

ТЕХНОЛОГИЯ ПОСТРОЕНИЯ АВТОМАТИЗИРОВАННЫХ СИСТЕМ ХРАНЕНИЯ СПУТНИКОВЫХ ДАННЫХ

В.Ю. Ефремов, Е.А. Луян, А.А. Мазуров, А.А. Прошин, Е.В. Флитман

*Институт космических исследований РАН
E-mail: info@d902.iki.rssi.ru*

Описывается разработанная в ИКИ РАН технология построения автоматизированных систем хранения спутниковых данных. Технология предназначена для создания систем хранения спутниковых данных с целью решения различных исследовательских и прикладных задач. Представлена предлагаемая архитектура построения автоматизированных систем хранения на основе использования базовых программных компонент. Описаны базовые программные компоненты и их возможности. Представлена методика оценки требований к ресурсам, необходимым для функционирования систем хранения. Описаны схемы построения архивов исходных спутниковых данных и результаты их обработки. Приведен список систем, при построении которых была использована данная технология, и анализируется опыт их создания и эксплуатации.

Данные спутникового дистанционного зондирования Земли (далее спутниковые данные) в последние годы все активнее используются для решения различных задач, при этом во многих системах, обеспечивающих мониторинг состояния окружающей среды, эти данные и результаты их обработки стали играть достаточно важную роль. Поэтому практически во всех таких системах сегодня созданы или разрабатываются блоки работы со спутниковыми данными. Для упрощения процесса создания таких блоков и их дальнейшей эксплуатации в ИКИ РАН была разработана специальная технология, которая достаточно подробно описана в [1, 2], а также в работе *Ефремов В.Ю., Луян Е.А., Мазуров А.А.* «Технология построения автоматизированных систем сбора, обработки, хранения и распространения спутниковых данных» в настоящем сборнике. Одним из ключевых ее элементов является технология построения автоматизированных систем хранения спутниковых данных, описанию которой и посвящена настоящая работа.

Реализация конкретной системы хранения спутниковых данных (СХСД), безусловно, определяется типами используемых в системе спутниковых данных и результатов их обработки, а также специфическими для данной системы требованиями по организации работы с данными и доступа к ним. Однако основные задачи, решаемые различными системами хранения, схожи, что позволяет выработать общие принципы и создать базовое программное обеспечение для построения подобных систем.

В настоящей работе мы проанализируем особенности систем архивации спутниковых данных и сформулируем основные требования к функционированию таких систем, рассмотрим архитектуры построения автоматизированных СХСД и кратко опишем характеристики разработанных в ИКИ базовых программных компонент для построения СХСД.

Особенности систем архивации спутниковых данных

Рассмотрим особенности спутниковых данных, влияющие на организацию их хранения и обеспечение к ним оперативного доступа. К таким особенностям, в первую очередь, относится их большой объем. Так, например, объем данных, полученных за один сеанс, может составлять несколько гигабайт. Хотя при современных возможностях вычислительной техники работа с данными таких объемов не вызывает больших сложностей, существующие сегодня глобальные компьютерные сети в подавляющем большинстве случаев не позволяют оперативно передать такой объем данных всем пользователям. К сожалению, эту проблему не удастся решить, просто сжимая весь объем данных, так как при сжатии без потерь объем данных, как правило, уменьшается не более чем в два раза.

Естественный путь уменьшения объемов передаваемых данных без ухудшения их качества — это их фрагментация. Как правило, спутники при пролете получают данные о достаточно большой территории. Использование фрагментации позволяет пользователям получать данные только по интересующим их регионам, что существенно уменьшает объем передаваемых данных. Во многих случаях уменьшение объема достигается также в случае, если пользователю передаются не исходные спутниковые данные, а результаты их обработки. Однако для этого обычно требуется организация устойчивой автоматической обработки данных. Кроме этого, в большинстве случаев пользователю сначала может быть предоставлена возможность получения некоторого обзорного изображения (quick look), на основе которого он может принять решение, какие конкретные данные ему необходимы, не скачивая больших объемов данных. Поэтому в системах хранения должна формироваться и храниться подобная информация.

Кроме проблем, связанных с оперативной передачей данных пользователям, большой объем спутниковых данных приводит в общем случае к невозможности долгого хранения всех данных в режиме непосредственного доступа на жестких дисках сервера архивации. Поэтому для долговременного хранения данных необходимо использовать автономные носители информации, такие как магнитные и лазерные диски. Естественно, что при этом в системе хранения должен быть реализован удобный механизм доступа пользователей к данным, находящимся в долговременном архиве.

