Организация контроля и анализа работоспособности систем автоматизированной обработки спутниковых данных с использованием BI-технологий

Д.А. Кобец, И.В. Балашов, И.Г. Сычугов, В.А. Толпин

Институт космических исследований РАН, Москва, 117997, Россия

E-mail: kobets@d902.iki.rssi.ru

В работе рассматривается опыт применения технологий Business intelligence (BI) для построения уровня контроля и анализа работоспособности систем автоматизированной обработки спутниковых данных. Приведён веб-интерфейс мониторинга хода выполнения автоматической обработки спутниковых данных. Разбирается назначение и функционал блоков системы автоматической обработки спутниковых данных на всех её уровнях (уровень подготовки данных, формирования заданий, обработки, контроля и управления). Приведены примеры интерактивных отчётных форм для анализа статистических данных, ориентированных на разные задачи и используемых для оперативного выявления неполадок в цепях обработки. Представлена структура распределённых кластеров вычислительных узлов, а также серверов и баз данных, участвующих в обработке спутниковых данных. Особое внимание уделено описанию функционирования блока контроля (в том числе OLAP-кубов (online analytical processing), являющихся его важной частью). Рассматриваются особенности собираемой с вычислительных узлов статистической информации об их состояниях. Приведён перечень проектов (сервисов анализа данных и информационных систем дистанционного мониторинга), в составе которых задействованы описываемые в статье блоки контроля и управления.

Ключевые слова: анализ спутниковых данных, системы дистанционного мониторинга, параллельные вычисления, автоматическая обработка данных, обработка данных дистанционного зондирования, BI-технологии, OLAP, интерактивная отчётность, отчётные формы

Одобрена к печати: 05.06.2017

DOI: 10.21046/2070-7401-2017-14-3-92-103

Введение

В настоящее время данные, поступающие со спутников дистанционного зондирования Земли, всё шире используются в различных областях. Поэтому обеспечивать возрастающий спрос на спутниковые данные становится возможным только с использованием автоматизированных систем обработки. Рост объёмов анализируемой спутниковой информации приводит к необходимости использования многопоточных систем обработки, базирующихся на вычислительных кластерах, в том числе и территориально распределённых. Естественным, что при росте как числа выполняемых процессов, так и задействованных вычислительных ресурсов возрастает сложность задач организации их управления, в том числе сбора и анализа информации о состоянии процессов и ресурсов.

Разработка систем обеспечения контроля работоспособности различных комплексов обработки спутниковых данных, создаваемых и поддерживаемых Институтом космических исследований Российской академии наук (ИКИ РАН), ведётся достаточно давно. Так, в начале двухтысячных годов в ИКИ РАН была разработана и начала активно использоваться система мониторинга процессов PMS (Process Monitoring System). Она позволяет отслеживать процессы архивации данных, выполнение процедур обработки на различных вычислительных узлах, состояние ресурсов и т.д. (Балашов и др., 2011; Ефремов и др., 2004). Логическим продолжением системы PMS стала технология автоматического контроля работоспособности, позволяющая собирать и хранить информацию обо всех про-

цессах и ресурсах, использующихся в таких комплексах. Подробно принципы построения этой технологии описаны, в частности, в работе (Кобец и др., 2015).

Следует отметить, что в настоящее время в системах, поддерживаемых ИКИ РАН, за сутки собирается и заносится в базы данных информация о тысячах процессов. Поэтому, не имея специальных механизмов анализа получаемой информации, уже невозможно обнаруживать сбойные или критические ситуации, организовывать их устранение и, в конечном итоге, обеспечивать управление и устойчивую работу систем обработки спутниковых данных. Для проведения эффективного анализа собираемой информации потребовалось создание специальных инструментов. Основными их задачами становятся сбор и организация хранения информации в виде, удобном для проведения её быстрого анализа, и создание удобных интерфейсов для проведения такого анализа. Отметим, что наиболее быстро развивающимися технологиями создания подобных инструментов сегодня являются BI-технологии (Business intelligence), основанные на построении хранилищ структурированной информации на основе OLAP-структур (online analytical processing) (Паклин, Орешков, 2013).

