

Система контроля функционирования распределённого программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг»

А. А. Прошин, М. А. Бурцев, Д. А. Кобец, И. Г. Сычугов, М. В. Радченко

*Институт космических исследований РАН, Москва, 117997, Россия
E-mail: andry@d902.iki.rssi.ru*

Современные системы доступа к данным дистанционного зондирования Земли и результатам их обработки представляют собой сложные, территориально распределённые программно-аппаратные комплексы, в работе которых участвуют многие десятки компьютеров, функционирующих в автоматизированном режиме. К таким сложным комплексам, безусловно, относится созданный в 2012 г. в Институте космических исследований РАН (ИКИ РАН) центр коллективного пользования системами архивации, обработки и анализа данных спутниковых наблюдений ЦКП «ИКИ-Мониторинг» (<http://ckp.geosmis.ru/>). Для стабильного и качественного функционирования программно-аппаратного комплекса необходимо обеспечить многоуровневый контроль работы всех входящих в его состав компонент от отдельных программ до целых программных подсистем, оперативное оповещение специалистов о возникающих неполадках и отслеживание процесса их устранения. Для решения этих задач в отделе «Технологии спутникового мониторинга» ИКИ РАН была разработана многокомпонентная система контроля функционирования программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг». В статье приводятся краткие сведения о реализации этого комплекса, после чего анализируются основные задачи контроля его функционирования. Затем кратко описываются программные компоненты, предназначенные для их решения. В заключение рассматриваются направления дальнейшего развития системы.

Ключевые слова: дистанционное зондирование Земли, спутниковые данные, информационная система, система коллективного пользования, сверхбольшие массивы данных, контроль функционирования распределённых программно-аппаратных комплексов

Одобрена к печати: 13.12.2022

DOI: 10.21046/2070-7401-2023-20-1-95-110

Введение

Современные системы доступа к данным дистанционного зондирования Земли (ДЗЗ) и результатам их обработки представляют собой сложные, территориально распределённые программно-аппаратные комплексы, в работе которых участвуют многие десятки компьютеров, функционирующих в автоматизированном режиме. К таким сложным комплексам, безусловно, относится созданный в 2012 г. в Институте космических исследований РАН (ИКИ РАН) центр коллективного пользования системами архивации, обработки и анализа данных спутниковых наблюдений ЦКП «ИКИ-Мониторинг» (Лупян и др., 2015а, 2019; <http://ckp.geosmis.ru/>). С момента создания центра его функциональные возможности многократно выросли, в том числе за счёт кратного увеличения аппаратных и программных составляющих. В настоящее время в функционировании ЦКП задействовано свыше 50 серверов архивации и специализированных систем хранения данных (NAS — *англ.* Network Attached Storage), более 100 серверов (включая виртуальные машины), предназначенных для обработки спутниковых данных, а также около 10 серверов, обеспечивающих работу сервисов доступа к данным. Архивы центра содержат свыше ста типов информационных продуктов, полученных в результате обработки данных более 40 различных спутниковых приборов наблюдения. На сегодняшний день пользователям предоставляется доступ более чем к 6 Пб данных, а также к широкому спектру инструментов для их анализа и обработки. Ежедневно в архивы ЦКП «ИКИ-Мониторинг» может поступать более 4 Тб новых данных различных типов. В настоящее время в работе ЦКП также принимают участие шесть территориально разнесённых информационных центров. В частности, благодаря сотрудничеству ИКИ РАН и Научно-исследовательского центра космической гидрометеорологии «Планета» (Бурцев и др., 2019)

в работе ЦКП участвуют три центра приёма, обработки и распространения данных, расположенные в Москве, Новосибирске и Хабаровске, что даёт пользователям доступ к данным целого ряда российских спутниковых систем. Все процессы получения, усвоения, обработки и предоставления данных в рамках ЦКП полностью автоматизированы. Число таких уникальных автоматических процессов, работу которых необходимо контролировать, в настоящее время исчисляется сотнями.

Стремительное развитие ЦКП «ИКИ-Мониторинг» на протяжении последних десятилетий привело к качественному повышению требований к обеспечению контроля его работы. Предоставляемые центром возможности по работе с данными в настоящее время активно используются в целом ряде исследовательских и прикладных проектов, поэтому задача обеспечения бесперебойного функционирования ЦКП выступает одной из наиболее приоритетных. Для решения этой задачи необходимо реализовать контроль работы всех компонент программно-аппаратного комплекса ЦКП на уровнях от отдельных программ до сложных программных подсистем, входящих в его состав, а также оперативное оповещение специалистов о возникающих в его работе неполадках и отслеживание процесса их устранения.

Настоящая статья посвящена описанию сложной многокомпонентной системы контроля функционирования программно-аппаратного комплекса (далее — ПАК) ЦКП «ИКИ-Мониторинг» (Балашов и др., 2011; Кобец и др., 2017; Мамаев и др., 2008). Система основана на программных решениях, разработанных в отделе «Технологии спутникового мониторинга» ИКИ РАН с учётом многолетнего опыта построения различных систем доступа к спутниковым данным.

В статье приводятся краткие сведения о реализации ПАК «ИКИ-Мониторинг» и подробно описываются основные задачи, которые необходимо решить для обеспечения его бесперебойной и эффективной работы. Рассматриваются основные программные компоненты, разработанные в отделе «Технологии спутникового мониторинга» для решения каждой из них. В заключение приводятся оценки эффективности реализованной системы и рассматриваются основные направления дальнейшей модернизации системы контроля функционирования ПАК ЦКП «ИКИ-Мониторинг».

Краткие сведения о реализации ЦКП «ИКИ-Мониторинг»

Программно-аппаратный комплекс ЦКП «ИКИ-Мониторинг» реализован на основе использования технологий и базового программного обеспечения, разработанных в отделе «Технологии спутникового мониторинга» ИКИ РАН (Кобец и др., 2015; Лупян и др., 2015а; Прошин и др., 2016; Толпин и др., 2011). В соответствии с разработанной архитектурой в состав комплекса входят следующие системы:

- система сбора данных из различных источников;
- система архивации данных;
- система обработки данных;
- система обеспечения доступа к данным;
- система контроля функционирования комплекса.

Серверы, обеспечивающие работу систем сбора, архивации, обеспечения доступа к данным, а также основных элементов системы контроля функционирования комплекса, работают под управлением операционных систем (ОС) FreeBSD и Linux. В рамках системы обработки используется большой кластер серверов (почти 100) под управлением ОС Windows, на которых установлен стандартный пакет разработанного в ИКИ РАН программного обеспечения, предназначенного для проведения широкого спектра обработки спутниковых данных. При этом задачи управления процессами обработки данных и диспетчеризации потоков данных реализованы на выделенных серверах под управлением операционных систем семейства Unix. Часть задач по обработке данных также решается на серверах под управлением ОС Unix, что обусловлено использованием стороннего программного обеспечения, реализованного на этой платформе.

