

УПРАВЛЕНИЕ И КОНТРОЛЬ РАБОТОСПОСОБНОСТИ СИСТЕМ АВТОМАТИЗИРОВАННОЙ ОБРАБОТКИ 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

Представлена технология, которая разрабатывалась в ИКИ РАН для организации управления и контроля работоспособности распределенных системы обработки спутниковых данных. Технология включает системы контроля: состояния вычислительной техники, процессов обработки данных, потоков данных, БД и систем доступа к данным, а также системы представления информации операторам. Все эти системы рассчитаны на работу с удаленными элементами на основе Интернет-технологий. Технология также рассчитана на то, что контроль работоспособности различных элементов систем обработки спутниковых данных осуществляется удаленным оператором. Созданная технология была использована для организации систем контроля работоспособности в различных центрах приема, обработки и архивации спутниковых данных.

На протяжении последних лет в ИКИ РАН активно ведутся разработки систем автоматизированной обработки спутниковых данных (см., например, [1–5]). За это время созданы архитектура и базовые программные компоненты, использующиеся в настоящее время в различных системах мониторинга окружающей среды (см. работу «Технология построения автоматизированных информационных систем сбора, обработки, хранения и распространения спутниковых данных для решения научных и прикладных задач» в настоящем сборнике). Ключевым принципом разработки этих систем является максимальная автоматизация процедур сбора, архивации, обработки и представления данных. Это позволяет исключить участие человека в рутинных процедурах обработки данных, свести к минимуму ошибки и сбои вследствие «человеческого фактора», минимизировать время от момента поступления исходных данных до получения и архивации конечного продукта, а также снизить затраты на эксплуатацию и поддержание спутниковых информационных систем.

При реализации системы автоматической обработки данных, естественно, встает проблема мониторинга и контроля работоспособности системы обработки. На практике наиболее частыми первыми признаками неполадок или сбоев в системе обработки являются отсутствие выходного продукта в надлежащие сроки или получение неадекватного выходного продукта (плохого качества, сбойного и т. д.). Причин для возникновения сбоев в автоматической обработке данных весьма много, основные из них можно классифицировать следующим образом:

1. Сбои аппаратного и системного уровня:
 - а) «зависание» компьютера (станции обработки) или сервера;
 - б) потеря сетевого соединения (вследствие неисправности сетевого кабеля, сетевого адаптера, сбоя программной поддержки сети, зависания концентратора или коммутатора и т. д.);
 - в) пропадание питания компьютеров или сетевых устройств (неисправность источников бесперебойного питания, разряд батарей ИБП из-за долгого отсутствия напряжения в сети, неисправность кабеля питания и т. д.);
 - г) сбои операционной системы (переполнение дисков, сбои и/или выгрузка по ошибке системных процессов, переполнение доступной оперативной памяти и т. д.).
2. Сбои тематической обработки данных:
 - а) ошибки в программах тематической обработки;
 - б) ошибки в программах автоматизации обработки;
 - в) остановка программы обработки в нештатной ситуации с сообщением об ошибке, приводящая к остановке цепочки обработки;

г) ошибки в синхронизации процессов обработки, приводящие к сбоям и/или порче данных;

д) отсутствие исходных данных для обработки в результате неполадок спутниковой приемной станции, либо задержка в получении данных от внешних источников;

е) сбои и повреждения файлов исходных данных, приводящие к сбоям в работе программ обработки.

Для обеспечения успешного функционирования системы автоматической обработки данных, естественно, необходимо предусмотреть средства контроля и мониторинга, позволяющие своевременно обнаружить эти сбойные ситуации, выяснить причины их возникновения и предпринять действия по исправлению и восстановлению потоков обработки данных. При этом для обеспечения работы достаточно больших систем эти средства должны позволять в максимально автоматическом режиме обнаруживать сбои в работе системы, диагностировать их и оперативно предоставлять информацию о них операторам системы. Поэтому фактически параллельно с созданием систем обработки спутниковых данных в ИКИ РАН разрабатывались средства автоматизированного контроля их работоспособности. [6, 7]. Естественно, что по мере развития систем эти средства постоянно модифицируются и дорабатываются. Настоящая работа посвящена описанию архитектуры и возможностей версии системы контроля работоспособности, которая используется в ИКИ РАН.