В настоящей работе представлено описание подходов, схем организации и инструментария на основе BI-технологии для анализа информации о работе распределённых комплексов обработки спутниковых данных. Данные подходы и инструменты сегодня фактически являются частью технологии автоматизированной обработки спутниковых данных, созданной и развиваемой в ИКИ РАН (Лупян и др., 2015а). В настоящее время такие подходы и инструменты используются в различных системах автоматизированной обработки спутниковых данных, в том числе в центре коллективного пользования «ИКИ-Мониторинг» (Лупян и др., 2015в).

Особенности организации систем автоматизированной обработки спутниковых данных и задачи контроля работоспособности её элементов

Системы обработки спутниковых данных, создаваемые и поддерживаемые ИКИ РАН, имеют блочную архитектуру (Лупян и др., 2011а, 2015а), так как подобная архитектура является наиболее удобной для дальнейшего масштабирования, развития и поддержки. В такой схеме каждый блок, проектируемый под решение определённого спектра задач, имеет описанный интерфейс, посредством которого осуществляется взаимодействие с другими блоками системы.

Система обработки спутниковых данных имеет в своём составе следующие основные уровни (рис. 1):

1. *Уровень подготовки данных.* К этому уровню относятся блоки формирования комплектов файлов для обработки, которые являются уникальными для каждой тематической обработки. Они отслеживают полноту и комплектность необходимых данных и формируют очередь готовых заданий.

2. *Уровень формирования заданий.* На нём для сформированных комплектов данных определяются задания для их обработки, которые уже передаются на вычислительные комплексы, проводящие обработку. На этом уровне, в том числе, отслеживается максимально допустимое одновременное количество выполнений комплектов, относящихся к заданию одного типа, отслеживается наличие свободных вычислительных ресурсов, производится мониторинг количества комплектов переданных на обработку данных (в том числе для того, чтобы предотвратить загрузку всех мощностей выполнением одного типа заданий в ущерб выполнению других заданий), определяются приоритеты заданий и т. д.

3. *Уровень выполнения заданий (обработки).* Работа этого уровня обеспечивается распределёнными кластерами вычислительных машин («обработчиков»). Каждая вычислительная машина, задействованная в кластерах обработки, обычно находится под управлением унифицированного комплекта программ, что позволяет легко расширять парк обрабатывающих компьютеров. Каждый обработчик может выполнять несколько комплектов одного задания, которые он берет вместе с заданиями (сценариями) на обработку с серверов-диспетчеров заданий. Ход обработки каждого комплекта документируется в базе данных уровня контроля.

4. *Уровень контроля и управления.* На этом уровне агрегируется информация (рис. 1, пунктирные линии) о ходе процессов обработки со всех блоков системы, что позволяет операторам, используя, в том числе, и инструменты BI-аналитики, оперативно выявлять и устранять неполадки, обеспечивая работоспособность системы в реальном времени.

Таким образом, получается логически разграниченная на уровни схема с удобным и достаточно понятным управлением потоками обработки. Система контроля позволяет в реальном времени через веб-интерфейс проводить анализ хода обработки за выбранный промежуток времени (на каких обработчиках производилась обработка, количество удачных и ошибочных обработок и т. п.).

Рис. 1. Схема автоматизированной обработки спутниковых данных

Особенности информации, получаемой о различных блоках и элементах системы, необходимых для их контроля и управления

Для осуществления контроля над системой автоматизированной обработки спутниковых данных необходимо организовать контуры документирования процессов от блоков

на всех её уровнях. Перечень контролируемых параметров, однозначно характеризующих работу каждого уровня, логически вытекает из решаемых уровнем задач:

1. *Уровень подготовки данных* – наличие исходных данных для формирования комплектов файлов на обработку, количество комплектов, которое возможно сформировать, количество уже сформированных комплектов.

2. *Уровень формирования заданий* – типы заданий, для которых сформированы комплекты файлов на обработку, количество сформированных комплектов файлов на обработку, количество комплектов находится в данный момент в обработке, количество комплектов, обработка которых может быть начата (исходя из их приоритета и максимально возможного количества комплектов, обработка которых одновременно разрешена), количество успешно обработанных комплектов, количество ошибочно обработанных комплектов.