Программное обеспечение UNIX-серверов реализовано в основном на языках программирования Perl, Python и C с использованием большого количества специально разработанных библиотечных модулей. Автоматический запуск программ на серверах производится по расписанию на основе стандартного Unix-сервиса cron. Частота запуска конкретных программ, естественно, зависит от их назначения. К примеру, пакетные процедуры архивации данных и актуализации состояния распределённых архивов запускаются каждую минуту, а программы для создания резервных копий баз данных — один раз в сутки. На серверах используются СУБД (системы управления базами данных) MySQL, MariaDB, а также PostgreSQL и веб-серверы Apache и nginx. Хранение данных в архивах реализовано на базе кластера UNIX-серверов и специализированных систем хранения данных, которые в рамках одного информационного центра объединяются в единое дисковое пространство с использованием протокола NFS (*англ.* Network File System).

Программный пакет для обработки спутниковых данных, устанавливаемый на Windows-серверах, разработан на языках C и C++ с использованием свободно распространяемого программного обеспечения. Кроме того, на таких серверах установлена специальная программа, которая периодически обращается к управляющим Unix-серверам для получения очередного назначенного на выполнение задания по обработке спутниковых данных. После получения такого задания программа запускает его выполнение (сессии) в многопоточном режиме, при этом количество потоков зависит как от свойств конкретной обработки, так и от имеющихся на сервере аппаратных ресурсов. В рамках каждого из потоков последовательно запускаются отдельные процессы, предназначенные для скачивания на сервер требуемых данных, проведения различных этапов их обработки и передачи результатов на сервер архивации.

При разработке программного обеспечения особое внимание уделялось его унификации, что позволяет использовать одни и те же программные компоненты на большом числе различных серверов, в том числе территориально удалённых и развёрнутых на разных платформах. Потоки данных и метаданных между информационными узлами распределённых архивов реализованы на основе использования протоколов FTP (*англ.* File Transfer Protocol) и SFTP (*англ.* Secure File Transfer Protocol).

Основные задачи системы контроля

Спектр основных задач системы контроля сложного ПАК включает в себя:

- мониторинг и контроль процессов, протекающих в системе;
- детектирование и дальнейший контроль сбоев и отказов компонентов комплекса;
- ведение справочной информации о компонентах комплекса.

Рассмотрим эти группы задач подробнее.

Контроль выполнения всех процессов в распределённом программно-аппаратном комплексе ЦКП «ИКИ-Мониторинг» является наиболее очевидной задачей системы контроля. Под процессами в данном случае понимаются все автоматические запуски программ на контролируемых серверах, не относящихся к работе ядра операционной системы и сопутствующего программного обеспечения, т.е. все пользовательские процессы. Прежде всего это программы, предназначенные для получения, обработки, архивации и обеспечения доступа к спутниковым данным и данным из других источников. Кроме этого, важную роль играют процессы, обеспечивающие выполнение различных служебных задач, таких как создание резервных копий данных, монтирование сетевых дисковых разделов, диспетчеризация потоков данных и др. Ещё одна группа процессов отвечает за процедуры, предназначенные непосредственно для контроля тех или иных характеристик работы серверов, входящих в состав распределённого ПАК. К таким контрольным процессам, в частности, относятся программы, отслеживающие степень заполнения дискового пространства, целостность установленных на сервере программных компонент, работу сервисов синхронизации времени и многие другие. Функции контроля могут также выполнять и сервисные процессы. К примеру, программа, отвечающая за монтирование внешних дисковых разделов, используется также для контроля их

доступности. Состав и количество автоматически выполняемых программ на каждом из серверов зависит от специфики решаемых на нём задач, при этом в типичном случае на одном сервере производятся тысячи запусков десятков различных процедур в сутки. Кроме того, задача контроля выполнения процессов сильно осложняется тем, что в состав программно-аппаратного комплекса входят компьютеры, работающие под управлением различных операционных систем и выполняющие разные функции.

Для обеспечения контроля всех выполняемых процессов необходимо вести протоколы их выполнения и фиксировать их коды возврата, сигнализирующие либо об успешном их выполнении, либо о тех или иных ошибках, возникших в процессе их работы. Естественно, что для контроля многих десятков и даже сотен различных процессов их коды возврата должны быть стандартизированы. Важно отметить, что контроль выполнения элементарных процессов реализуется на базовом уровне рассматриваемой системы, на основе которого решаются и более сложные задачи, связанные с контролем работы ключевых компонент и подсистем программно-аппаратного комплекса.

Второй ключевой задачей является автоматическое детектирование критических неполадок в работе программно-аппаратного комплекса (сбоёв). Как уже было отмечено выше, в рамках функционирования распределённого ПАК ЦКП «ИКИ-Мониторинг» в сутки на каждом из серверов в среднем производится более 1000 запусков различных процедур. В результате суммарное число запусков процессов для всего комплекса в сутки заведомо превышает сотню тысяч. Как показывает многолетний опыт разработки и эксплуатации сложных распределённых систем, контроля на уровне отдельных процессов для поддержки бесперебойной работы комплекса оказывается явно недостаточно. Связано это с целым рядом причин. Во-первых, эффективный контроль выполнения процессов на многих десятках серверов практически невозможен без централизованных решений, позволяющих агрегировать и анализировать информацию об ошибках на серверах. Во-вторых, ошибки при выполнении различных процессов не являются одинаково критичными для функционирования комплекса и должны быть классифицированы по типу и ранжированы по степени значимости. В-третьих, отсутствие явных ошибок на том или ином сервере ещё не гарантирует, что все выполняемые на нём задачи выполняются должным образом. Это может быть связано как с программными ошибками, так и с влиянием широкого спектра факторов, в том числе внешних. Например, требуемых данных может не оказаться в архивах из-за неполадок в центрах распространения данных. Это, в свою очередь, приводит к тому, что для эффективного контроля функционирования комплекса целесообразно отслеживать различные интегральные характеристики его работы, такие как своевременное поступление требуемых наборов данных в архивы, допустимое количество ошибок при проведении того или иного процесса обработки данных, доступность серверов по сети и т.п. В результате возникает необходимость в создании централизованных механизмов автоматического детектирования различных типов сбоев в работе комплекса, вызываемых программными ошибками, аппаратными проблемами и различными внешними факторами. При этом практика показывает, что от того, насколько исчерпывающим и информативным является набор автоматически контролируемых критических неполадок в работе комплекса, напрямую зависит качество поддержки его бесперебойного функционирования.

