

УПРАВЛЕНИЕ ДАННЫМИ В ОПЕРАТИВНЫХ ТЕХНОЛОГИЯХ ГИДРОМЕТЦЕНТРА РОССИИ

Все виды прогностической деятельности в области гидрометеорологии основываются на использовании данных, оперативно получаемых от различных наблюдательных систем, а также архивных данных. При составлении краткосрочных прогнозов в основном используются текущие данные о состоянии атмосферы, а при разработке долгосрочных и среднесрочных прогнозов используется также информация о прошлых состояниях атмосферы и гидросферы — за периоды времени от нескольких месяцев до ста лет.

Большой объем обрабатываемой информации, разнообразие наблюдательных платформ, широкий круг решаемых задач как оперативного, так и исследовательского характера требуют организации и совершенствования общей системы управления данными начиная от приема входной информации и заканчивая доведением результатов численных прогнозов до рабочих мест прогнозистов — синоптиков, океанологов, гидрологов и других специалистов — через локальные и глобальные сети.

Система управления данными главного вычислителя является важнейшей составляющей общей системы информационного обеспечения прогностической деятельности Гидрометцентра России. Она состоит из следующих разделов: технологическая линия обеспечения вычислителя данными наблюдений и продукцией зарубежных центров, объединенный банк прогностических, прикладных и исследовательских задач, автоматизированная система оперативной обработки информации, обеспечивающая регламентированный процесс совместной обработки данных, распределенная база данных коллективного доступа, обеспечивающая накопление, хранение и обмен информацией, технологические средства отображения и распространения в международных форматах произведенной продукции, информационное обеспечение исследовательских задач.

Все перечисленные технологии являются продуктом научной, оперативной и организационной деятельности нескольких поколений ученых и программистов Гидрометцентра России.

Первые системы расчета полей барической топографии на ЭВМ были разработаны и внедрены С. Л. Белоусовым почти 40 лет назад. В дальнейшем, к концу 1960-х годов, эти работы получили развитие на ЭВМ М-20 (М-220, М-222) и ВЕСНА. Тогда впервые был внедрен полный цикл оперативной обработки данных, начинавшийся с приема исходной информации, ее первичной обработки и заканчивавшийся расчерчиванием карт прогностических метеорологических полей на графопостроителе, которые затем передавались в оперативные подразделения. Реализация цикла проводилась вручную операторами ЭВМ и требовала их высокой квалификации вследствие низкой надежности вычислительной техники.

Установка в 1974 г. ЭВМ БЭСМ-6, одной из лучших в мире в то время, обладавшей высоким быстродействием, большим объемом магнитных барабанов и, главное, существенно более высокой надежностью, позволила начать работы по созданию автоматизированных систем обработки оперативной информации. Эти работы велись под руководством К. А. Семендяева и О. М. Кастина.

Кроме того, в это время произошел очень значимый для научного сообщества переход на язык программирования Фортран, что привело к резкому увеличению количества задач, выполнявшихся на ЭВМ.

В этот же период под руководством Б. Г. Буравцева делались первые шаги по технической и технологической организации распределенной обработки данных: прием и раскодирование данных наблюдений производились на ЭВМ МИНСК-32, раскодированные данные передавались по специальному каналу на БЭСМ-6, где выполнялись прогностические задачи, их результаты на магнитных лентах переносились на М-222 для расчерчивания, а также выдавались на перфоленду в одной из первых версий кода GRID для распространения в глобальной сети.

Появление в 1978—1980 гг. ЭВМ серии ЕС и Cyber-172 позволило поднять на новый, более высокий уровень разработку оперативных систем управления данными. В этот период была развита концепция банка данных ПРОГНОЗ, в рамках которой создавалось и внедрялось программное обеспечение автоматизированных систем оперативной обработки информации (АСОИИ-ЕС и АСОИИ-Cyber), баз данных гидрометеорологической информации и результатов прогнозов оперативного и исследовательского

назначения, средств визуализации данных разной степени обработки, архивации данных, формирования продукции прогностических моделей для дальнейшего распространения. Тогда же сформировалась организационная структура системы управления данными. Она включала группы администраторов баз данных, АСООИ, продукции, визуализации, которые проводили разработку, испытание и совместно с Главным вычислительным центром Росгидромета внедрение новых программных средств, консультировали авторов задач и пользователей ЭВМ.