Другой характерной особенностью спутниковых данных является разнородность форматов представления данных различных спутниковых систем. И, хотя в настоящее время разработаны достаточно универсальные форматы хранения спутниковых данных, на практике, даже при использовании этих форматов, работа с различными типами данных имеет свою специфику. В результате для использования конкретных типов данных необходимы специализированные модули для работы с этими данными. В то же время следует отметить, что во многих случаях для работы с результатами обработки спутниковых данных можно использовать достаточно простые графические форматы, легко импортируемые в специализированное программное обеспечение, которое используется при построении различных систем мониторинга.

Основные требования к функционированию СХСД

Исходя из перечисленных выше особенностей и на основании опыта построения систем хранения спутниковых данных, в ИКИ РАН были сформулированы следующие основные требования к функционированию этих систем:

- система хранения должна быть сравнительно легко адаптируема под архивацию новых типов спутниковых данных;
- система хранения должна обеспечивать оперативный доступ к данным для широкого круга удаленных пользователей;
- система должна быть сравнительно недорогой в построении и эксплуатации;
- все процедуры, связанные с аннотацией, архивацией, хранением данных и предоставлением к ним доступа должны быть максимально автоматизированы;
- система хранения должна быть масштабируема при увеличении количества поддерживаемых архивов и числа серверов, используемых под архивы;
- система должна обеспечивать возможность хранения данных как на локальных дисках, так и на автономных носителях информации, таких как CD-диски, магнитные ленты;
- отдельные компоненты СХСД могут находиться в территориально разнесенных информационных центрах;
- должен быть реализован удаленный контроль за функционированием СХСД.

Архитектура построения автоматизированных СХСД

Для создания систем хранения спутниковых данных, позволяющих решать вышеперечисленные задачи, была разработана общая архитектура построения автоматизированных СХСД (см. рисунок). В рамках архитектуры основные задачи решаются на основе использования базовых программных компонент: системы ведения архивов данных, системы администрирования архивов, системы доступа к данным на основе Web-интерфейсов и системы контроля за функционированием СХСД.

Система ведения архивов отвечает за организацию архивов спутниковых данных. Архив спутниковых данных реализуется в виде базы данных, содержащей описание спутниковых данных, локального файлового хранилища на жестких дисках сервера архивации и долговременного хранилища файлов данных на магнитных лентах, лазерных дисках и т. п. Для построения архива спутниковых данных был разработан программный пакет FDB (File Data Base) [1, 2]. Он позволяет на основе использования специализированных FDB утилит проводить операции с файлами данных в архиве согласованно с соответствующими изменениями в базе данных.