3. *Уровень выполнения заданий* – характеристика каждого процесса, участвующего в обработке комплекта (время начала и окончания, код возврата, лог ошибок, строка системного вызова, родитель процесса, в обработке какого комплекта участвовал данный процесс), в какой последовательности процессы следовали друг за другом при обработке данного комплекта, к какому типу заданий принадлежит данный комплект, на каком обработчике происходила обработка данного комплекта, обработка каких комплектов происходила на обработчике параллельно с данным.

На основе данных, передаваемых контурами документирования процессов, необходимо сформировать и оперативно пополнять архивы статистической информации. На основе всей поступившей информации на уровне контроля и управления необходимо сформировать удобные для анализа интерактивные отчётные формы, при работе с которыми оператор сможет менять управляющие параметры для проведения анализа различных процессов, выявлять и устранять неполадки в работе системы автоматизированной обработки спутниковых данных: выявлять ошибочные комплекты (обработка которых не возможна), выявлять ошибки в сценариях обработки, выявлять ошибки в работе обработчиков, отслеживать периодичность поступления комплектов, анализировать производительность обработчиков и т. д.

Возможности использования VI-технологий при организации работы с информацией о состоянии различных блоков и элементов систем автоматической обработки спутниковых данных

На всех этапах обработки спутниковых данных контуры документирования процессов от всех уровней системы автоматизированной обработки пополняют архивы статистической информации уровня контроля и управления. В этих архивах накапливается внушительный объём статистической информации. Без специальных механизмов и инструментов оператор не в состоянии провести её эффективный анализ. Поэтому должны быть решены задачи структурирования накопленной информации и представления результатов в виде различных удобных интерактивных отчётов. VI-технологии, которые использовались

нами для решения этих задач, позволяют на основе единого механизма создавать широкий спектр интерактивных инструментов анализа различной информации, представляемой пользователю в виде отчётов (таблиц, графиков и карт). Это позволяет пользователю самостоятельно решать, какие анализируемые данные должны быть ему представлены, а также определять вид и состав этих данных. Для реализации такой возможности собранная информация должна пройти агрегацию, и на её основе формируются OLAP-кубы. OLAP-куб представляет собой многомерный массив, созданный из соединённых таблиц архивов статистической информации, структурированных по многомерному принципу. Куб имеет преимущество в скорости обработки запросов по сравнению с обычными реляционными базами данных. Высокая скорость выполнения запросов достигается за счёт того, что в центре схемы хранения данных в OLAP-кубе располагается таблица ключевых фактов, по которым делаются запросы. К таблице ключевых фактов присоединено множество таблиц с измерениями, показывающими, как факты могут быть проанализированы. Преобразование статистической информации, хранящейся в реляционных базах данных, в OLAP-куб осуществляется посредством многотабличных SQL-запросов. В зависимости от объёма анализируемой информации OLAP-кубы могут как формироваться «на лету», так и постоянно храниться на сервере и пополняться по мере появления новых статистических данных. Схема процесса построения OLAP-кубов представлена на рис. 2.

Рис. 2. Схема уровня контроля и управления системы автоматизированной обработки спутниковых данных

Пример использования созданной системы для выявления и диагностики сбоев

Уровень контроля и управления системы автоматической обработки спутниковых данных, построенный с использованием инструментов VI-технологий, на данный момент успешно внедрён в ИКИ РАН и обеспечивает обширный функционал, позволяющий оператору выявлять и устранять возникшие нештатные ситуации. В нём реализовано значительное число интерактивных отчётных форм для анализа статистических данных, ориентированных на разные задачи. Ниже на конкретных примерах мы проиллюстрируем возможности созданной системы.

Рассмотрим реальную ситуацию, возникающую в системе автоматической обработки спутниковых данных в момент введения в эксплуатацию нового обработчика. На начальном этапе необходимо удостовериться, что обработка комплектов файлов разрешённых типов заданий происходит на нём корректно. Для этого лучше всего подходит инструмент анализа в виде интерактивной таблицы, ключевые факты которой количественно характеризуют ход обработки в разрезе измерения «Статусов» (по горизонтали) и иерархии измерения «Тип задания» и «Обработчик» (по вертикали). Отфильтровав факты по измерению «Обработчиков», оставив только контролируемый (рис. 3), можно увидеть, что за анализируемое время контролируемый обработчик выполнил шесть типов заданий.