Следующей неотъемлемой задачей системы контроля является документирование сбоев, детектированных в работе комплекса, оповещение о них специалистов и контроль их устранения. Только систематизированный учёт всех обнаруженных неполадок и процесса их своевременного устранения может обеспечить качественную поддержку бесперебойной работы сложного программно-аппаратного комплекса. Для организации эффективной работы по устранению сбоев, детектированных в работе комплекса, информация о них должна поступать в централизованную систему ведения сбоев, позволяющую оперативно оповещать о них ответственных специалистов и отслеживать процесс их устранения. Кроме автоматически детектируемых сбоев в эту систему должны поступать сообщения об ошибках, обнаруженных пользователями ЦКП «ИКИ-Мониторинг», а также специалистами, задействованными в развитии и поддержке центра. Наиболее эффективным является адресное на-

правление сбоя непосредственно тем специалистам, которые могут их устранить. Однако это не всегда возможно, и, кроме того, в процессе устранения ошибок могут возникнуть обстоятельства, требующие участия и других специалистов. Поэтому система ведения сбоев также должна поддерживать совместную работу специалистов и коммуникацию между ними.

Задача документирования аппаратных и программных компонент и другой нужной информации непосредственно вытекает из необходимости реализации централизованного контроля работы распределённого ПАК. По сути, задача сводится к ведению набора справочников компонент и параметров ПАК и к реализации веб-интерфейсов для работы с этими данными. Существенно, что возможности ЦКП «ИКИ-Мониторинг» в настоящее время используются в рамках целого ряда научных и прикладных проектов, поэтому в справочниках также содержится информация о том, в интересах каких проектов используются те или иные компоненты. В первую очередь для работы системы контроля требуется детальная информация о компьютерах, входящих в состав распределённого ПАК. Также необходимой является информация о базах данных, используемых в работе ЦКП «ИКИ-Мониторинг». Кроме этого, в справочниках содержится информация о специалистах, осуществляющих поддержку функционирования работы комплекса, с указанием их зоны ответственности по различным направлениям работ. Ещё одним актуальным для задачи контроля типом информации является каталог реализованных в рамках различных проектов веб-интерфейсов, как пользовательских, так и служебных, включая те, что непосредственно предназначены для поддержки бесперебойной работы комплекса.

Решение вышеперечисленных задач в целом позволяет контролировать бесперебойную работу программно-аппаратного комплекса. Однако существует ряд дополнительных задач, решение которых может обеспечить более качественную поддержку его функционирования. К таким, в частности, относится задача формирования ежедневных сводных отчётов о работе наиболее критичных компонент комплекса, которые отсылаются специалистам по электронной почте к началу рабочего дня. Отчёты должны содержать краткую агрегированную информацию, позволяющую быстро оценить работу по соответствующему направлению, и в первую очередь предназначены для руководителей, которым необходимо контролировать множество различных направлений работ. Состав и содержание отчётов могут меняться в соответствии с текущими приоритетными задачами и проектами. В качестве примера такого отчёта можно привести сводную информацию о поступлении наиболее критичных типов информации в архивы ЦКП «ИКИ-Мониторинг». Ещё одной дополнительной задачей, которая позволяет лучше контролировать работу комплекса, является реализация различных инструментов для детального анализа функционирования отдельных его подсистем. Такие инструменты, как правило, реализуются в виде веб-интерфейсов и позволяют специалистам быстрее разобраться в причинах возникших неполадок, выявить скрытые ошибки и найти «узкие» места в работе сложных подсистем. В частности, необходимы инструменты, позволяющие проанализировать время выполнения различных процедур обработки спутниковых данных на разных серверах обработки, оценить степень загруженности имеющихся вычислительных ресурсов и проследить за динамикой заполнения имеющегося дискового пространства. К инструментам анализа также относятся интерфейсы, позволяющие получить разнообразные статистические данные. В качестве примеров можно привести формы для получения информации в виде таблиц о наличии в архивах разных типов спутниковых данных и интерфейсы, позволяющие визуально оценить пространственное заполнение данными заданной области интереса.

Контроль выполнения процедур

Контроль выполнения процедур на Unix-серверах реализуется на основе использования специально разработанного для этого программного пакета PMS (*англ.* Process Monitoring System). Этот пакет в обязательном порядке устанавливается на все такие сервера вне зависимости от их конкретного назначения. Все контролируемые процессы запускаются на сервере при

помощи специальной программы, которая заносит протоколы их выполнения и коды возврата в специализированную базу данных. При реализации всех запускаемых процедур используется стандартный набор кодов возврата, основные из которых сигнализируют о следующих результатах выполнения процесса:

- успешное выполнение;
- ошибка при выполнении;
- холостой запуск;
- блокировка запуска другим процессом;
- прекращение выполнения по сигналу.

Также для каждого процесса могут быть заданы критерии детектирования состояния ошибки в его работе: отсутствие успешных запусков за заданный интервал времени, повторяющиеся в течение указанного времени ошибки, превышение заданного максимального времени выполнения.

Для доступа к контрольной информации и выполнения ряда операций по управлению запусками процессов создан специализированный веб-интерфейс, в котором реализованы следующие основные возможности:

- конфигурация и описание процессов, включая:
 - название, принадлежность к группам процессов,
 - критерии детектирования состояния ошибки и их критичность,
 - способы оповещения специалистов;
- предоставление информации о протоколах выполнения и кодах возврата для серии запусков процессов за заданное последнее время;
- различные режимы фильтрации отображаемой информации, в частности:
 - отображение процессов в различных состояниях (сбойные, приостановленные, выполняемые в текущий момент);
 - предоставление информации о запусках процессов с заданными кодами возврата (успешные, ошибочные и др.);
- возможность остановки или временного прекращения запуска указанного процесса;
- возможность остановки или аварийного завершения процессов, выполняемых дольше установленного интервала времени;
- сводная статистика по выполнению процессов за указанный диапазон дат.

На *рис. 1* (см. с. 101) приведён пример использования основного интерфейса системы PMS, предназначенного для получения информации о последних запусках выбранных процессов, а на *рис. 2* (см. с. 101) — пример интерфейса для описания процесса.