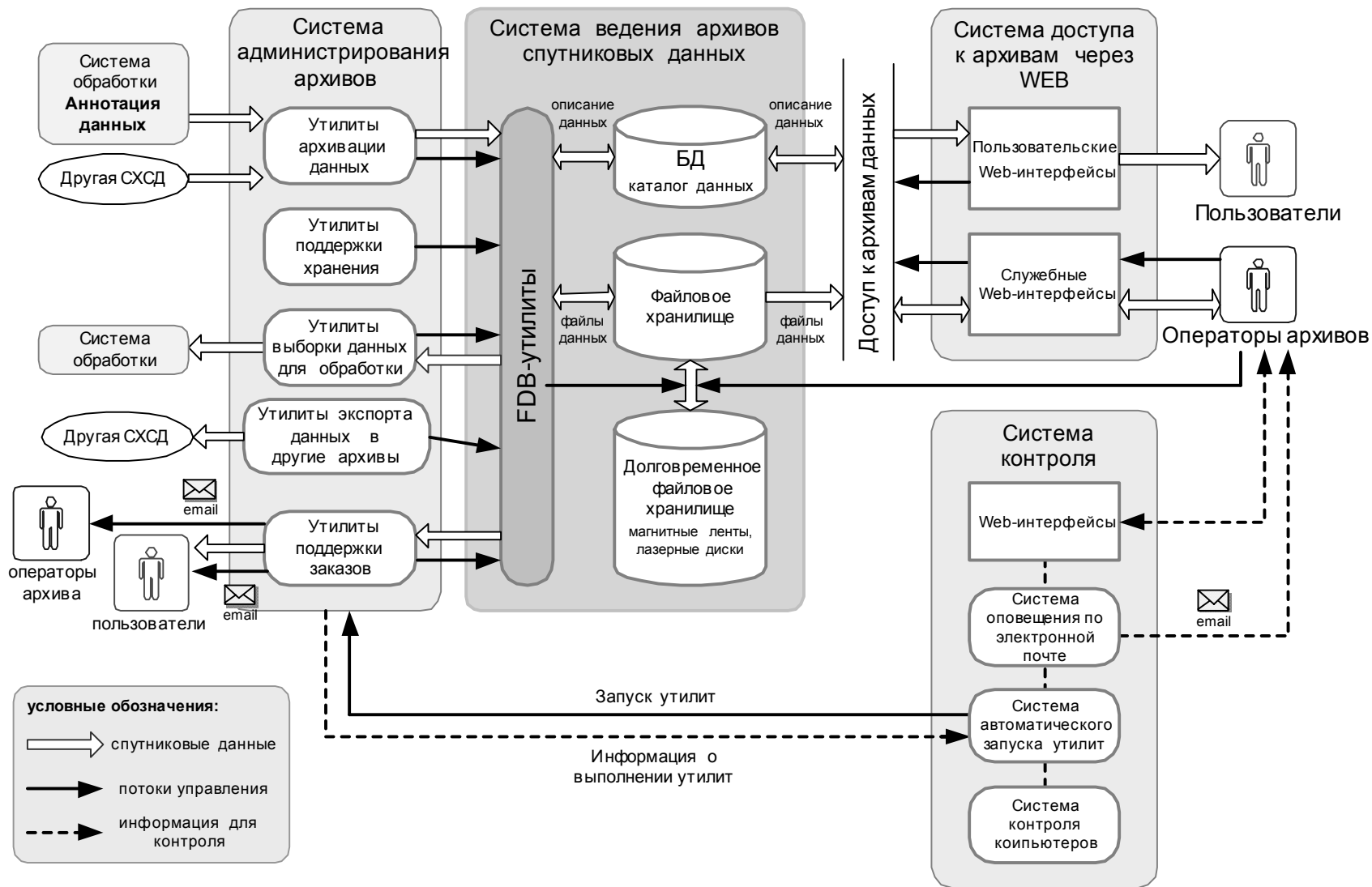
Система администрирования архивов предназначена для автоматизированной архивации спутниковых данных, их хранения, предоставления данных системам обработки, поддержки выполнения заказов пользователей и выполнения любых других задач, связанных с хранением файлов данных в архивах. За выполнение каждой из этих задач отвечают соответствующие группы утилит администрирования, представленные на рисунке.

Система доступа к архивам на основе Web-интерфейсов предоставляет пользовательские и служебные Web-интерфейсы для доступа к данным в архивах. Пользовательские интерфейсы предназначены для обеспечения доступа к архивам спутниковых данных и доступны для любого пользователя Интернет. Служебные Web-интерфейсы используются операторами архивов для выполнения различных операций, связанных с поддержкой архивов спутниковых данных, таких как визуальная оценка качества спутниковых снимков, регистрация пользователей, контроль за выполнением заказов пользователей и т. п.

Система контроля за функционированием СХСД позволяет удаленно контролировать выполнение автоматизированных процедур и производить мониторинг таких параметров используемых компьютеров, как объем свободного места на дисках, загруженность процессора, объем используемой оперативной памяти и прочих. Реализована также система активного оповещения оператора архива о возникших сбоях в работе системы по электронной почте.

Ниже показано, как на основе использования базовых программных компонент решаются основные задачи системы хранения: аннотация и архивация спутниковых данных, поддержка хранения спутниковых данных, обеспечение доступа к данным и контроль за функционированием СХСД. На рисунке сплошными стрелками обозначены потоки управления, пунктирными — потоки информации, используемой для контроля функционирования отдельных компонент СХСД, полыми стрелками — потоки спутниковых данных.

Аннотация и архивация спутниковых данных. Процедура аннотации предназначена для подготовки спутниковых данных и продуктов их обработки к архивации. В рамках этой процедуры, как правило, производится фрагментация данных, изготовление обзорных снимков и создание файлов аннотации, содержащих сопроводительную информацию о спутниковых данных.



Архитектура построения автоматизированной системы хранения спутниковых данных

Процедура архивации спутниковых данных включает входной контроль, предварительную подготовку и усвоение данных в архиве. Программно она реализуется при помощи специальных утилит архивации. Сначала поступившие аннотированные данные проверяются на комплектность наборов файлов, после чего производится разбор и проверка по формату файлов аннотации. Проверенные и подготовленные данные заносятся в архив при помощи утилит архивации. При этом в базу данных заносится вся описательная информация о новых данных, а сами спутниковые данные помещаются в локальное файловое хранилище.