Структура системы контроля работоспособности и мониторинга обработки спутниковых данных

В рамках единой архитектуры систем обработки данных, разрабатываемых в ИКИ РАН, стало возможно создать унифицированную систему мониторинга и контроля работоспособности систем обработки данных. Система решает следующие задачи:

- контроль работоспособности компьютеров (серверов и рабочих станций);
- контроль сетевой среды (сетевых соединений между компьютерами);
- контроль процессов тематической обработки данных;
- мониторинг выполнения потоков обработки данных;
- протоколирование автоматической работы процессов обработки данных;
- обеспечение интерфейсов оператора для контроля состояния и оповещения об ошибках;
- обеспечение интерфейсов администратора системы для конфигурирования и настройки системы.

Разработка системы базировалась на следующих предпосылках и принципах:

- использование унифицированной архитектуры обработки данных, что позволяет вести автоматическое протоколирование обработки и анализ ее результатов;
- взаимодействие программных компонентов через сеть по протоколу IP для обеспечения возможности контроля и мониторинга как внутри локальной сети, так и удаленно через Интернет;
- разработка унифицированных специализированных программных компонентов (приложений, скриптов, модулей) для контроля состояния систем и процессов обработки данных;
- использование (наряду со специализированным ПО) Web-интерфейсов для визуализации информации, что позволяет обеспечить большую гибкость системы и возможность удаленного авторизованного доступа к системе контроля.

Программное обеспечение системы контроля работоспособности включает:

- программы запуска процессов обработки на рабочих станциях и серверах;

- программы-агенты сбора информации о состоянии локального компьютера и обработки и рассылки ее потребителям;
- программы-клиенты для получения информации от программ-агентов и ее архивации;
- средства доступа к информации (web-интерфейсы, cgi-скрипты и т. д.).

Принципиальная схема архитектуры системы контроля работоспособности и мониторинга обработки приведена на рис. 1. Ниже приводится более подробное описание различных элементов разработанной системы контроля.