Задания	Обработчики	Количество удачных потоков	Количество ошибочных потоков	Количество пустых сессий	Количество crash потоков	Количество незавершенных потоков
calculate mean ndvi from sin aqua modifs (session)		18	0	0	0	0
calculate mean ndvi from sin modifs (session)		2	0	0	0	0
create 4dc v05 terra (session)		32	0	0	0	0
create modis ds 6 collection (session)		4	1	0	0	0
egorovs aqua weekly slide mosaic v05 (session)		6	0	0	0	0
snap_stream_ew(session)	CASTOR	46	10	438	0	0

Рис. 3. Интерактивный инструмент анализа статистики выполненных заданий (применен фильтр по измерению обработчиков)

При этом ошибочно были выполнены комплекты, относящиеся к типу заданий snap_stream_ew(session). Чтобы понять, является ли типичным такое соотношение удачно обработанных комплектов файлов для данного типа задания к ошибочным, необходимо убрать фильтр с измерения «Обработчиков» и проанализировать, как комплекты данного типа заданий обрабатывались на других обработчиках (рис. 4).

Можно наблюдать, что такое соотношение не является характерным, и ошибочно обработанные комплекты файлов для данного типа заданий обычно (на других обработчиках) составляют менее 1% от числа удачно обработанных комплектов. Таким образом, проанализировав данную интерактивную таблицу и применив всего лишь один фильтр, легко увидеть, какой тип заданий некорректно обрабатывается на контролируемом обработчике.

Рассмотрим ещё одну характерную ситуацию, когда необходимо проанализировать, как обработка комплектов файлов заданий влияет на загруженность обработчиков.

Для этого лучше всего подходит инструмент анализа, представляющий собой интерактивную гистограмму, ключевые факты которой отражают производительность обработчика в процентах (за выбранный период времени) в разрезе измерения «Обработчиков» (отложенного по оси абсцисс) и измерения «Типов заданий» (отображённого отдельными цветами, приведёнными в легенде) (рис. 5).

Задания	Обработчики	Количество удачных потоков	Количество ошибочных потоков	Количество пустых сессий	Количество crash потоков	Количество незавершенных потоков
oper 4dc interpolation modis (session)		5	0	0	0	0
snap_stream_ew(session)	CASTOR	46	10	438	0	0
	PROC-U1-B2013	162	1	0	0	0
	PROC_U1_DL1	142	1	0	0	0
	PROC_U1_DL2	212	1	0	0	0
	PROC_U1_DL3	242	0	0	0	0
	PROC_U1_NATASHA	223	0	0	0	0
	PROC_U1_YULIA	250	1	1	0	0
sst for landsat oper (session)		5 562	0	33	0	0

Рис. 4. Интерактивный инструмент анализа статистики выполненных заданий (фильтр по измерению обработчиков снят)

Рис. 5. Интерактивный инструмент, позволяющий анализировать вклад процессов обработки комплектов заданий в загрузженность обработчиков

Как можно видеть, типы заданий, использующие наибольшее количество вычислительных ресурсов, это snap_stream_ew(session) (лиловый), create 4dc v05 terra(session) (коричневый) и Historic interpolation modis channels 1, 2 (session) (васильковый). Также очевид-

но, что обработчики BLADE_09_GZ, PROC_U1_ALLA, PROC_U1_C2013, PROC_U1_DL1 и WINMOD практически не загружены. Чтобы понять, является ли типичным для этих обработчиков такая низкая загруженность, необходимо проанализировать, как каждый из них был загружен в прошлом. Для этого лучше всего подходит анализ, представляющий собой интерактивную гистограмму, в качестве ключевых фактов которой также выступает загруженность, но только вместо измерения «Обработчиков» по оси абсцисс располагается измерение «времени». Сформировав для данного интерактивного отчёта OLAP-куб под обработчик PROC_U1_C2013 (рис. 6) становится ясным, что для данного обработчика такая картина загруженности не является характерной, так как на гистограмме ясно видно, как она падала последние несколько дней. Причём очевидно, что проблемы связаны именно с обработчиком, а не с отсутствием комплектов файлов для обработки, так как задание create 4dc v05 terra(session) (коричневый) продолжало выполняться на других обработчиках, что можно видеть на предыдущем отчёте (рис. 5).

