

ОБЩЕСИСТЕМНОЕ ОБЕСПЕЧЕНИЕ ОТРАСЛЕВОЙ СИСТЕМЫ МОНИТОРИНГА ГОСКОМРЫБОЛОВСТВА

*А.А. Нестеренко**, А.А. Романов**, М.В. Андреев**, Е.А. Лупян**

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

*** Федеральное государственное унитарное предприятие «Всероссийский научно-исследовательский и проектно-конструкторский институт экономики, информации и автоматизированных систем управления рыбного хозяйства» (ФГУП ВНИЭРХ). E-mail: nester@vnierrkh.ru*

Настоящая работа посвящена описанию принципиальных схем организации информационного обмена в отраслевой системе мониторинга (ОСМ) Госкомрыболовства [1–4]. В ней рассматриваются вопросы, связанные с архитектурой построения системы, которая позволяет организовать сбор различной информации, в том числе спутниковых данных, поступающей из различных источников в унифицированную базу данных (БД ОСМ), и организацией доступа пользователей к данным ОСМ.

Следует отметить, что одна из основных особенностей информационных ресурсов ОСМ,— то, что они являются распределенными. При этом выбранная сегодня общая архитектура построения системы опирается на системы информационных узлов, в которые постоянно поступает информация, первоначально собираемая из различных источников. Собираемая в системе информация агрегируется в центральных базах данных, а после этого рассылается в так называемые информационные узлы (ИУ ОСМ). При этом такие узлы могут располагаться и в организациях, которые являются поставщиками информации в ОСМ, т. е. в информационной системе возникает ситуация, когда отдельные ее узлы одновременно являются как получателями, так и поставщиками информации. Поэтому, для того чтобы, в частности, избежать противоречивости в информационных ресурсах в таких ситуациях, необходимо было разработать специальную архитектуру и принципы построения информационных потоков и взаимодействия информационных ресурсов ОСМ.

Для начала рассмотрим основные принципы сбора и рассылки данных, реализованные в ОСМ. Логическая схема сбора и распространения информации, поступающей в отраслевую систему мониторинга из различных источников, приведена на рис. 1. На ней представлено несколько информационных уровней, ответственных за сбор, организацию хранения и распространения данных. Разделение системы на такие уровни позволяет упростить логику работы системы и сделать работу ее достаточно прозрачной и непротиворечивой.

Уровень «Источники» представляет собой первичные с точки зрения системы источники информации. Эта информация передается на следующий уровень в виде, требующем дополнительной обработки для того, чтобы ее можно было ввести в БД ОСМ, т. е. ее необходимо перевести во входной формат ОСМ и согласовать с имеющимися в БД справочниками. Структура и формат этих данных могут быть различными. Соответственно по-разному может быть организован процесс перевода информации в вид, необходимый для дальнейшей обработки. Но выходным форматом этого процесса всегда является унифицированный формат БД ОСМ.

С точки зрения обеспечения процесса сбора информации и включения ее в единый отраслевой информационный ресурс (ЕОИР) ключевым уровнем, безусловно, является уровень с условным названием «Сборка». На этом уровне должны находиться информационные узлы, являющиеся источником части данных ЕОИР. Такие информационные узлы мы будем называть активными информационными узлами (АИУ). Для того чтобы упростить работу всей системы и сделать ее не противоречивой, необходимо предъявить к таким узлам следующее основное требование: каждый АИУ должен обеспечивать работу с фиксированным набором источников информации, и информация от любого источника

должна попадать в систему только через один АИУ. При таком подходе, безусловно, остается ряд проблем, связанных с включением в систему данных, требующих, например, предварительного наличия в системе определенных справочников (например, для включения в систему информации о позиции конкретных судов в системе уже должен существовать справочник, содержащий информацию об этих судах). Однако такой подход позволяет избежать проблем, связанных с синхронизацией однотипной информации, приходящей из разных узлов.