Контроль выполнения процедур обработки данных на Windows-серверах обработки реализован на основе протоколов системы управления многопоточковой обработкой спутниковых данных. Программа запуска заданий по обработке спутниковых данных, установленная на Windows-серверах, протоколирует информацию обо всех выполняемых процессах в централизованную базу данных, предназначенную для этой цели. При этом для каждого из элементарных процессов фиксируется следующая информация: идентификатор сервера, тип обработки, идентификатор сессии, номер потока выполнения, время начала и окончания, а также код возврата. В случае возникновения ошибки при выполнении процесса фиксируется сообщение о её причинах, а выполнение соответствующего потока выполнения задания по обработке данных прекращается. Следует отметить, что в настоящее время за сутки в централизованную базу данных поступает информация о сотнях тысяч элементарных процессов, относящихся к различным типам обработки и выполняемых на разных серверах, в сессиях и потоках. В результате использование обычных веб-интерфейсов, наподобие тех, что применяются для контроля процессов, выполняемых на Unix-серверах, оказывается практически невозможным. Таким образом, возникает необходимость в интерфейсах, позволяющих просматривать информацию о выполнении процессов в различных агрегированных представлениях.

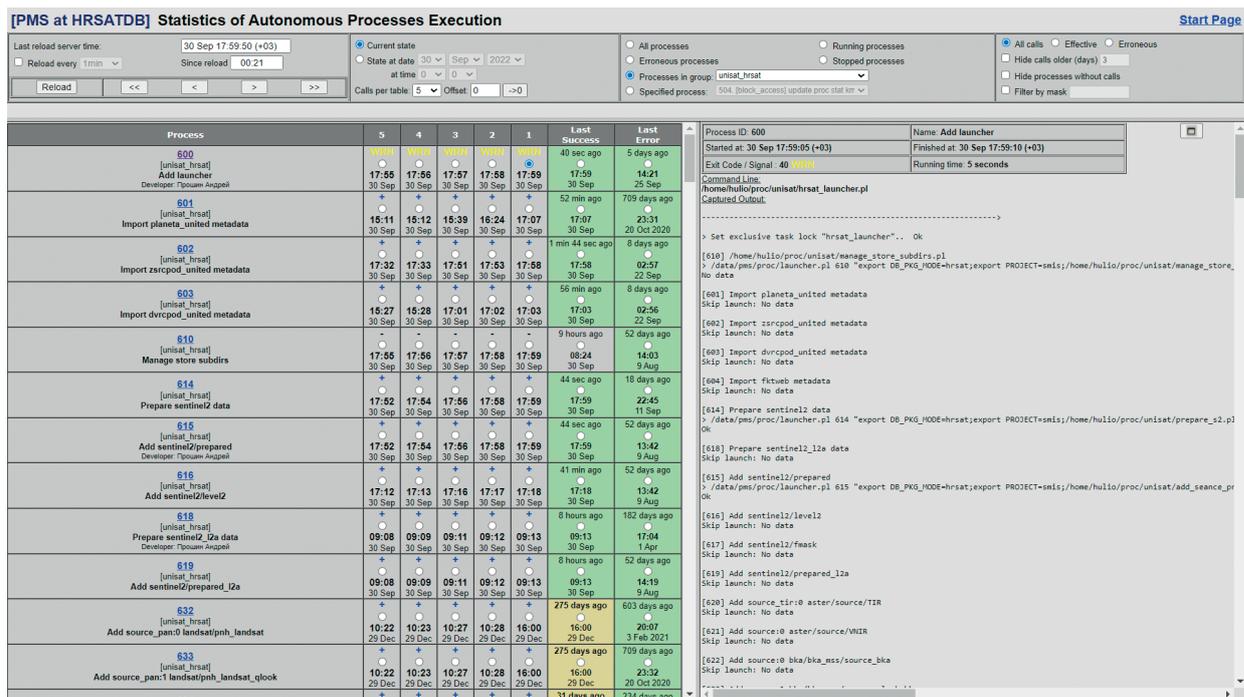


Рис. 1. Интерфейс системы PMS. Просмотр информации о последних запусках процессов