Рис. 6. Интерактивный инструмент, позволяющий анализировать вклад процессов обработки комплектов заданий в загруженность контролируемого обработчика (PROC_U1_C2013) в зависимости от времени

Разобранные выше примеры показывают, что созданный уровень контроля и управления системой автоматической обработки спутниковых данных предоставляет инструменты, позволяющие комплексно анализировать все этапы процесса обработки данных, без которых очень сложно выявить реальные причины сбоев и нештатных ситуаций.

Заключение

До внедрения блоков контроля и контуров документирования процессов, единственная информация, доступная в системе и характеризующая ход обработки – это факт наличия ошибочно обработанного комплекта файлов. Не представлялось никакой возможности проанализировать полноту использования вычислительных ресурсов или выявить причины, препятствующие обработке комплектов файлов, так как система вообще не располагала никакой статистической информацией и инструментами её анализа. После внедрения блоков контроля генерируемая ими интерактивная отчётность, позволяющая анализировать статистическую информацию, передаваемую контурами документирования процессов, предоставляет исчерпывающие сведения о ходе обработки спутниковых данных, что позволяет выявлять в цепях автоматической обработки ошибки более высокого уровня, обнаружение которых до внедрения системы не представлялось возможным. Оперативное устранение всех выявленных ошибок позволило повысить эффективность обработки данных, что положительно повлияло на уменьшения сроков их доставки к потребителю. Возможность контроля, предоставляемая данными блоками, даёт возможность расширить в будущем парк имеющихся вычислительных машин.

На данный момент система внедрена в следующих проектах.

1. Сервисы анализа данных спутниковых наблюдений для оценки и мониторинга возобновляемых биологических ресурсов ВЕГА-Science (<http://sci-vega.ru/>) Vega-PRO (<http://pro-vega.ru/>) (Лупян и др., 2011б; Барталев и др., 2012).
2. Информационная система дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства ИСДМ-Рослесхоз (<http://www.nffc.aviales.ru>) (Лупян и др., 2015б).
3. Система проведения комплексного анализа данных спутникового дистанционного зондирования в интересах исследования Мирового океана SeeTheSea (<http://ocean.smislab.ru/static/index.sht>) (Лупян и др., 2012а).
4. Система мониторинга вулканической активности Камчатки и Курил VolSatView (<http://volcanoes.smislab.ru/static/index.sht>) (Гордеев и др., 2016).
5. Отраслевая система мониторинга Федерального агентства по рыболовству ОCM Росрыболовства (<http://cfmc.ru/>) (Нестеренко и др., 2004).
6. Объединённая система работы с данными центров НИЦ «Планета» Федерального агентства по гидрометеорологии и мониторингу окружающей среды (<http://moscow.planeta.smislab.ru/>) (Лупян и др., 2014).
7. Система комплексного дистанционного мониторинга лесов Приморского края «Вега-Приморье» (<http://primorsky.geosmis.ru/>) (Лупян и др., 2016).
8. Веб-сервис глобального спутникового мониторинга сельского хозяйства VEGA-GEOGLAM (<http://vega.geoglam.ru/>).

Работа выполнялась при поддержке ФАНО в рамках темы «Мониторинг» (госрегистрация № 01.20.0.2.00164).

