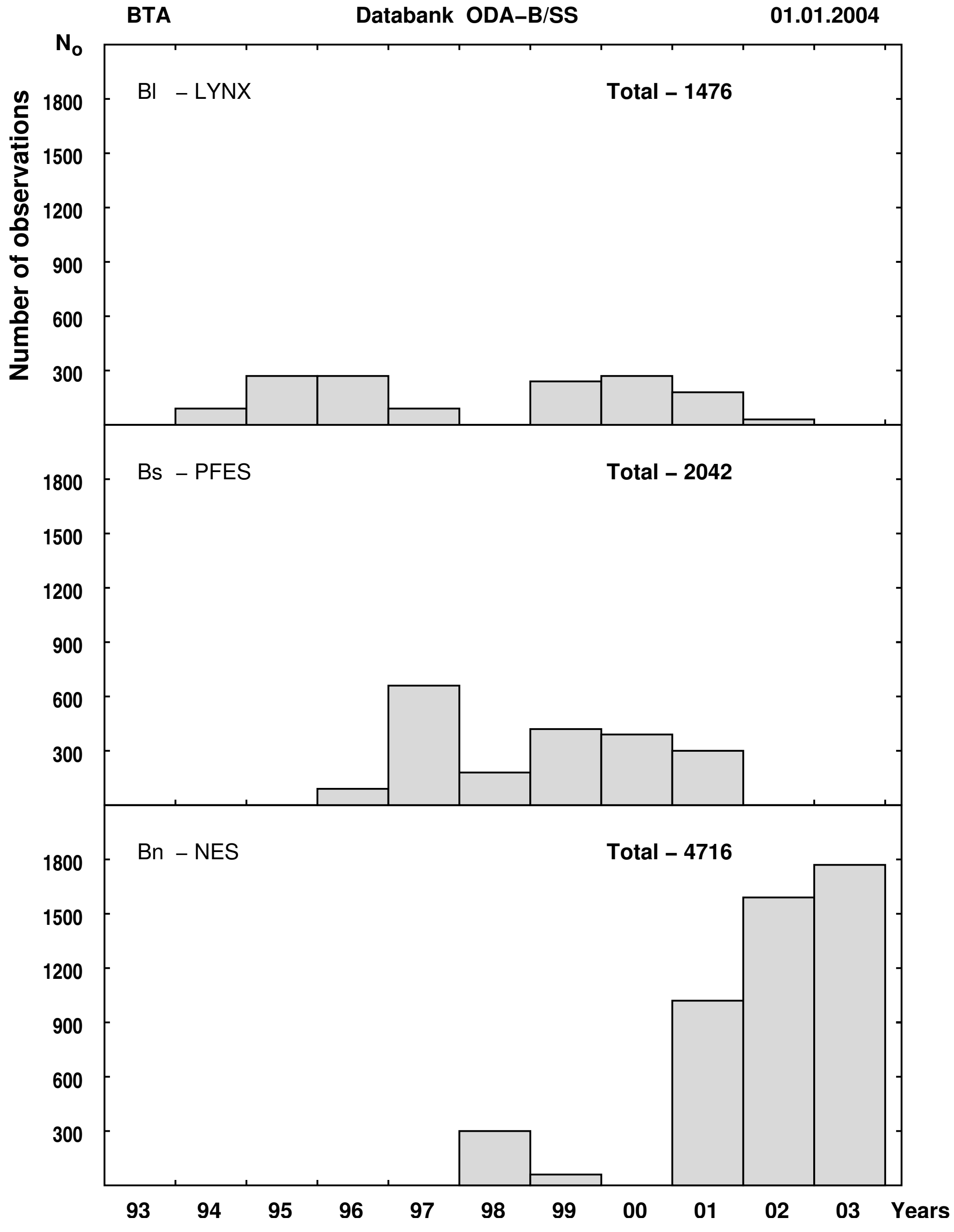

Рис. 14: Распределение по годам числа наблюдений для 3-х локальных архивов.

Список литературы

Кононов В.К., 1995а, Препринт САО РАН, **110T**, 13

Кононов В.К., 1995б, Препринт САО РАН, **111T**

Кононов В.К., 1995в, Препринт САО РАН, **110T**

Кононов В.К.: 1996а, Архивизация наблюдательных данных телескопов РАТАН–600 и БТА, Кандид. диссерт., САО РАН, 248с.

Кононов В.К.: 1996б, Препринт САО РАН, **112T**

Кононов В.К., 1996в, Препринт САО РАН, **114T**, 15

Кононов В.К., 2002а, Препринт САО РАН, **176**, 14

Кононов В.К., 2002б, Препринт САО РАН, **177**

Кононов В.К., 2003, Препринт САО РАН, **191**

Кононов В.К., Евангели А.Н.: 1991, Сообщ. САО, **67**, 87

Кононов, Панчук (Kononov V.K., Panchuk V.E.): 2000a, Bull. Spec. Astrophys. Obs., **49**, 128

Кононов В.К., Панчук В.Е.: 2000б, Препринт САО РАН, **139T**

Кононов В.К., Панчук В.Е.: 2001, Препринт САО РАН, **163**

Кононов В.К., Клочкова В.Г., Панчук В.Е.: 1996, Препринт САО РАН, **115T**

Ханиш и др. (Hanisch R.J., Farris A., Greisen E.W., Pence W.D., Schlesinger B.M., Teuben P.J., Thompson R.W. and Warnock A. III), 2001, Astron. & Astrophys., **376**, 359