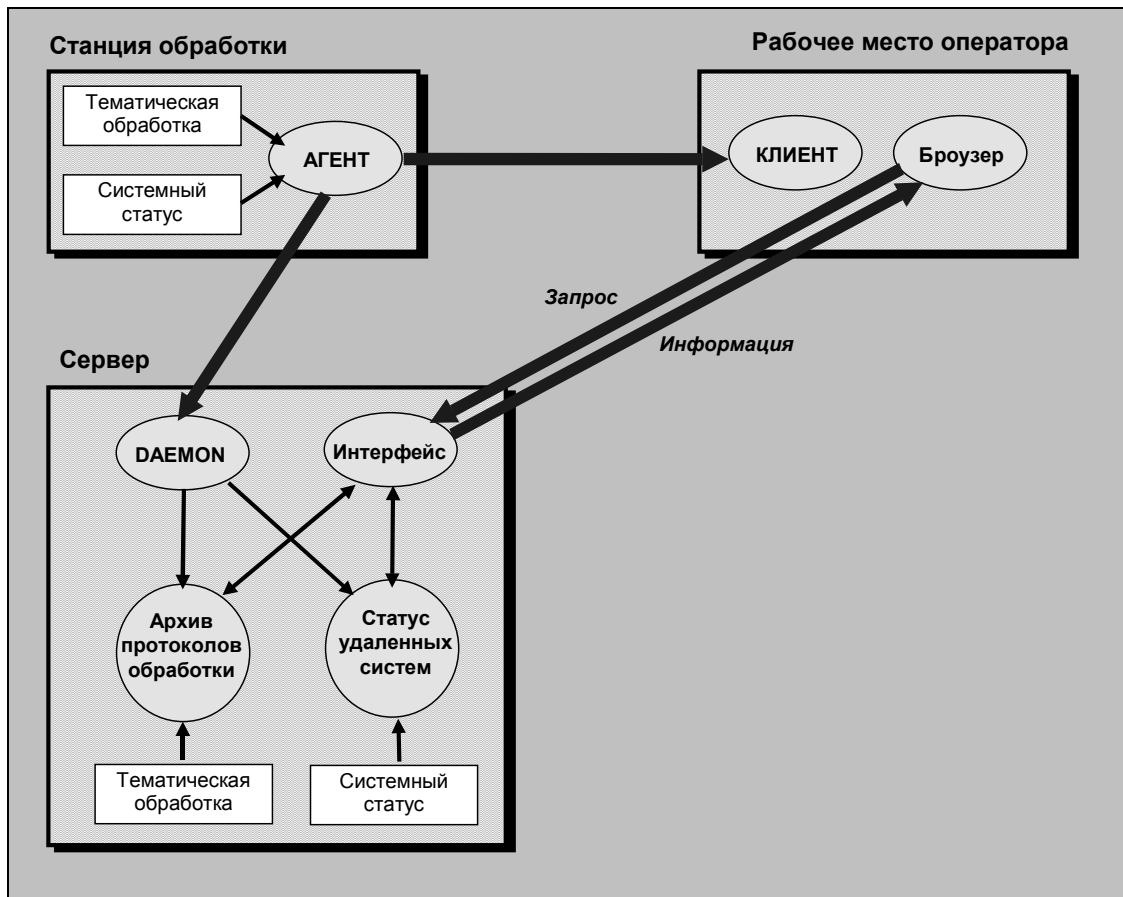


Рис. 1. Принципиальная схема архитектуры системы контроля работоспособности и мониторинга обработки

Контроль работоспособности аппаратуры (компьютеров и сетевой среды)

Программы контроля работоспособности аппаратных средств отслеживают параметры функционирования локального компьютера и передают их на центральный сервер системы, а также на рабочее место оператора, осуществляющего текущий контроль. При этом контролируются следующие параметры:

- работоспособность ОС (загрузка ЦП, доступная память, свободное дисковое пространство, состояние системных серверных процессов, время с последней перезагрузки компьютера);
- состояние серверных процессов тематической обработки;
- наличие сетевого соединения между серверами и рабочими станциями;
- состояние устройств бесперебойного питания (UPS).

При возникновении ошибочных ситуаций в web-интерфейсе соответствующие значения выделяются красным цветом, а на рабочем месте оператора, кроме визуального (цветового) оповещения, включается также и звуковой сигнал для привлечения внимания оператора.

Некоторые примеры служебных интерфейсов контроля работы аппаратуры приведены на рис. 2.



Рис. 2. Примеры служебных интерфейсов контроля работоспособности аппаратной среды обработки

Контроль и протоколирование работы процессов обработки данных

Контроль процессов обработки данных позволяет своевременно обнаружить нарушения и сбои в работе системы обработки данных, а протоколирование работы процессов дает возможность проанализировать ситуацию и найти причины неполадок. Подсистема мониторинга обработки данных выполняет следующие задачи:

1. Мониторинг выполнения процессов распределения данных на сервере:
 - а) контроль временных параметров обработки (время, прошедшее с последнего /успешного/ эффективного запуска процесса обработки);
 - б) обнаружение аварийных ситуаций (ошибки в работе программ или в данных, взаимная блокировка процессов);
 - в) задержки в поступлении данных или запуске обработки;
 - г) хранение протоколов работы процессов и обеспечение доступа к ним;
2. Мониторинг обработки данных на станциях обработки данных:
 - а) текущая активность станции обработки;
 - б) контроль зависания или остановки программ обработки данных (ограничение длительности обработки);

- в) контроль аварийных ситуаций в программах, поддерживающих систему мониторинга (остановка программы с сообщением для оператора);
- г) протоколирование запусков программ обработки данных.

Для всех временных параметров устанавливаются предельные значения, превышение которых создает ошибочную ситуацию. Возникновение такой ситуации инициирует визуальное и звуковое оповещение оператора в служебных интерфейсах контроля обработки. Примеры таких интерфейсов приведены на рис. 3.