Литература

1. *Балашов И.В., Ефремов В.Ю., Мазуров А.А., Мамаев А.С., Матвеев А.М., Прошин А.А.* Особенности организации контроля и управления распределенных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 3. С. 161–166.
2. *Барталев С.А., Еришов Д.В., Лупян Е.А., Толпин В.А.* Возможности использования спутникового сервиса Вега для решения различных задач мониторинга наземных экосистем // Современные проблемы дистанционного зондирования Земли из космоса. 2012. Т. 9. № 1. С. 49–56.
3. *Гордеев Е.И., Гирина О.А., Лупян Е.А., Сорокин А.А., Крамарева Л.С., Ефремов В.Ю., Мельников Д.В.* Информационная система VolSatView для решения задач мониторинга вулканической активности Камчатки и Курил // Вулканология и сейсмология. 2016. № 6. С. 62–77.
4. *Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Прошин А.А., Флитман Е.В.* Управление и контроль работоспособности систем автоматизированной обработки спутниковых данных // Современные проблемы дистанционного зондирования Земли из космоса. 2004. Вып. 1. Т. 1. С. 467–475.
5. *Кобец Д.А., Матвеев А.М., Мазуров А.А., Прошин А.А.* Организация автоматизированной многопоточной обработки спутниковой информации в системах дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 1. С. 145–155.
6. *Лупян Е.А., Балашов И.В., Бурцев М.А., Ефремов В.Ю., Кашицкий А.В., Кобец Д.А., Крашенинникова Ю.С., Мазуров А.А., Назиров Р.Р., Прошин А.А., Сычуглов И.Г., Толпин В.А., Уваров И.А., Флитман Е.В.* (2015а) Создание технологий построения информационных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 5. С. 53–75.
7. *Лупян Е.А., Барталев С.А., Еришов Д.В., Котельников Р.В., Балашов И.В., Бурцев М.А., Егоров В.А., Ефремов В.Ю., Жарко В.О., Ковганко К.А., Колбудаев П.А., Крашенинникова Ю.С., Прошин А.А., Мазуров А.А., Уваров И.А., Стыценко Ф.В., Сычуглов И.Г., Флитман Е.В., Хвостиков С.А., Шуляк П.П.* (2015б) Организация работы со спутниковыми данными в информационной системе дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства (ИСДМ-Рослесхоз) // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 5. С. 222–250.
8. *Лупян Е.А., Барталев С.А., Балашов И.В., Барталев С.С., Бурцев М.А., Егоров В.А., Ефремов В.Ю., Жарко В.О., Кашицкий А.В., Колбудаев П.А., Крамарева Л.С., Мазуров А.А., Оксюкевич А.Ю., Плотников Д.Е., Прошин А.А., Сенько К.С., Уваров И.А., Хвостиков С.А., Ховратович Т.С.* Информационная система комплексного дистанционного мониторинга лесов «Вега-Приморье» // Современные проблемы дистанционного зондирования Земли из космоса. 2016. Т. 13. № 5. С. 11–28.
9. *Лупян Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В., Крашенинникова Ю.С.* (2011а) Технологии построения информационных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 1. С. 26–43.
10. *Лупян Е.А., Матвеев А.М., Уваров И.А., Бочарова Т.Ю., Лаврова О.Ю., Митягина М.И.* (2012а) Спутниковый сервис See the Sea — инструмент для изучения процессов и явлений на поверхности океана // Современные проблемы дистанционного зондирования Земли из космоса. 2012. Т. 9. № 2. С. 251–262.
11. *Лупян Е.А., Милехин О.Е., Антонов В.Н., Крамарева Л.С., Бурцев М.А., Балашов И.В., Толпин В.А., Соловьев В.И.* Система работы с объединёнными информационными ресурсами, получаемыми на основе спутниковых данных в центрах НИЦ «Планета» // Метеорология и гидрология. 2014. № 2. С. 89–97.
12. *Лупян Е.А., Прошин А.А., Бурцев М.А., Балашов И.В., Барталев С.А., Ефремов В.Ю., Кашицкий А.В., Мазуров А.А., Матвеев А.М., Суднева О.А., Сычуглов И.Г., Толпин В.А., Уваров И.А.* (2015в) Центр коллективного пользования системами архивации, обработки и анализа спутниковых данных ИКИ РАН для решения задач изучения и мониторинга окружающей среды // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 5. С. 263–284.
13. *Лупян Е.А., Савин И.Ю., Барталев С.А., Толпин В.А., Балашов И.В., Плотников Д.Е.* (2011б) Спутниковый сервис мониторинга состояния растительности («Вега») // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 1. С. 190–198.
14. *Нестеренко А.А., Романов А.А., Андреев М.В., Лупян Е. А.* Общесистемное обеспечение отраслевой системы мониторинга Госкомрыболовства // Современные проблемы дистанционного зондирования Земли из космоса. 2004. Т. 1. С. 204–212.
15. *Паклин Н., Орешков В.* Бизнес-аналитика. От данных к знаниям. СПб: Питер, 2013. 704 с.