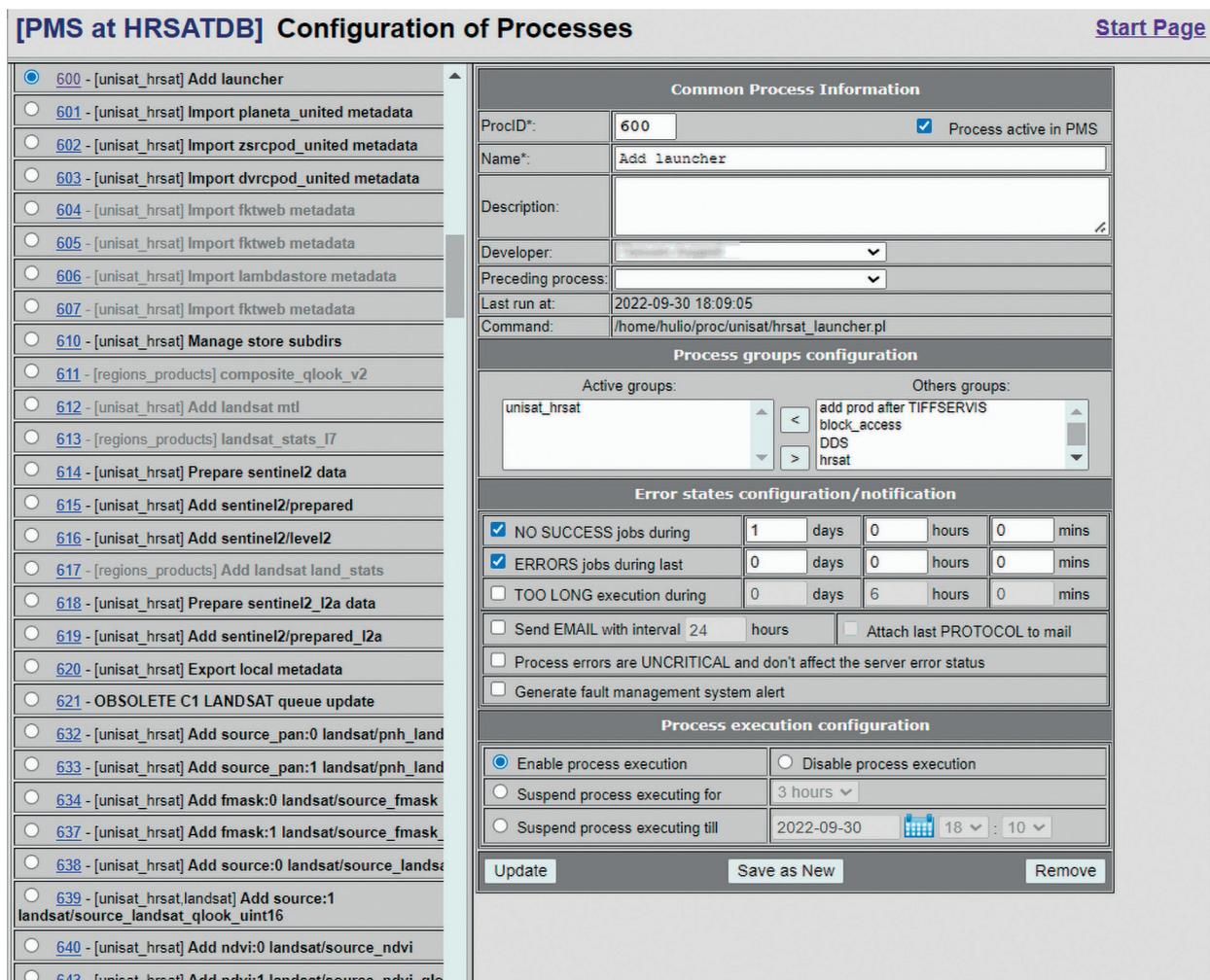


Рис. 2. Интерфейс системы PMS. Конфигурация процессов

Группы заданий	Задания	Обработчики	.rdy	Удачных	Продолжительность сессий [мин]	Среднее количество потоков в сессии	Среднесессионная продолжительность потока [мин] min	Среднесессионная продолжительность потока [мин] avg	Среднесессионная продолжительность потока [мин] max	Процесс	Ошибки
регулярные-приоритетные	alpha Unisat composite qlook (session)		724	819	4 643	3	0,2	5,7	49,6	1	
	AM_SIN_1DC daily composite oper (session)			46	55	1	0,6	1,2	2,2		
	Burns 250 TV product (session)		3								
	Calc vegetation indexes from AQUA weekly composite (session)			94	56	1	0,2	0,6	2,8		
	Calc vegetation indexes from weekly composite (session)			132	78	1	0,2	0,6	3,2		1
	Calculate mean ndvi from sin 193.232.9.113 (session)	BLADE_06_GZ									
		BLADE_10_GZ		1	1	1	0,8	0,8	0,8		
		BLADE_11_GZ		6	8	2	0,4	1,4	3,1		
		BLADE_12_GZ		1	1	1	1,1	1,1	1,1		
		BLADE_15_GZ		6	9	2	0,7	1,5	2,1		
		BLADE_16_GZ		10	14	3	0,4	1,4	3,1		
		P-ALLA-28									
		P-B2013-183									
		P-GACRUX-75									
		P-HATASHA-29		11	7	6	0,4	0,6	0,8		
		P-TANIYA-154									
		P-YULIA-30		3	15	3	5,0	5,0	5,0		
		PDL-6C-DIONE		11	10	1	0,4	0,9	2,2		
		PDL-6C-ENCCLADUS		14	11	2	0,4	0,8	1,9		
		PDL-6C-MIMAS		7	7	2	0,4	1,0	1,7		
	PDL-ALPHARD-33		9	22	2	0,6	2,4	5,3			
	PDL-ALTAIR-80										
	PDL-ALTAIR-70										

Рис. 3. Пример использования BI-интерфейса для получения информации о выполнении заданий обработки спутниковых данных

В результате решение было найдено в применении современных BI-технологий (англ. Business Intelligence), использующих хранилища структурированной информации на основе OLAP-структур (англ. online analytical processing). Благодаря возможности предварительной агрегации информации по различным признакам и параметрам, эти инструменты позволяют достаточно быстро проводить анализ больших объёмов многомерных данных. На основе BI-технологий был реализован целый ряд различных интерактивных отчётных форм (таблиц, гистограмм, графиков) (Кобец и др., 2017), позволяющих визуально анализировать информацию о выполнении как элементарных процессов, так и комплексных процедур обработки. На рис. 3 приведён пример одной из таких форм для получения агрегированной информации о выполнении заданий по обработке на различных серверах.

Система ведения сбоев

Разработанная в ИКИ РАН в отделе «Технологии спутникового мониторинга» система ведения сбоев предназначена для решения двух основных задач системы контроля:

- автоматическое детектирование критических сбоев в работе комплекса;
- документирование сбоев, детектированных в работе комплекса, оповещение о них специалистов и контроль их устранения.

Универсальные системы управления проектами, позволяющие отслеживать выполнение задач специалистами, широко используются в работе самых разных организаций, поэтому при проектировании системы ведения сбоев рассматривались варианты с использованием готового программного обеспечения, в частности Redmine (<https://www.redmine.org/>) и Microsoft Project (<https://www.microsoft.com/>). Однако, как показала практика, имеющиеся программные решения недостаточно хорошо подходят для специфических задач по организации контроля функционирования распределённого ПАК ЦКП «ИКИ-Мониторинг». Несмотря на декларируемую гибкость, их использование неминуемо ограничивает разрабатываемый функционал системы. Также существенную сложность представляет собой интеграция таких систем с остальными компонентами многоуровневой системы контроля. В ре-

зультате было решено сосредоточиться на разработке собственной специализированной программной системы, что позволяет модифицировать её в соответствии с постоянно изменяющимися требованиями и обстоятельствами.

По факту система ведения сбоев является базовым инструментом, используемым специалистами для поддержки бесперебойной работы программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг». Ниже кратко рассмотрены основные программные компоненты этой системы.

Централизованная база данных системы ведения сбоев предназначена для ведения информации о сбоях, детектированных как автоматически, так и пользователями. Она позволяет сохранять следующие основные атрибуты для каждого из сбоев:

- время детектирования и устранения сбоя;
- источник сбоя, в качестве которого может выступать либо один из агентов детектирования, либо пользователь;
- описание сбоя и признак его критичности;
- логин специалиста, отвечающего за исправление сбоя;
- информацию об оповещении специалистов;
- информацию о процессе устранения сбоя;
- проект и направление работ;
- текущий статус сбоя: новый, направлен, принят, исправлен и др.

Специальные программные компоненты (агенты) предназначены для автоматического детектирования сбоев различных типов. Непосредственное занесение сбоев в систему производится при помощи специального программного интерфейса. Ниже кратко описана реализация и назначение основных используемых в настоящее время агентов.