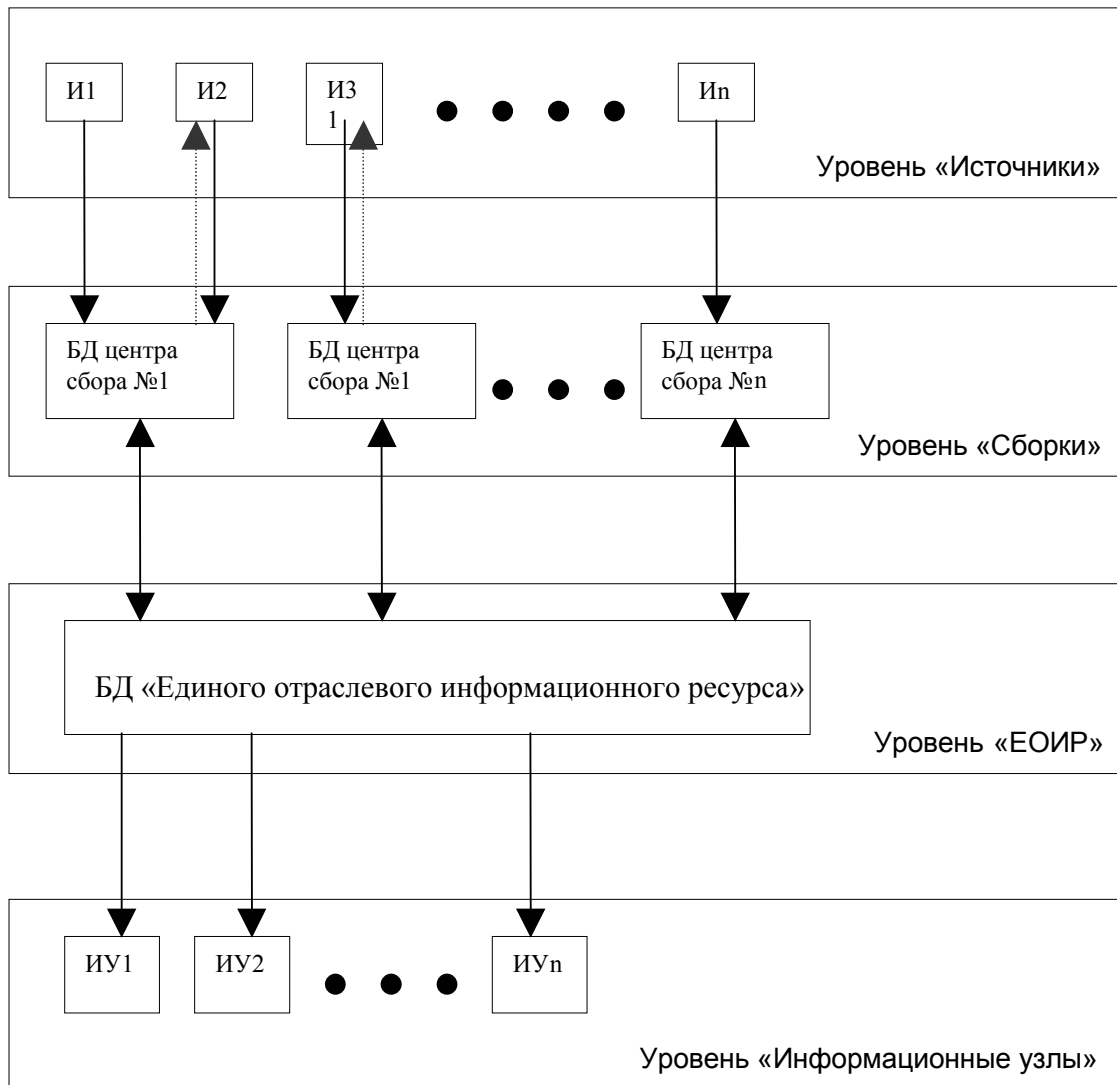


Рис. 1. Общая логическая схема сбора и распространения информации, поступающей в отраслевую систему мониторинга из различных источников

Далее данные передаются на уровень «Единого отраслевого информационного ресурса». На этом уровне в БД формата ОСМ собирается информация из всех источников системы и организовывается рассылка данных конечным пользователям на уровень «Информационные узлы». Каждый информационный узел имеет БД в формате ОСМ, в которой содержится часть данных БД ЕОИР, необходимая пользователям ИУ для работы. Таким образом, содержание БД ИУ поддерживается в актуальном состоянии.