Organization of control and performance analysis of systems for automated processing of satellite data

D.A. Kobets, I.V. Balashov, I.G. Sychugov, V.A. Tolpin

Space Research Institute RAS, Moscow 117997, Russia
E-mail: kobets@d902.iki.rssi.ru

Experience in application of Business intelligence (BI) technologies for designing of control and performance analysis of systems for satellite data automated processing is described in this paper. A web-interface was designed on the basis of this experience for monitoring the execution progress of automated processing of satellite data. Satellite data processing system blocks assignment and functionality are analyzed at every level (data preparation, task organization, processing, control and management). There are examples of interactive report forms for statistic data analysis, which are intended for different tasks and used for operational detection of malfunctions in data processing circuits. The structure of distributed processing nodes, servers and data bases involved in processing of satellite data is shown. Careful attention is paid to control block functionality description and OLAP-cubes. Special aspects of collecting from the nodes statistical data on their states are discussed. There is a list of projects (data analysis services and remote sensing information systems) employing the control and management blocks described in the paper.

Keywords: satellite data analysis, remote monitoring systems, parallel processing, automatic data processing, processing of remote sensing data, business intelligence, OLAP, interactive reports, reporting forms

Accepted: 05.06.2017

DOI: 10.21046/2070-7401-2017-14-3-92-103

References

1. Balashov I.V., Efremov V.Yu., Mazurov A.A., Mamaev A.S., Matveev A.M., Proshin A.A., Osobennosti organizatsii kontrolya i upravleniya raspredelennykh sistem dstantsionnogo monitoring (Features of remote monitoring distributed systems control and monitoring), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 3, pp. 161–166.
2. Bartalev S.A., Ershov D.V., Loupian E.A., Tolpin V.A., Vozmozhnosti ispol'zovaniya sputnikovogo servisa Vega dlya resheniya razlichnykh zadach monitoringa nazemnykh ekosistem (Possibilities of Satellite Service VEGA Using for Different Tasks of Land Ecosystems Monitoring), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, Vol. 9, No. 1, pp. 49–56.
3. Gordeev E.I., Girina O.A., Loupian E.A., Sorokin A.A., Kramareva L.S., Efremov V.Yu., Mel'nikov D.V., Informatsionnaya sistema VolSatView dlya resheniya zadach monitoringa vulkanicheskoi aktivnosti Kamchatki i Kuril (VolSatView information system for solving the tasks of Kamchatka and the Kuril Islands volcanic activity monitoring), *Vulkanologiya i seismologiya*, 2016, No. 6, pp. 62–77.
4. Efremov V.Yu., Loupian E.A., Mazurov A.A., Proshin A.A., Flitman E.V., Upravlenie i kontrol' rabotosposobnosti sistem avtomatizirovannoi obrabotki sputnikovyykh dannykh (Control and performance analysis of systems for automated processing of satellite data), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2004, Vol. 1, No. 1, pp. 467–475.
5. Kobets D.A., Matveev A.M., Mazurov A.A., Proshin A.A., Organizatsiya avtomatizirovannoi mnogopotokovoi obrabotki sputnikovoi informatsii v sistemakh dstantsionnogo monitoring (Organization of automated multithreaded processing of satellite information in remote monitoring systems), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 1, pp. 145–155.
6. Loupian E.A., Balashov I.V., Burtsev M.A., Efremov V.Yu., Kashnitskii A.V., Kobets D.A., Krashenninnikova Yu.S., Mazurov A.A., Nazirov R.R., Proshin A.A., Sychugov I.G., Tolpin V.A., Uvarov I.A., Flitman E.V., Sozdanie tekhnologii postroeniya informatsionnykh sistem dstantsionnogo monitoring (Development of information systems design technologies), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 5, pp. 53–75.
7. Loupian E.A., Bartalev S.A., Ershov D.V., Kotelnikov R.V., Balashov I.V., Burtsev M.A., Egorov V.A., Efremov V.Yu., Zharko V.O., Kovganko K.A., Kolbudaev P.A., Krashenninnikova Yu.S., Proshin A.A., Mazurov A.A., Uvarov I.A., Stytsenko F.V., Sychugov I.G., Flitman E.V., Khvostikov S.A., Shulyak P.P., Organizatsiya raboty so sputnikovymi dannymi v informatsionnoi sisteme dstantsionnogo monitoringa lesnykh pozharov Federal'nogo agentstva lesnogo khozyaistva (ISDM-Rosleskhoz) (Satellite data processing management in Forest Fires Remote Monitoring Information System (ISDM-Rosleskhoz) of the Federal Agency for Forestry), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 5, pp. 222–250.
8. Loupian E.A., Bartalev S.A., Balashov I.V., Bartalev S.S., Burtsev M.A., Egorov V.A., Efremov V.Yu., Zharko V.O., Kashnitskii A.V., Kolbudaev P.A., Kramareva L.S., Mazurov A.A., Oksyukevich A.Yu., Plotnikov D.E., Proshin A.A., Sen'ko K.S., Uvarov I.A., Khvostikov S.A., Khovratovich T.S., Informatsionnaya sistema kompleksnogo dstantsionnogo monitoringa lesov "Vega-Primor'e" (Vega-Primorie: complex remote forest monitoring information system), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2016, Vol. 13, No. 5, pp. 11–28.