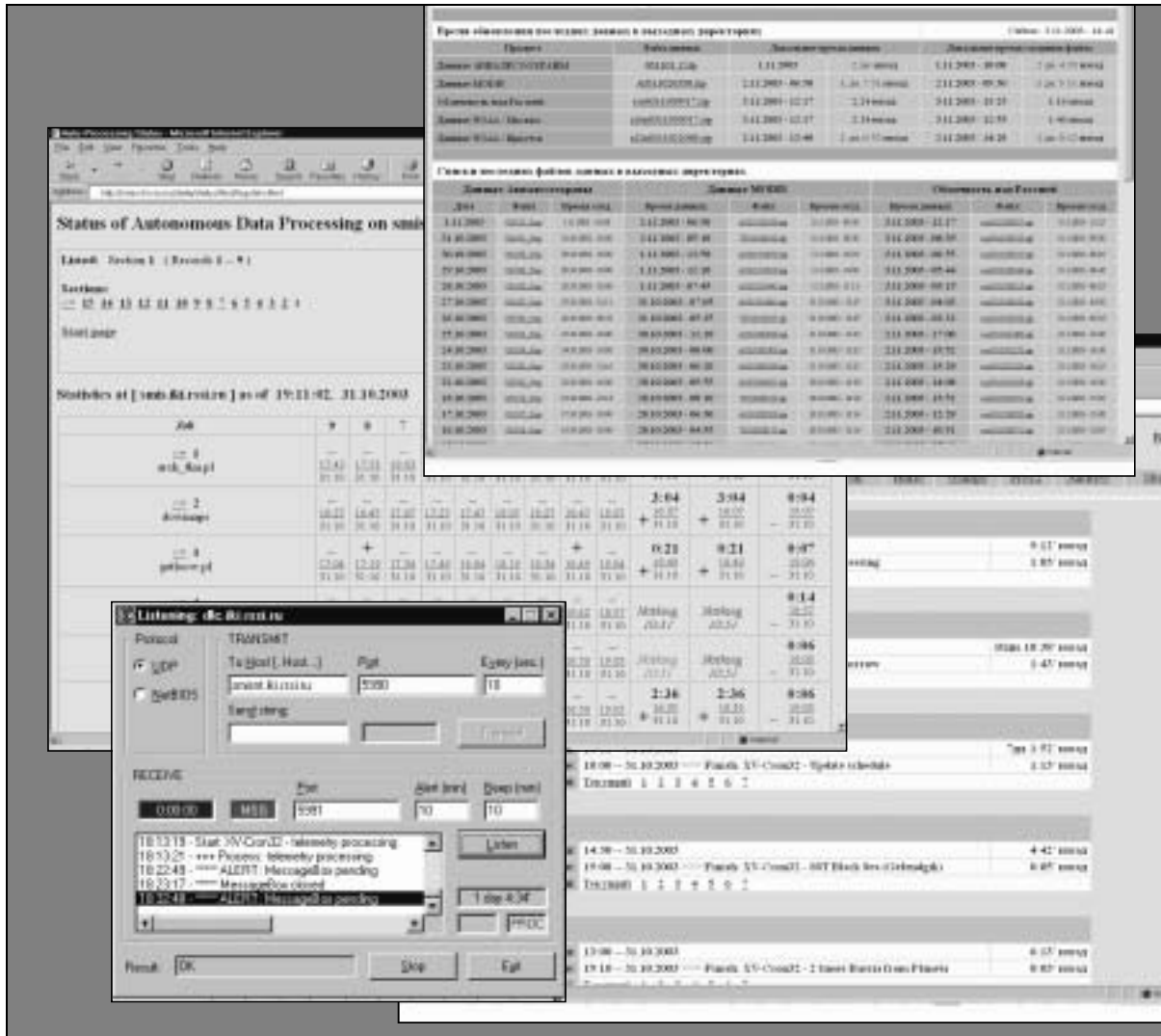


Рис. 3. Примеры служебных интерфейсов мониторинга процессов автоматической обработки данных

Программные компоненты системы контроля и мониторинга

В процессе разработки системы контроля были созданы специализированные программные продукты и интерфейсы, которые позволили решить задачи, стоящие перед системой, достаточно простыми и универсальными способами. Их можно разделить на четыре основных класса в зависимости от типа решаемых задач: контроль обработки, выполняемой на компьютерах, работающих под управлением ОС Windows, контроль работоспособности серверов данных, обеспечение функционирования центрального сервера мониторинга.