Поддержка хранения спутниковых данных. Ключевым вопросом поддержки хранения спутниковых данных является организация хранения файлов данных в режиме непосредственного доступа на локальных дисках сервера архивации и в режиме долговременного хранения на автономных носителях информации, таких как лазерные диски и магнитные ленты. Как правило, при добавлении данных в архив файлы данных помещаются в режим непосредственного доступа. Чтобы избежать переполнения диска, по истечении некоторого времени в зависимости от используемой схемы хранения, файлы данных или их часть перемещаются в долговременный архив, либо вообще удаляются из архива. При этом описание данных и их обзорные снимки, как правило, остаются в режиме непосредственного доступа, так что, в случае необходимости, например, по заказу пользователей, файлы данных могут быть опять восстановлены в режим непосредственного доступа. К задачам, связанным с поддержкой хранения спутниковых данных, относится также резервное копирование архива.

Обеспечение доступа к спутниковым данным. Базовый вариант доступа к данным в архиве реализуется на основе использования пользовательских Web-интерфейсов. При этом в большинстве СХСД реализуются следующие типовые интерфейсы:

- каталог имеющихся в архиве данных по годам, месяцам и т. п.;
- поиск данных в архиве по различным критериям (дата, время, координаты и т. п.);
- интерфейс просмотра выбранных данных;
- интерфейс заказа данных.

Для выполнения различных задач, связанных с поддержкой архивов спутниковых данных, реализуются служебные WEB-интерфейсы, такие как:

- интерфейсы визуальной оценки качества спутниковых снимков, позволяющие занести соответствующую информацию в архив спутниковых данных;
- интерфейс географической привязки спутниковых снимков на основе java-апплета.

Предоставление данных системам обработки осуществляется при помощи утилит выборки данных для обработки, позволяющих получать файлы данных из архива по определенному набору критериев, например, по диапазону дат или географических координат.

Контроль за функционированием СХСД. Контроль за функционированием СХСД осуществляется «Системой контроля», являющейся отдельной программной компонентой. Основной задачей системы является контроль за выполнением утилит, запускаемых в автоматическом режиме. В рамках системы контроля производится также мониторинг состояния различных характеристик компьютеров, участвующих в работе системы хранения, таких как объем свободного места на диске, загруженность процессора, объем свободной памяти и др. Вышеперечисленная информация предоставляется операторам архива посредством использования служебных Web-интерфейсов, позволяющих осуществлять удаленный контроль за функционированием системы. Оперативное оповещение операторов архива о возникновении различных ситуаций сбоя в работе СХСД производится при помощи сообщений по электронной почте.

Общие сведения о реализации базовых программных компонент

Базовые программные компоненты созданы для работы под операционной системой UNIX. В качестве сервера БД в системе хранения используется реляционная СУБД MySQL фирмы MySQL AB, а в качестве WWW-сервера — сервер Apache. Программный пакет FDB, являющийся основой системы ведения архивов, реализован на языке C. При построении остальных программных компонент использовался язык программирования Perl. Для решения типичных программных задач, возникающих при реализации программных компонент СХСД, был разработан специализированный пакет Perl библиотек SDB. Для резервного копирования данных разработаны утилиты backup_utils, а для организации передачи данных по FTP — утилиты ftp_utils.

Внедрение представленной технологии

На основе данной технологии был разработан целый ряд автоматизированных систем хранения различного назначения, которые условно можно разделить на два основных класса: системы хранения исходных спутниковых данных и системы хранения продуктов тематической обработки спутниковых данных.

Системы хранения исходных спутниковых данных в основном устанавливаются в специализированных центрах приема спутниковых данных, где необходимо обеспечить долговременную архивацию принимаемых данных. При этом к исходным спутниковым данным условно можно отнести не только сами файлы исходных данных, но и стандартные продукты их обработки различных уровней. Для обеспечения доступа к исходным спутниковым данным, как правило, используются типовые пользовательские Web-интерфейсы. В настоящее время по представленной технологии реализованы, в частности, следующие системы хранения спутниковых данных:

- Система хранения данных метеорологических спутников серии NOAA в ИКИ РАН (<http://smis.iki.rssi.ru>) [7].
- Система хранения данных российских спутников «Ресурс», «Океан-О» и «Метеор-3М» в Центре космических наблюдений (Росавиакосмос) (<http://resurs.cpi.space.ru>) [3].
- Система хранения данных российских и зарубежных спутников в НИЦ «Планета» <http://sputnik.infospace.ru> [4].
- Система архивации данных, поступающих со спутника «Метеор-3М», в Госцентре «Природа».