Как уже отмечалось выше, в ОСМ параллельно фактически осуществляются два процесса: сбор данных в ЕОИР и рассылка информации получателям. При этом данные каж-

дой копии БД ОСМ являются подмножеством данных ЕОИР, и каждое подмножество данных поступает в ЕОИР только из одного источника. Рассмотрим предлагаемую технологическую схему реализации системы сбора и рассылки данных в ОСМ (рис. 2). На ней можно выделить следующие функциональные блоки:

- *центральный узел (ЦУ)*, аккумулирующий в своей БД все данные ЕОИР и осуществляющий рассылку необходимой информации всем потребителям данных;
- *информационные узлы (ИУ)*, являющиеся только получателями данных, и не поставляющие информацию в центральный узел;
- *активные информационные узлы (АИУ)*, являющиеся как потребителями, так и получателями данных;
- *узлы рассылки (УР)*, не обрабатывающие данные, но использующиеся для оптимизации сетевого трафика и уменьшения времени реакции системы в качестве дополнительного этапа пересылки данных.

Кроме того, информационный узел может реализовывать функциональность узла рассылки. В этом случае может быть проведена оптимизация потоков данных. Наиболее сложным случаем является активный узел, с несколькими получателями, часть из которых также является активными информационными узлами.

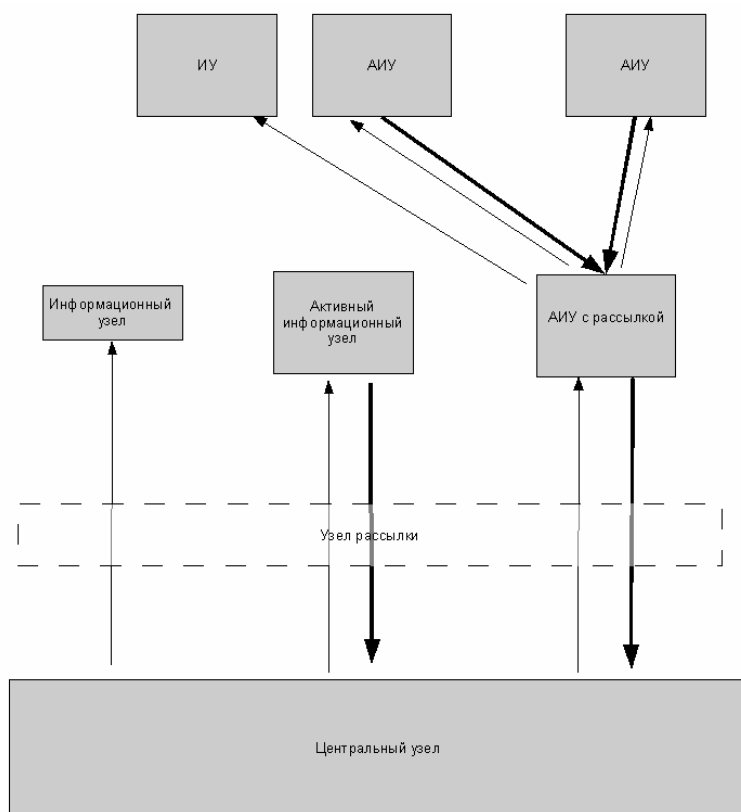


Рис 2. Технологическая схема сбора и распространения информации, поступающей в отраслевую систему мониторинга из различных источников

Остановимся более подробно на обсуждении схемы потоков информации в каждом из этих блоков.

Самый простой случай — это обычный информационный узел (рис. 3). ИУ получает данные из центрального узла или из узла рассылки и вводит эти данные в свою БД. При этом никаких дополнительных проверок не происходит, и в БД вводятся все данные из потока, которые удовлетворяют правилам синтаксиса формата потока и условиям целостности БД.