Агент контроля наличия ошибок (КНО) позволяет опрашивать состояние специальных контрольных процессов на Unix-серверах, каждому из которых соответствует определённый тип проверок. Сбой детектируется в том случае, когда текущим состоянием контрольного процесса является «ошибка». Существенно, что протокол выполнения таких процессов должен содержать блок в формате XML с информацией, необходимой для описания конкретного сбоя, помимо обычного текста. Состав проверок, проводимых на конкретных серверах, конфигурируется при помощи специального интерфейса. Наиболее используемыми в настоящее время являются следующие типы проверок:

- контроль доступности сервера по сети;
- контроль функционирования требуемых «демонов» на сервере;
- проверка состояния дисковых массивов (RAID — *англ.* redundant array of independent disks);
- диагностика и прогнозирование переполнения дисковых разделов сервера;
- проверка синхронизации времени на сервере;
- проверка доступности по NFS всех требуемых удалённых томов архивов данных;
- проверка целостности тематического программного обеспечения.

Агент контроля наличия данных (КНД) играет очень важную роль, так как позволяет контролировать оперативное поступление в архивы требуемых для выполнения основных задач ЦКП «ИКИ-Мониторинг» информационных продуктов. Контроль реализуется посредством централизованного выполнения SQL-запросов (*англ.* Structured Query Language) к базам данных, при этом для каждого контролируемого типа данных в них указывается допустимая «давность» данных.

Агент контроля состояния обработки (КСО) предназначен для автоматического детектирования сбоев при выполнении задач по обработке спутниковых данных на кластере Windows-серверов. Программно он является частью системы управления многопоточковой обработкой спутниковых данных. Агент позволяет реализовать следующие основные типы проверок:

- контроль процента ошибок на каждом из серверов обработки;

[СДКП] Редактирование информации о сбое

Логин: Дежурный:

Общая информация

Идентификатор сбоя	59042
Кто обнаружил	Пользователь: [Имя]
Ответст. дежурный	[Имя]
Текущий статус	Принят
Время обнаружения	2022-09-29 12:58:39

Проекты

- ИСДМ-Рослесхоз
- Планета - Геофизика-Центры
- Сервис БЕГА
- БЕГА-Лес
- Тайловый интерфейс
- Вулканы
- БЕГА ЕФИС ЗСН

Центры

- ЦЭПП РАН
- ЦСМС
- ДВ РЦПОД
- ВЦ ДВО РАН
- ГАЗКОМ Щелково
- ИКИЗ
- ИКИ РАН

Проинформировать

- [Имя]

Определение типа работ и ответственного

Направление работ	Система доступа
Тип работ	Формы отчетности по пожарам (-ЛО, Ясень)
Ответственный	[Имя] Куратор направления
Критический сбой:	<input type="checkbox"/>
Изменить статус:	Принят

Описание и изменения:

Сбой от:
 Организация: ФБУ "Авиалесоохрана"
 e-mail: [адрес электронной почты]
 По адресу: Ссылка в описании
 Хочу опубликовать Акт № к-1614 проверки достоверности сведений о площади лесных пожаров с использованием данных дистанционного зондирования Земли на территории: Амурская

Новые изменения:

Сбой был занесен пользователем системы, который не может просматривать журнал сбоев.
 Чтобы послать ему письмо с новыми изменениями поставьте галочку

Рис. 5. Система ведения сбоев. Интерфейс редактирования информации о сбое

После регистрации нового сбоя сразу же производится оперативное оповещение о нём специалистов по электронной почте и через внутреннюю систему обмена сообщениями. Если в информации о сбое ответственный специалист не указан, то оповещение направляется на дежурного специалиста. В его задачи входит отслеживание всех детектируемых за сутки сбоев, перенаправление их специалистам, а также контроль их устранения. Также ежедневно всем ответственным специалистам рассылается информационное письмо с указанием актуальных для них сбоев.

Система документирования и контроля проектов

Система документирования и контроля проектов (СДКП) предназначена для ведения и предоставления специалистам информации об используемых в рамках различных проектов аппаратных и программных компонентах, включая ссылки на инструменты, предназначенные для решения различных задач контроля. По существу, она является базовым интерфейсом описываемой в настоящей статье системы контроля, в рамках которого доступны все основные интерфейсы, предназначенные для решения задач контроля функционирования программно-аппаратного комплекса, сгруппированные по четырём основным разделам. В разделе «Инструменты СДКП» представлены ссылки на наиболее часто используемые инструменты контроля. Раздел «Конфигурация проекта» содержит ссылки на интерфейсы ведения основных справочников, реализованных в описываемой системе. В третьем разделе находятся ссылки на разнообразные служебные интерфейсы, предназначенные в том числе и для решения задач контроля. Четвёртый раздел содержит ссылки на пользовательские интерфейсы, реализованные в рамках выбранного проекта.

Формирование ежедневных сводных отчётов

Ежедневные отчёты о функционировании различных компонент комплекса позволяют в компактной форме предоставить руководителям направлений и соответствующим специалистам агрегированную контрольную информацию. Такие отчёты рассылаются специалистам по электронной почте рано утром до начала рабочего дня и, как правило, содержат информацию о функционировании той или иной компоненты комплекса на протяжении предшествующих суток. Естественно, что наиболее актуальны ежедневные отчёты для контроля решения ключевых задач, от выполнения которых зависит функционирование комплекса в целом. В частности, на ежедневной основе формируются отчёты о поступлении в архивы ЦКП «ИКИ-Мониторинг» спутниковых данных различных типов. При этом такие отчёты содержат не только данные за предыдущие сутки, но и агрегированную информацию за предыдущие несколько месяцев, что позволяет проследить динамику поступления в архивы разных данных. Причинами перебоев с поступлением данных могут быть как внешние факторы, связанные с работой центров распространения спутниковой информации, так и внутренние неполадки в функционировании систем для их подготовки и архивации. Сходный отчёт ежедневно формируется и для тематических информационных продуктов, получаемых в результате проведения обработки данных разных типов, поступающих в центр. Ещё один очень важный для задач контроля ежедневный отчёт содержит информацию о текущих критических сбоях в работе комплекса с указанием ответственных за их исправление, времени с момента детектирования и типа сбоя. Также ежедневные отчёты при необходимости могут формироваться для более пристального контроля за работой компонент, которые в настоящий момент находятся в состоянии разработки, или компонент, которые наиболее критичны для выполнения текущих обязательств по различным проектам.