Системы хранения продуктов тематической обработки спутниковых данных обычно используются для решения прикладных задач мониторинга. При этом реализуются специализированные системы хранения спутниковых данных, обеспечивающие доступ к соответствующим результатам тематической обработки. К таким системам хранения предъявляются особенно высокие требования по уровню автоматизации их работы и оперативности представления данных пользователям. В этом случае, наряду с интеграцией СХСД с прикладным программным обеспечением, используемым в системе, обычно создаются также специализированные Web-интерфейсы, позволяющие удаленным пользователям получать и анализировать результаты обработки спутниковых данных без использования какого-то дополнительного программного обеспечения. Можно привести следующие примеры системы хранения продуктов тематической обработки спутниковых данных, реализованных по описанной технологии:

- Оперативный архив данных российской системы спутникового мониторинга лесных пожаров (<http://www.nffc.aviales.ru>) [5].
- Общеотраслевой архив спутникового мониторинга океанографической обстановки Госкомрыболовства (<http://xsat.vnierkh.ru>) [6].

- Система хранения спутниковых данных действующего макета спутникового мониторинга сельскохозяйственных районов (<http://www.agrocosmos.gvc.ru>).
- Система хранения спутниковых данных для обеспечения мониторинга состояния окружающей среды Подмосковского региона (проект ведется при поддержке Правительства Москвы и РФФИ № 02-07-96006) (<http://mosreg.infospace.ru>).

В заключение отметим, что отдельные элементы представленной технологии были созданы при поддержке РФФИ (проекты № 95-07-19329, 96-07-89217, 00-07-90010, 01-07-90172, 03-07-90371).

Литература

1. *Лупян Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В.* Универсальная технология построения систем хранения спутниковых данных. М.: ИКИ РАН. Препринт Пр-2024. 2000. 22 с.
2. *Loupian E., Mazurov A., Nazirov R., Proshin A., Flitman E.* A Universal Technology for Development of Satellite Data Storage Systems // 4th Intern. Symp. on “Reducing the Cost of Spacecraft Ground Systems and Operations”. 2001.
3. *Андреев М.В., Ефремов В.Ю., Гостев М.В., Дмитриев Г.А., Лупян Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В.* Система оперативного удаленного доступа к архивам данных российских природоресурсных спутниковых систем. М.: ИКИ РАН. Препринт Пр-2055. 2002. 42 с.
4. *Асмус В.В., Бурцева Т.Н., Лупян Е.А., Мазуров А.А., Милехин О.Е., Прошин А.А., Флитман Е.В.* Система для оперативного доступа удаленных пользователей к спутниковым данным центра приема НИЦ «ПЛАНЕТА» // Тез. докл. Международ. шк.-семинара по компьютерной автоматизации и информатизации ACS-2000. С. 82.
5. *Абушенко Н.А., Барталев С.А., Беляев А.И., Еришов В.В., Коровин Г.Н., Кошелев В.В., Лупян Е.А., Крашенинникова Ю.С., Мазуров А.А., Минько Н.П., Назиров Р.Р., Прошин А.А., Флитман Е.В.* Система сбора, обработки и доставки спутниковых данных для решения оперативных задач службы пожароохраны лесов России // Научные технологии. 2000. Т. 1. № 2. С. 4–18.
6. *Андреев М.В., Егоров В.А., Ильин О.В., Лупян Е.А., Мазуров А.А., Прошин А.А., Флитман Е.В.* Использование технологий построения автоматизированных систем приема, обработки, архивации и распространения данных для мониторинга состояния морей // Тез. докл. конф. «Информ. ресурсы об океане — актуальные проблемы формирования, распространения и использования в науч. исслед. и в морской деятельности». ОИР. Обнинск: ВНИИГМИ-МЦД, 2002. С. 146.
7. *Захаров М.Ю., Лупян Е.А., Мазуров А.А., Матюшкин Е.В., Назиров Р.Р., Флитман Е.В.* Построение архивов данных метеорологических спутников на основе технологий глобальных сетей ИНТЕРНЕТ. М.: ИКИ РАН. Препринт Пр.-1978. 1998. 27 с.