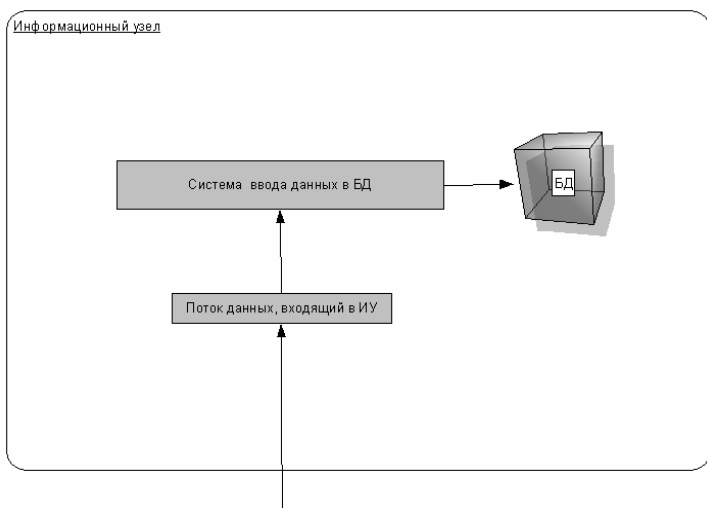


Рис 3. Схема неактивного информационного узла

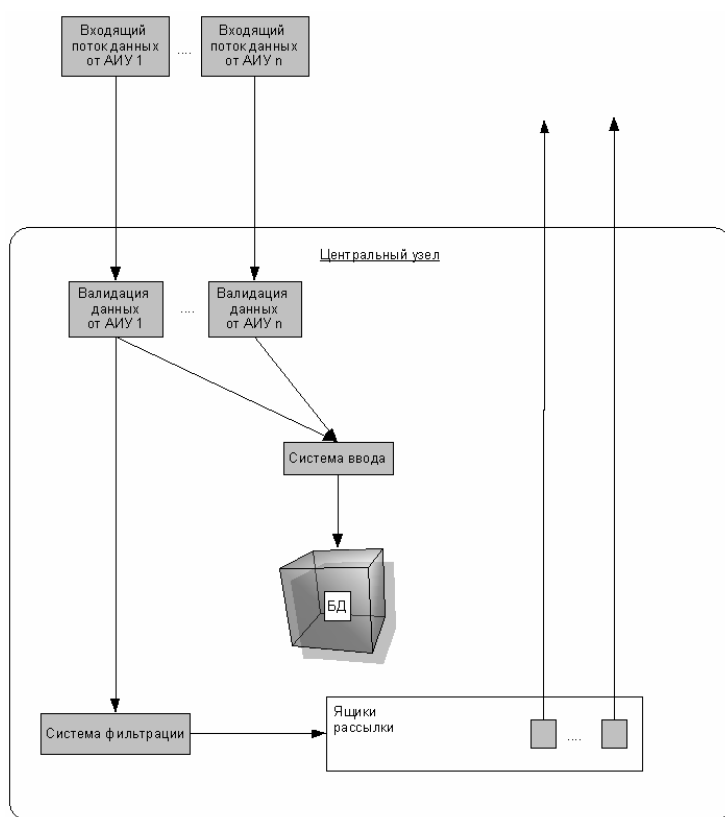


Рис 4. Схема центрального информационного узла

Центральный узел (рис. 4) получает данные от всех активных узлов. При этом данные могут проходить через один или несколько узлов рассылки, но обязательным требованием является то, что вся информация из всех входных источников системы обязательно приходит в БД центрального узла. Таким образом, БД ЦУ содержит полный объем данных ЕОИР, а любая другая БД ОСМ содержит то или иное подмножество этой информации. Перед тем как вводить данные, полученные от АИУ, в свою БД, ЦУ производит валидацию этих данных, целью которой является дополнительная проверка того, что данные, пришедшие от АИУ, лежат в пределах его компетенции. Далее собранная и проверенная информация вводится в БД ЦУ, а полученный поток данных проходит через систему фильтрации, формирующую поток данных для каждого ИУ. Таким образом, каждый пользователь данных системы получает необходимую информацию.

Активный ИУ (рис. 5), кроме информации из ЦУ, получает данные из первичных источников информации и реализует ИСР конвертор для каждого из них. Полученные из ИСР конвертора данные поступают в систему верификации и ввода в БД АИУ. Здесь, кроме проверок синтаксиса формата и правильности порядка поступления данных в сформированном потоке, происходит проверка данных на согласованность с БД. Данные, прошедшие верификацию, вводятся в БД и записываются в файл фильтра. Такие файлы фильтра составляют поток данных, поступающий из АИУ в ЦУ.