га работоспособности системы, представление данных операторам системы. Ниже приводится краткое описание этих компонент.

Программные компоненты для контроля обработки данных на компьютерах под управлением ОС Windows

Программа автозапуска процедур обработки XVCron32 позволяет формировать пакеты процедур обработки и осуществляет их запуск в соответствии с расписанием и/или с условиями старта (наличие/отсутствие указанных файлов, их обновление с момента последнего запуска). Программа поддерживает работу с сетевыми файлами, при работе на спутниковой приемной станции может синхронизировать запуски с сессиями приема данных со спутников в соответствии с расписанием приема, загружаемым автоматически.

Программа-агент RemW32 — сбор информации о ходе обработки и состоянии локального компьютера и пересылка ее на сервер. Программа пересылает данные по сети с использованием протоколов UDP или NetBIOS. Интегрируется с программой автозапуска XVCron32, что позволяет осуществлять мониторинг запусков процедур обработки и передавать эту информацию на сервер или другим клиентам.

Программные модули для контроля работоспособности серверов данных (процессы обработки и распределения данных)

Система автозапуска процессов обработки и протоколирования работы. Все процессы обработки запускаются по расписанию (посредством демона cron(8) в ОС UNIX) стандартизованным образом с использованием специализированного сценария autorun.pl, написанного на языке Perl. Сценарий позволяет осуществлять прозрачное для программ обработки протоколирование их работы и сохранение протоколов в файлы для последующего доступа к ним. Кроме того, все процедуры обработки возвращают стандартизованные коды завершения, что в дальнейшем обеспечивает автоматический анализ результатов их работы.

Модули сбора информации о состоянии сервера представляют собой CGI-сценарии на языке Perl, осуществляющие сбор информации о состоянии сервера для отображения в Web-интерфейсе мониторинга.

Специализированные процедуры мониторинга и активного оповещения об ответственных серверных процессах или аппаратных сбоях. Для контроля состояния особо ответственных процессов (таких, как серверы баз данных) написаны Perl-сценарии, осуществляющие периодический мониторинг их работоспособности. При сбоях в работе такого процесса осуществляется протоколирование ошибочной ситуации и дополнительно производится рассылка сообщения по электронной почте лицам, отвечающим за функционирование процесса, для максимально оперативной коррекции ошибки. Подобный подход применяется также и для мониторинга некоторых нежелательных состояний аппаратуры (в частности, переполнение файловых систем на сервере).

Специализированные Web-интерфейсы и CGI-сценарии для мониторинга ответственных процессов обработки разрабатываются индивидуально в тех случаях, когда необходимо максимально оперативно реагировать на задержки или сбои в обработке данных, либо если выходные данные имеют сложную структуру, и стандартные интерфейсы контроля не обеспечивают должной информативности.

Центральный сервер мониторинга работоспособности системы

Как правило, он совпадает с одним из серверов данных, что обеспечивает простоту интеграции всей системы, при необходимости, может быть отдельным компьютером, при условии организации пересылки данных мониторинга с серверов данных. Для организации его работы созданы:

Серверный демон (процесс) сбора информации мониторинга rwd.pl. Perl-скрипт осуществляет прием информации по протоколу UDP от программ-агентов RemW32, работающих на станциях обработки данных, и сохранение протоколов в файлы.

Набор служебных web-интерфейсов и CGI-сценариев для удаленного доступа к информации мониторинга, статистике обработки данных, протоколам обработки, а также для администрирования системы.

Рабочее место оператора

На рабочем месте оператора установлены и работают постоянно (как сервис Windows) программы **RemW32** в режиме получения информации. Они позволяют осуществлять контроль хода обработки на станциях обработки данных в оперативном режиме с использованием активного оповещения об ошибках обработки (визуальное и звуковое оповещение). Кроме того, оператор имеет возможность использовать стандартный браузер для более детального анализа ситуации, мониторинга обработки и контроля работы аппаратных средств через служебные web-интерфейсы, которые предоставляют более подробную и обширную информацию.