Инструменты для анализа функционирования компонент комплекса

Для анализа функционирования отдельных подсистем комплекса было разработано большое количество специализированных инструментов, предоставляющих возможность в интерактивном режиме исследовать различные показатели работы подсистем. Использование таких инструментов позволяет определить причины обнаруженных неполадок, выявить скрытые ошибки и найти пути для оптимизации работы различных подсистем комплекса. Наиболее востребованные на текущий момент инструменты перечислены ниже:

- Комплексная информация о состоянии и степени наполнения дисковых хранилищ.
- Информация о задержках при обработке данных по сеансам, принимаемым в различных информационных центрах.
- Реализованные на основе программного обеспечения Zabbix (<https://www.zabbix.com>) интерфейсы для мониторинга аппаратной нагрузки на серверы комплекса.
- Множество интерактивных отчётных форм, реализованных на базе VI-технологий, для анализа информации о выполнении заданий по обработке спутниковых данных на кластере Windows-серверов.
- Формы запроса статистики наличия разных типов данных в архивах с возможностью фильтрации по целому ряду различных критериев, таких как диапазон дат, географическая область, станции приёма и т. п.
- Статистика по данным, поступившим в архивы за указанный диапазон времени.
- Статистика по аппаратным сбоям серверов.
- Инструменты для визуального анализа пространственного покрытия различными типами спутниковой информации.

На *рис. 6* (см. с. 107) приведён пример интерфейса, построенного по VI-технологии и предназначенного для сравнительного анализа времени, затраченного на выполнение наиболее ресурсоёмких заданий по обработке спутниковых данных.

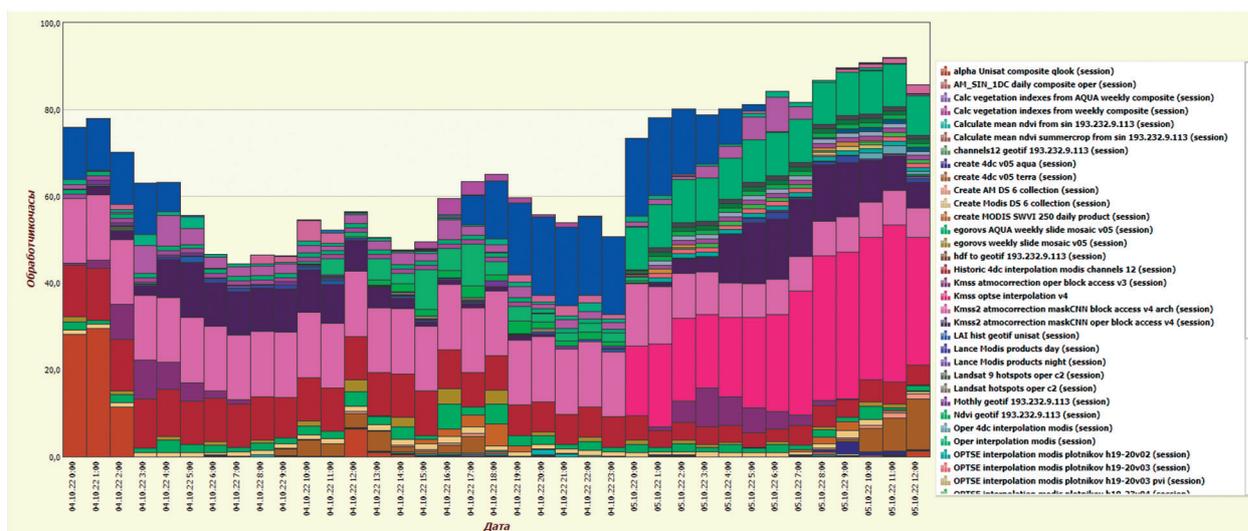


Рис. 6. Пример использования ВІ-интерфейса для сравнительного анализа времени, затраченного на выполнение различных задач по обработке спутниковых данных

Заключение

Разработанная в отделе «Технологии спутникового мониторинга» ИКИ РАН многоуровневая система контроля функционирования ПАК ЦКП «ИКИ-Мониторинг» позволяет эффективно решать поставленные перед ней задачи, что подтверждается многолетним опытом её использования. Важно отметить, что подобная система просто необходима для эффективного решения задач работы с большими данными в космических исследованиях. Использование рассматриваемой системы контроля позволяет осуществлять поддержку бесперебойной работы центра коллективного пользования усилиями менее чем десяти разработчиков, для которых эта деятельность не является основной. Это достигается благодаря максимальной автоматизации процессов детектирования сбоев в работе системы и высокого уровня их информативности. Для адаптации к постоянно меняющимся условиям и возрастающим требованиям реализованная система контроля регулярно модернизируется. Одним из основных направлений модернизации является дальнейшее повышение информативности сообщений об автоматически детектированных сбоях, что позволяет предоставить специалистам не только общие сведения о возникших неполадках, но и диагностику потенциальных причин их возникновения, а также сразу правильно адресовать сбой нужному специалисту. Ещё одним актуальным направлением является разработка и внедрение новых инструментов для анализа функционирования различных подсистем комплекса. В первую очередь речь идёт о новых интерфейсах, базирующихся на использовании ВІ-технологий, а также интерфейсов, позволяющих визуально анализировать пространственное распределение и комплектность информационных продуктов, содержащихся в архивах. Также постоянно ведутся работы по совершенствованию уже имеющихся компонент системы контроля.

Работы по развитию технологий контроля функционирования сложного распределённого программно-аппаратного комплекса ЦКП «ИКИ-Мониторинг» с 2019 г. выполняются при поддержке темы Министерства науки и высшего образования Российской Федерации «Большие данные в космических исследованиях: астрофизика, солнечная система, геосфера» (№ 122042500019-6).