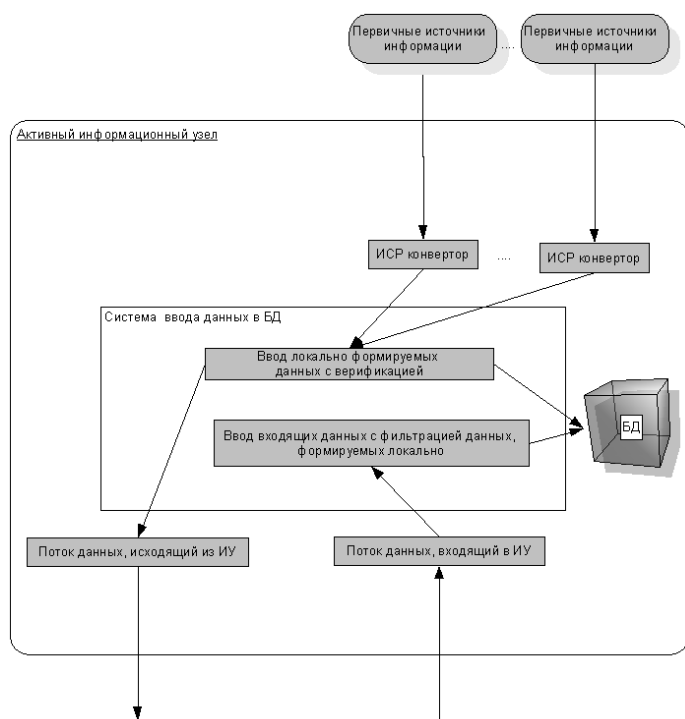


Рис 5. Схема активного информационного узла

Следует обратить внимание на важность этапа верификации данных в АИУ, и, как следствие, необходимость наличия в активном информационном узле самосогласованной БД ОСМ, содержащей всю необходимую для этого информацию. Это может привести, например, к тому, что во входящий в АИУ поток будут включаться данные (справочники, отчетная информация и т. п.), не нужные непосредственно для работы АИУ. Может показаться, что хорошо реализованный ИСР конвертор не нуждается в дополнительных проверках. Но процесс верификации данных является необходимой частью формирования потока пополнения, поскольку он происходит одновременно с вводом данных и только так можно гарантировать, что поток данных, формируемый АИУ, будет без ошибок усваиваться системой ввода как в ЦУ, так и в информационных узлах. В случаях же возникновения проблем только АИУ, производящий поток данных, может их решить. Возможно также реализовать систему контроля на другом уровне, скажем ЦУ. Но в этом случае время реакции на ошибки сильно возрастает, особенно если принять во внимание сильно распределенный характер ОСМ.

Поскольку АИУ является для системы в целом источником определенного рода информации, то его БД все время содержит актуальные данные, относящиеся к его компетенции. Следовательно, такого рода данные, приходящие к нему во входящем потоке данных из ЦУ или узла рассылки, следует игнорировать. Имеет смысл реализовать этот контроль в таком виде, чтобы проверки на верификацию локально формируемых данных и фильтр входящего потока осуществлялись однотипно и конфигурировались одними и теми же настройками системы ввода данных.

Наиболее сложным функциональным блоком системы является активный узел, с несколькими получателями, часть из которых также является активными информационными узлами (рис. 6). Такой узел взаимодействует с другими блоками системы следующим образом: он получает информацию из ЦУ; отправляет необходимую информацию узлам, получающим данные ЕОИР через него (такие ИУ мы дальше будем называть зависящими от рассматриваемого АИУ); принимает от некоторых из этих узлов формируемые ими данные и отправляет в ЦУ данные, полученные им от зависящих от него узлов, а также данные, формируемые им самим. Такой узел совмещает в себе функциональность АИУ и узла рассылки. При этом представляется целесообразным провести некоторую оптимизацию потоков данных.