Развитие системы контроля и мониторинга обработки

Накопленный опыт использования системы контроля выявил некоторые ограничения и узкие места текущей реализации. На основе их анализа была начата разработка нового поколения системы, в настоящее время близится к завершению реализация ее пилотной версии. Архитектура системы контроля нового поколения является развитием первоначальной концепции, с существенным расширением некоторых ее возможностей (рис. 4). Главными особенностями новой версии системы являются:

Использование базы данных для хранения всей текущей и исторической информации, включая статистику обработки, протоколы работы программ и информацию о сбоях и ошибках обработки. Основная цель этого — снять имеющиеся ограничения на объем хранимой в системе информации при сохранении эффективности доступа к ней. Это позволяет:

- хранить информацию в течение большого периода времени для эффективного анализа сбоев в работе системы;
- обеспечить быструю выборку данных мониторинга по различным запросам и условиям;
- организовать хранение истории ошибочных состояний, а также предпринятых корректирующих действий для последующего анализа и отчетности;
- избежать проблем синхронизации при сохранении информации в случаях одновременного запуска многих процессов.

Интеграция средств настройки систем автозапуска и контроля в единую систему. Новая интегрированная система конфигурации позволяет избежать трудоемкого процесса редактирования таблиц автозапуска (crontab), а также снимает проблему их синхронизации с конфигурацией системы контроля.

Расширение набора ошибочных состояний обработки данных, распознаваемых автоматически. Это позволит использовать активное оповещение оператора о сбоях в обработке для большинства сбойных ситуаций, что увеличит оперативность их коррекции.

Поддержка назначения групп процессов для удобства анализа обработки в рамках различных проектов и задач.

Унификация процедур обработки ошибочных состояний аппаратуры и процессов обработки данных в рамках единого подхода.