9. Loupian E.A., Mazurov A.A., Nazirov R.R., Proshin A.A., Flitman E.V., Krashennnikova Yu.S., Tekhnologii postroeniya informatsionnykh sistem dstantsionnogo monitoring (Technologies for building remote monitoring information systems), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 1, pp. 26–43.
10. Loupian E.A., Mamveev A.M., Uvarov I.A., Bocharova T.Yu., Lavrova O.Yu., Mimiagina M.I., Sputnikovyi servis See the Sea-instrument dlya izucheniya protsessov i yavlenii na poverkhnosti okeana (The Satellite Service See the Sea – a tool for the study of oceanic phenomena and processes), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, Vol. 9, No. 2, pp. 251–262.
11. Loupian E.A., Milekhin O.E., Antonov V.N., Kramareva L.S., Burtsev M.A., Balashov I.V., Tolpin V.A., Solov'ev V.I., Sistema raboty s ob'edinennymi informatsionnymi resursami poluchaemymi na osnove sputnikovykh dannykh v tsentrakh NITs "Planeta" (System of operation of joint information resources based on satellite data in the Planeta Research Centers for Space Hydrometeorology), *Russian Meteorology and Hydrology*, 2014, Vol. 39, No. 12, pp. 847–853.
12. Loupian E.A., Proshin A.A., Burtsev M.A., Balashov I.V., Bartalev S.A., Efremov V.Yu., Kashnitskii A.V., Mazurov A.A., Matveev A.M., Sudneva O.A., Sychugov I.G., Tolpin V.A., Uvarov I.A., Tsentr kollektivnogo pol'zovaniya sistemami arkhivatsii, obrabotki i analiza sputnikovykh dannykh IKI RAN dlya resheniya zadach izucheniya i monitoringa okruzhayushchei sredy (IKI center for collective use of satellite data archiving, processing and analysis systems aimed at solving the problems of environmental study and monitoring), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 5, pp. 263–284.
13. Loupian E.A., Savin I.Yu., Bartalev S.A., Tolpin V.A., Balashov I.V., Plotnikov D.E., Sputnikovyi servis monitoringa sostoyaniya rastitel'nosti ("Vega") (Satellite Service for Vegetation Monitoring VEGA), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 1, pp. 190–198.
14. Nesterenko A.A., Romanov A.A., Andreev M.V., Loupian E.A., Obshchesistemnoe obespechenie otraslevoi sistemy monitoringa Goskomrybolovstva (Systemic support of the State Committee for the Fishing Industry monitoring departmental system), *Sovremennye problemy dstantsionnogo zondirovaniya Zemli iz kosmosa*, 2004, Vol. 1, pp. 204–212.
15. Paklin N., Oreshkov V., *Biznes-analitika. Ot dannykh k znaniyam* (Business analytics. From data to knowledge), Saint Petersburg: Piter, 2013, 704 p.