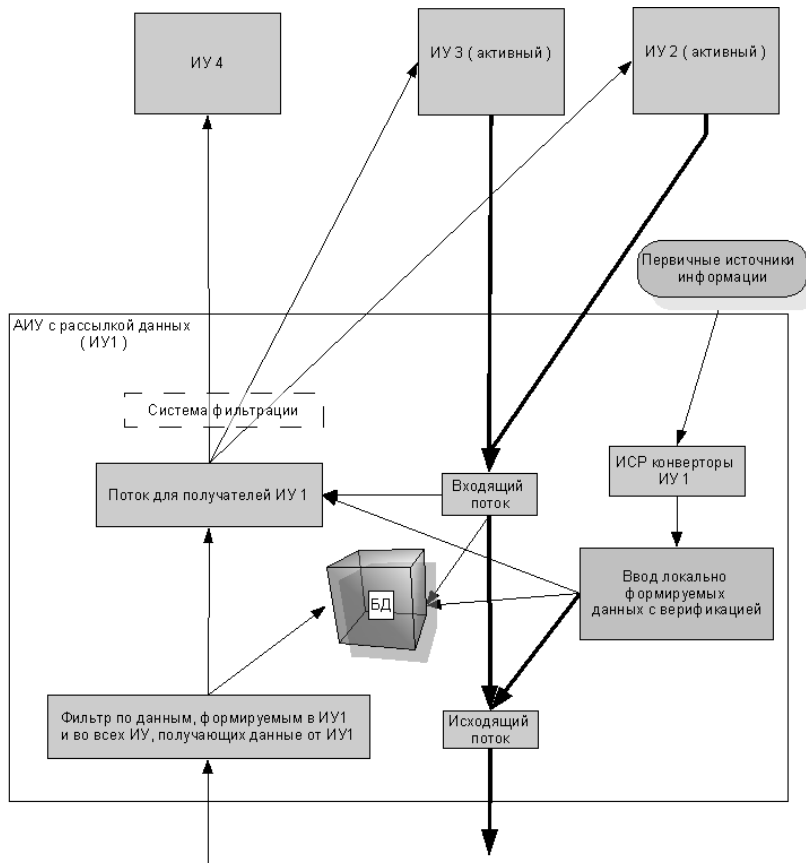


Рис 6. Схема активного информационного узла с рассылкой

Так же, как и обычный, АИУ с рассылкой данных реализует некоторое количество ИСР конверторов и валидацию получаемых от них данных одновременно с вводом в локальную БД. Полученный таким образом фильтрат попадает в исходящий из узла поток данных для ЦУ и в поток данных, предназначенный для зависящих ИУ. Этот поток проходит через систему фильтрации, и каждый ИУ получает необходимый ему объем данных.

Далее узел получает данные из зависящих узлов. Эти данные вводятся в локальную БД, попадают в поток данных для зависящих ИУ и передаются ЦУ. Таким образом, данные от одного из зависящих узлов попадают в локальную БД узла и всем узлам-получателям без необходимости перекачки этих данных в ЦУ и обратно. При этом возможно при помощи системы фильтрации избежать перекачки данных, полученных от одного из зависящих узлов обратно ему же. Но такой оптимизации можно и не делать, если это упростит настройки системы, и дополнительный входящий трафик в зависящий узел является приемлемым. Это не приведет ни к каким последствиям, так как АИУ, формирующий данные, не будет усваивать эти данные из входного потока.

Данные, полученные рассматриваемым узлом от ЦУ, проходят через фильтр, удаляющий оттуда информацию, находящуюся в компетенции узла и зависящих узлов. Это необходимо, так как и в локальной БД и в БД зависящих ИУ такие данные вследствие описанной выше оптимизации потоков поступают всегда раньше, чем они приходят из ЦУ, и таким образом информация, уже находящаяся в указанных БД, всегда более актуальна. После этого данные усваиваются в локальной БД и попадают в поток данных для зависящих узлов, после чего они проходят через систему фильтрации и передаются в зависящие узлы.

Учитывая, что может потребоваться верификация любой БД системы с БД ЦУ, представляется необходимым, чтобы вся информация, проходящая через рассматриваемый узел к зависящим узлам, усваивалась в локальной БД. Даже если эти данные не нужны в рассматриваемом АИУ для работы, а требуются только для одного из зависящих узлов, их нужно ввести в локальную БД. В таком случае, как правило, пользователи рассматриваемого узла вообще не работают с таблицами БД, для которых предназначаются эти данные. Следовательно, дополнительные расходы такого подхода связаны только с дисковым пространством сервера БД и не повлияют на быстродействие при работе пользователей узла с БД. Взамен мы получаем существенное упрощение схемы в целом. Аналогично при пересылке данных от зависящих узлов в ЦУ эти данные необходимо ввести в локальную БД. Это не значит, однако, что такие данные обязательно нужно передать во все зависящие узлы.