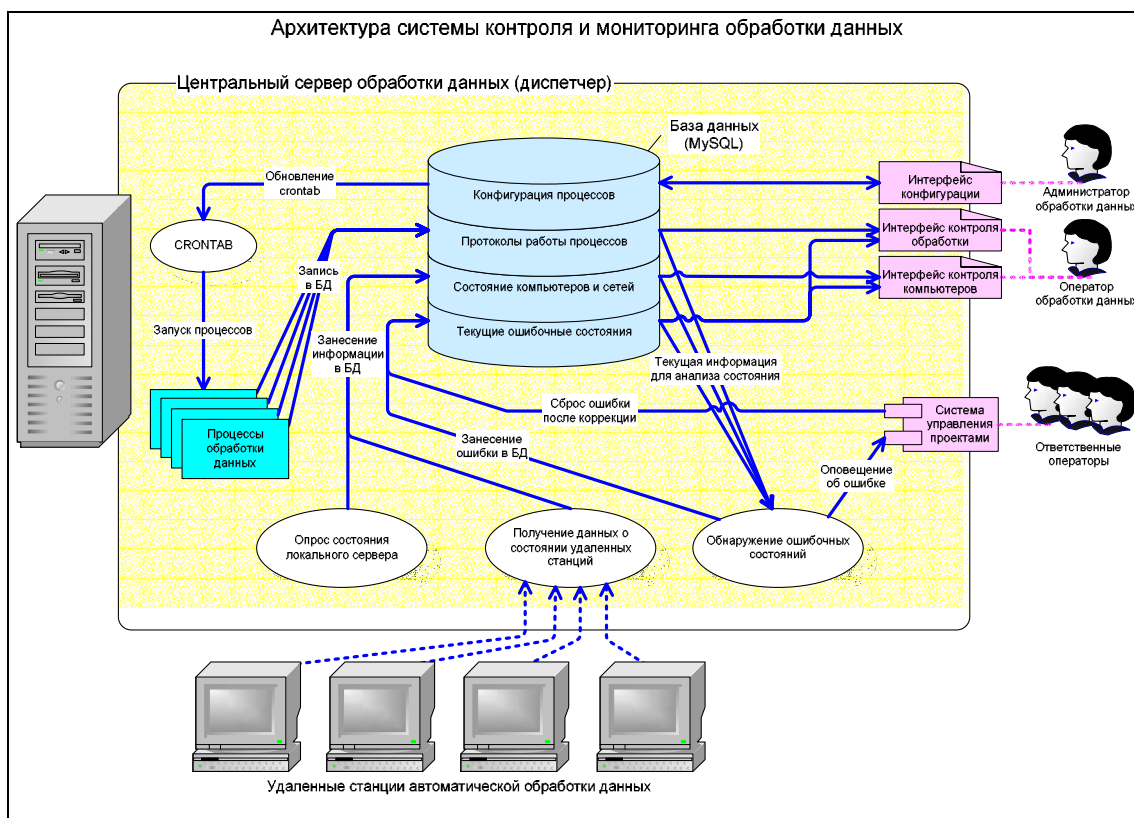


Рис. 4. Архитектура системы контроля и мониторинга обработки 2-го поколения.

Модификация системы операторского контроля, построенная на базе системы управления проектами, позволит решить организационные задачи контроля обработки данных:

- назначение ответственных за различные группы процессов и разграничение ответственности по проектам;
- контроль за действиями операторов по коррекции сбоев в обработке;
- использование активного оповещения ответственных операторов (по электронной почте, SMS и т. д.) о сбоях в обработке.

Широкое использование активного оповещения оператора об ошибках путем внедрения стандартизированной системы оповещения об ошибках.

Расширение возможностей и гибкости операторских интерфейсов.

Увеличение набора контролируемых параметров работоспособности аппаратной среды.

Описанная система контроля была создана, опробована и в настоящее время используется во многих системах обработки данных, разработанных и поддерживаемых ИКИ РАН. Следует отметить, что достаточная универсальность созданной технологии и базовых программных элементов позволяет использовать их не только для работы с системами обработки спутниковых данных, но и для контроля и управления блоками различных систем мониторинга. В настоящее время созданные на этой основе системы контроля обработки установлены и используются более чем в 30 кластерах обработки данных по всей территории России, таких как: Информационная система мониторинга лесных пожаров МПР РФ (Москва, Пушкино, Иркутск, Хабаровск), Центры приема и обработки спутниковых данных Госкомгидромета РФ (Москва, Хабаровск), Отраслевая система мониторинга Госкомрыболовства РФ (около 20 центров от Мурманска до Владивостока) и др. Опыт эксплуата-

ции этих систем позволяет сделать вывод о правильности выбранных подходов и эффективности созданных программных компонент.

Литература

1. Андреев М.В., Галеев А.А., Ефремов В.Ю., Ильин В.О., Крашенинникова Ю.С., Лурия Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В. Построение автоматизированных систем сбора, хранения, обработки и представления спутниковых данных для решения задач мониторинга окружающей среды. // Всерос. конф. «Дистанц. зондирование поверхности Земли и атмосферы». Иркутск, 2–6 июня 2003 С. 4.

2. Абушенко Н.А., Барталев С.А., Беляев А.И., Еришов В.В., Коровин Г.Н., Кошелев В.В., Лурия 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.

3. Асмус В.В., Бурцева Т.Н., Лурия Е.А., Мазуров А.А., Назиров Р.Р., Милехин О.Е., Прошин А.А., Флитман Е.В. Система «Спутник» для оперативного доступа удаленных пользователей к спутниковым данным // Тез. докл. 3-й Международ. науч.-технич. конф. «Космонавтика. Радиоэлектроника. Геоинформатика». Рязань, Россия, 6–8 сентября 2000. С. 307.

4. Андреев М.В., Ефремов В.Ю., Гостев М.В., Дмитриев Г.А., Крашенинникова Ю.С., Лурия 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.

5. Андреев М.В., Дегай А.Ю., Крашенинникова Ю.С., Лурия Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В., Гербек Э.Э., Проценко И.Г. Возможности организации отраслевого спутникового мониторинга // Рыбное хоз-во. Спец. вып. 2001. С. 35.

6. Захаров М.Ю., Лурия Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В. Система автоматического приема и архивирования спутниковых данных. М.: ИКИ РАН. Препринт Пр-1988. 1998. 19 с.

7. Loupian E., Mazurov A., Nazirov R., Proshin A., Flitman E. A Universal Technology for Development of Satellite Data Storage Systems // Proc. of the 4th Intern. Symp. on “Reducing the Cost of Spacecraft Ground Systems and Operations”, 2001.