Литература

1. *Балашов И. В., Ефремов В. Ю., Мазуров-мл. А. А., Мамаев А. С., Матвеев А. М., Прошин А. А.* Особенности организации контроля и управления распределённых систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 3. С. 161–166.
2. *Бурцев М. А., Успенский С. А., Крамарева Л. С., Антонов В. Н., Калашников А. В., Балашов И. В., Кашицкий А. В., Лупян Е. А., Матвеев А. М., Прошин А. А.* Современные возможности и перспективы развития Объединённой системы распределённой работы с данными НИЦ «Планета» // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 3. С. 198–212. DOI: 10.21046/2070-7401-2019-16-3-198-212.
3. *Кобец Д. А., Матвеев А. М., Мазуров А. А., Прошин А. А.* Организация автоматизированной многопоточковой обработки спутниковой информации в системах дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 1. С. 145–155.
4. *Кобец Д. А., Балашов И. В., Сычугов И. Г., Толпин В. А.* Организация контроля и анализа работоспособности систем автоматизированной обработки спутниковых данных с использованием VI-технологий // Современные проблемы дистанционного зондирования Земли из космоса. 2017. Т. 14. № 3. С. 92–103. DOI: 10.21046/2070-7401-2017-14-3-92-103.
5. *Лупян Е. А., Балашов И. В., Бурцев М. А., Ефремов В. Ю., Кашицкий А. В., Кобец Д. А., Крашениникова Ю. С., Мазуров А. А., Назиров Р. Р., Прошин А. А., Сычугов 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.
6. *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., Сычугов И. Г., Толпин В. А., Уваров И. А.* (2015б) 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.
7. *Лупян Е. А., Прошин А. А., Бурцев М. А., Кашицкий А. В., Балашов И. В., Барталев С. А., Константинова А. М., Кобец Д. А., Мазуров А. А., Марченков В. В., Матвеев А. М., Радченко М. В., Сычугов И. Г., Толпин В. А., Уваров И. А.* Опыт эксплуатации и развития центра коллективного пользования системами архивации, обработки и анализа спутниковых данных (ЦКП «ИКИ-Мониторинг») // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 3. С. 151–170. DOI: 10.21046/2070-7401-2019-16-3-151-170.
8. *Мамаев А. С., Прошин А. А., Флитман Е. В.* Создание системы документирования и контроля распределённых информационных систем // Современные проблемы дистанционного зондирования Земли из космоса. 2008. Вып. 5. Т. 2. С. 557–560.
9. *Прошин А. А., Лупян Е. А., Балашов И. В., Кашицкий А. В., Бурцев М. А.* Создание унифицированной системы ведения архивов спутниковых данных, предназначенной для построения современных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2016. Т. 3. № 3. С. 9–27. DOI: 10.21046/2070-7401-2016-13-3-9-27.
10. *Толпин В. А., Балашов И. В., Ефремов В. Ю., Лупян Е. А., Прошин А. А., Уваров И. А., Флитман Е. В.* Создание интерфейсов для работы с данными современных систем дистанционного мониторинга (система GEOSMIS) // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 3. С. 93–108.

Operation control system of the IKI-Monitoring distributed hardware and software complex

A. A. Proshin, M. A. Burtsev, D. A. Kobets, I. G. Sychugov, M. V. Radchenko

Space Research Institute RAS, Moscow 117997, Russia

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

Modern systems for remote sensing data and processing results access are sophisticated, geographically distributed hardware-software complexes consisting of many dozens of automatically operating computers. The IKI-Monitoring Center for Collective Use of systems for archiving, processing and analyzing satellite data at IKI RAS (<http://ckp.geosmis.ru/>), created in 2012, is certainly one of such sophisticated complexes. Stable and quality operation of the hardware and software system requires multi-level control of all its components, from individual programs to entire software subsystems, prompt notification of experts about system malfunction cases and their elimination process tracking. To solve these problems, the IKI RAS Department of Satellite Monitoring Technologies has developed a multi-component system to control the functioning of the hardware-software complex (HSC) of the IKI-Monitoring. The paper provides brief information on the HSC implementation, followed by its operation control main problems analysis. Next, the software components designed for highlighted problems solution are briefly described. In conclusion, further development trends of the system are discussed.

Keywords: Earth remote sensing, satellite data, information system, collaborative system, ultra-large data arrays, control of distributed software and hardware complexes functioning

Accepted: 13.12.2022

DOI: 10.21046/2070-7401-2023-20-1-95-110

References

1. Balashov I. V., Efremov V. Yu., Mazurov-m. A. A., Mamaev A. S., Matveev A. M., Proshin A. A., Features of Remote Monitoring Distributed Systems Control and Monitoring, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 3, pp. 161–166 (in Russian).
2. Bourtsev M. A., Uspensky S. A., Kramareva L. S., Antonov V. N., Kalashnikov A. V., Balashov I. V., Kashnitskii A. V., Loupian E. A., Matveev A. M., Proshin A. A., Actual features and evolution prospects of the SRC “Planeta” distributed data operation united system, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2019, Vol. 16, No. 3, pp. 198–212 (in Russian), DOI: 10.21046/2070-7401-2019-16-3-198-212.
3. Kobets D. A., Matveev A. M., Mazurov A. A., Proshin A. A., Organization of automated multithreaded processing of satellite information in remote monitoring systems, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 1, pp. 145–155 (in Russian).
4. Kobets D. A., Balashov I. V., Sychugov I. G., Tolpin V. A., Organization of control and performance analysis of systems for automated processing of satellite data, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2017, Vol. 14, No. 3, pp. 92–103 (in Russian), DOI: 10.21046/2070-7401-2017-14-3-92-103.
5. Loupian E. A., Balashov I. V., Bourtsev M. A., Efremov V. Yu., Kashnitskiy A. V., Kobets D. A., Krashe-ninnikova Yu. S., Mazurov A. A., Nazipov R. R., Proshin A. A., Sychugov I. G., Tolpin V. A., Uvarov I. A., Flitman E. V. (2015a), Development of information systems design technologies, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 5, pp. 53–75 (in Russian).
6. Loupian E. A., Proshin A. A., Bourtsev M. A., Balashov I. V., Bartalev S. A., Efremov V. Yu., Kashnitskiy A. V., Mazurov A. A., Matveev A. M., Sudneva O. A., Sychugov I. G., Tolpin V. A., Uvarov I. A. (2015b), 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 distantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 5, pp. 263–284 (in Russian).
7. Loupian E. A., Proshin A. A., Bourtsev M. A., Kashnitskii A. V., Balashov I. V., Bartalev S. A., Konstantinova A. M., Kobets D. A., Mazurov A. A., Marchenkov V. V., Matveev A. M., Radchenko M. V., Sychugov I. G., Tolpin V. A., Uvarov I. A., Experience of development and operation of the IKI-Monitoring center for collective use of systems for archiving, processing and analyzing satellite data, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2019, Vol. 16, No. 3, pp. 151–170 (in Russian), DOI: 10.21046/2070-7401-2019-16-3-151-170.

8. Mamaev A. S., Proshin A. A., Flitman E. V., Development of documentation and monitoring system for the distributed information systems, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2008, Issue 5, Vol. 2, pp. 557–560 (in Russian).
9. Proshin A. A., Loupian E. A., Balashov I. V., Kashnitskiy A. V., Burtsev M. A., Unified satellite data archive management platform for remote monitoring systems development, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2016, Vol. 13, No. 3, pp. 9–27 (in Russian), DOI: 10.21046/2070-7401-2016-13-3-9-27.
10. Tolpin V. A., Balashov I. V., Efremov V. Yu., Loupian E. A., Proshin A. A., Uvarov I. A., Flitman E. V., The GeoSMIS System: Developing Interfaces to Operate Data in Modern Remote Monitoring Systems, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 3, pp. 93–108 (in Russian).