Таким образом, мы видим, что взаимодействие функциональных блоков системы организуется в древовидную структуру. Подытожим основные принципы построения системы сбора и рассылки информации:

- Вся информация ЕОИР делится на подмножества. Любое подмножество информации формируется в одном и только одном АИУ.
- В БД ЦУ находится все множество информации ЕОИР.
- В каждый ИУ приходит только один поток информации от ЦУ или узла рассылки, и данные, исходящие из этого узла, передаются тому же узлу, от которого поступает входной поток.
- Все данные, проходящие через АИУ с рассылкой как в направлении от ЦУ к зависящим узлам, так и обратно, вводятся в локальную БД АИУ, т. е. данные, содержащиеся в БД узлов при прохождении дерева в направлении от ЦУ, представляют собой вложенные подмножества данных ЕОИР, БД каждого зависящего узла — это подмножество данных АИУ, через которое данный узел обменивается информацией с ЦУ.

Перечисленные выше требования позволят построить систему, в которой не будет возникать противоречивое дублирование информации, и информационный ресурс любого информационного узла будет асимптотически являться подмножеством информационных ресурсов центрального узла. Следует обратить внимание, что в конкретные моменты времени в системе, конечно, могут возникать расхождения в информации в узлах и центральном узле, но связано это будет только с конечным временем обмена информацией. К сожалению, в любой распределенной информационной системе избежать такой ситуации можно, фактически пожертвовав скоростью реакции системы на запросы пользователей, что может привести к большим неудобствам и практической невозможности использования такой системы. В то же время на практике, в частности, в ОСМ, во многих случаях не требуется постоянной целостности всей системы, поскольку реально быстрая актуализация отдельных ее ресурсов требуется лишь ограниченной группе пользователей, фактически ответственных за эти ресурсы. Для этой группы пользователей актуализация в предлагаемой схеме происходит фактически мгновенно. Для остальных же пользователей системы мгновенная актуализация этих ресурсов не является критичной. Важно, что в предлагаемой схеме даже та часть системы, которая не успела в конкретный момент времени пройти

полной актуализации, сохраняет свою локальную целостность, что позволяет локальным пользователям осуществлять работу с ней. Отметим также, что, по мере улучшения качества и быстродействия используемых в системе каналов обмена данными, скорость полной актуализации в предлагаемой системе будет постоянно расти, и временные расхождения в информации в различных узлах системы становятся все менее и менее существенными.

Следует отметить, что одним из типов данных, предоставляемых пользователям ОСМ, являются продукты обработки спутниковых изображений. Эти продукты содержат информацию о состоянии окружающей среды, такую как ледовая обстановка, облачность и т.п. Возможности использования спутниковых данных для решения задач ОСМ описаны, в частности, в [5, 6]. Для работы с этой информацией в ОСМ были включены центры приема и обработки спутниковых данных, установленные в КЦСМ (г. Петропавловск-Камчатский) и в ИКИ РАН (г. Москва). В этих центрах были созданы автоматизированные системы обработки, хранения и представления данных. Для организации доступа пользователей к данным в ОСМ созданы как стандартные интерфейсы работы с архивами спутниковых данных (см., например, [7]), так и специализированные интерфейсы, позволяющие проводить совместный анализ спутниковых данных и данных о позиционировании судов [8]. Пример такого интерфейса представлен на рис 7.

В заключение отметим, что работы по реализации отдельных элементов предложенной архитектуры были поддержаны проектом РФФИ № 03-07-90358.

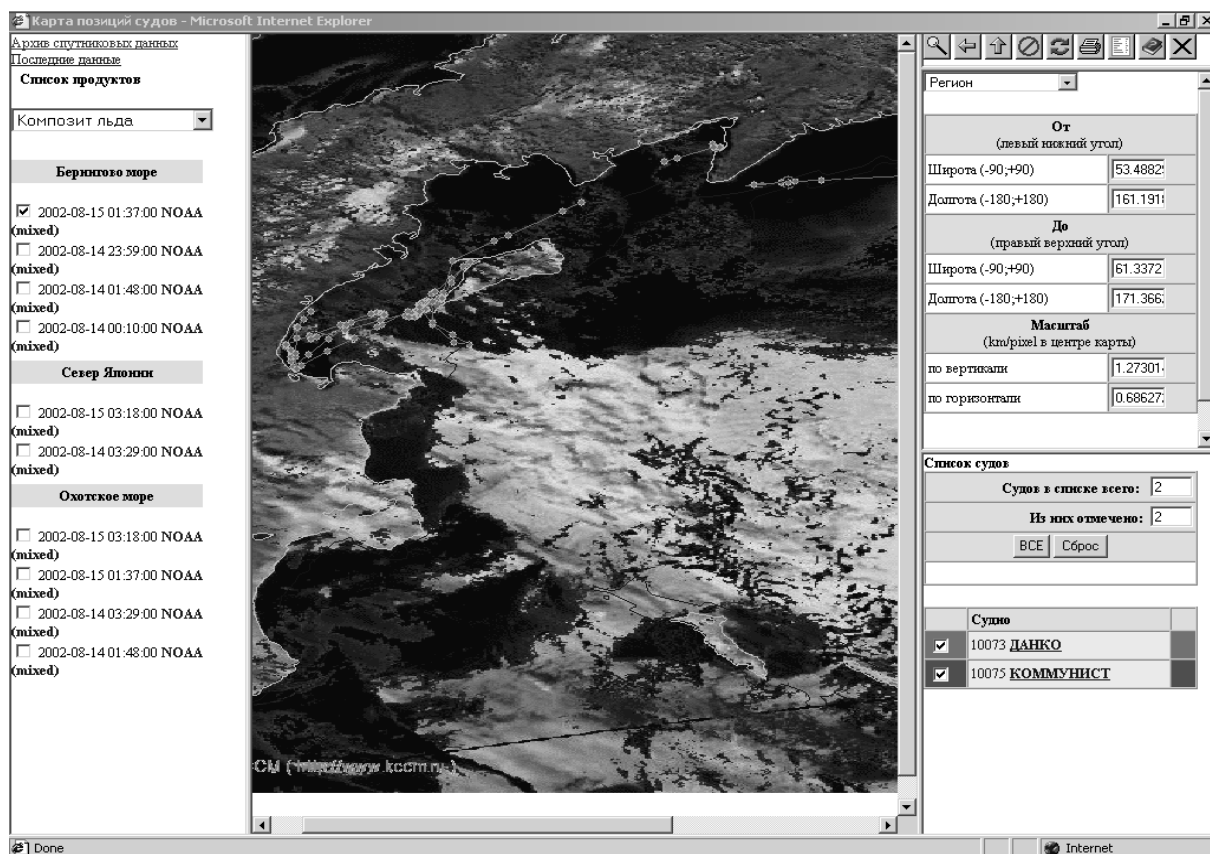


Рис. 7. Пример интерфейса, обеспечивающего одновременную работу с данными позиционирования судов и спутниковыми данными

Литература

1. Романов А.А. Современные возможности использования спутниковой информации в отраслевых задачах краткосрочного прогнозирования и управления флотом. Состояние и ближайшие перспективы // Новые информац. технологии и дистанц. методы отраслевого мониторинга промысловых районов Мирового океана. М.: ВНИРО, 1996. С. 19–35.
2. Романов А.А., Родин А.В., Мишкин В.М. Концепция отраслевой службы спутникового научно-производственного мониторинга промысловых районов Мирового океана // Дистанц. методы мониторинга промысловых районов Мирового океана в задачах информац. поддержки отраслевой научно-производств. деятельности. М.: ВНИРО, 1997. С. 7–32.
3. Проценко И.Г. Информационная система мониторинга рыболовства // Рыбное хозяйство. Спец. вып. 2001. С. 3.
4. Романов А.А., Матвеев С.В., Лупян Е.А., Мазуров А.А., Прошин А.А., Флитман Е.В. Перспективы создания и развития распределенного архива данных спутникового мониторинга океанографической обстановки Госкомрыболовства // Тез. докл. конф. «Информац. ресурсы об океане — актуальные проблемы формирования, распространения и использования в научных исслед. и в морской деятельности». ОИР. 2002. Обнинск, ВНИИГМИ-МЦД. С. 107.
5. Андреев М.В., Дегай А.Ю., Крашенинникова Ю.С., Лупян 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 хозяйство. Спец. вып. 2001. с. 35.
6. Романов А.А. Основы обработки и анализа данных космического дистанционного зондирования океана. М.: МФТИ, 2003. 272 с.
7. Лупян Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В. Универсальная технология построения систем хранения спутниковых данных. М.: ИКИ РАН. Препринт Пр-2024. 2000. 22 с.
8. Андреев М.В., Дегай А.Ю., Лупян Е.А., Мазуров А.А., Прошин А.А., Флитман Е.В. Картографический web-интерфейс на основе технологий DHTML и JavaScript. М.: ИКИ РАН. Препринт Пр-2051. 2002. 23 с.