Внедрение банка данных ПРОГНОЗ позволило существенно повысить эффективность использования гидрометеорологической информации, сделать ее доступной широкому кругу пользователей, во многом избежать таких непроизводительных затрат, как ручная подготовка данных, дублирование больших объемов информации на технических носителях, принадлежащих отдельным подразделениям, и др. Централизованные базы данных банка данных ПРОГНОЗ использовались в основных прогностических задачах, решаемых в Гидрометцентре: в объективном анализе полей метеорологических величин, краткосрочных региональных и полусферных прогнозах метеовеличин, прогнозах для авиации, прогнозах стока (ЭВМ ЕС), а также для среднесрочного прогноза погоды и модели общей циркуляции атмосферы (Cyber-172). Впоследствии на ЭВМ Hitachi в качестве внешних баз данных 15-уровневой спектральной модели (СМ-15) также использовались базы данных банка данных ПРОГНОЗ.

Система АСООИ-ЕС функционирует до настоящего времени на ЭВМ Compaq. В ее рамках выполняется несколько сеансов обработки данных, которые включают задачи расчета фактических и прогностических полей метеовеличин по региону и северному полушарию, комплекс задач анализа температуры поверхности океана, а также ряд специализированных задач.

С 1997 г. главным вычислителем Мирового метеорологического центра Москва является суперЭВМ CRAY Y-MP8E. Она имеет пиковую производительность 2,6 Гфлоп, тактовую частоту 6 нс, динамическую оперативную память 2 Гбайта, быструю статическую SSD-память 4 Гбайта, 8 центральных процессоров, суммарный объем дисковой памяти 62 Гбайта, автоматизированную ленточную библиотеку емкостью 1,2 Гбайта. Локальная вычислительная сеть (ЛВС CRAY) включает ~ 100 ПЭВМ и две рабо-

чие станции, имеет опосредованный выход в Интернет, а также связь с НИЦ „Планета” и ЦКС-Москва.

На CRAY Y-MP8E разработана и эксплуатируется достаточно развитая система управления данными, которая включает:

- систему оперативных баз данных гидрометеорологической информации разного уровня обработки;

- банк оперативных задач;

- автоматизированную систему управления оперативным счетом прогностических задач (АСООН-CRAY);

- оперативную технологию приема информации от ЦКС-Москва;

- оперативную систему раскодирования данных наблюдений и продукции зарубежных центров;

- резервную технологическую линию обеспечения CRAY Y-MP8E раскодированными данными с сервера ГИС Метео;

- систему формирования и выдачи продукции в ГИС Метео, ЦКС-Москва и Интернет;

- прикладные технологии: архивация данных всех уровней обработки на платформах CRAY и ПЭВМ, публикация некоторых видов сообщений из входного потока ЦКС-Москва на ПЭВМ оперативных подразделений ЛВС CRAY, удаленный доступ к базам данных CRAY Y-MP8E с рабочих станций и ПЭВМ ЛВС CRAY, визуализация прогностических полей на рабочих станциях ЛВС CRAY, мониторинг поступления данных наблюдений и продукции зарубежных центров.

Информационное обеспечение оперативных технологий

Общая система функционирования оперативных технологий обработки гидрометеорологической информации на CRAY Y-MP8E представлена на рис. 1. Ядром системы является распределенная информационная база данных. Она включает:

- совокупность файлов, содержащих первичные данные наблюдений, продукцию зарубежных прогностических центров в коде GRIB и различную текстовую информацию, поступающую из УГМС и наблюдательной сети;

- раскодированные данные наблюдений, размещенные в базе данных наблюдений;

- базы данных фактических и прогностических полей метео-элементов, рассчитанных по различным моделям;
- архивные базы данных;
- исследовательские базы данных, предназначенные для проведения экспериментов при разработке и развитии алгоритмов прогностических моделей.

В основу разработки системы информационного обеспечения были положены следующие принципы: все данные наблюдений,

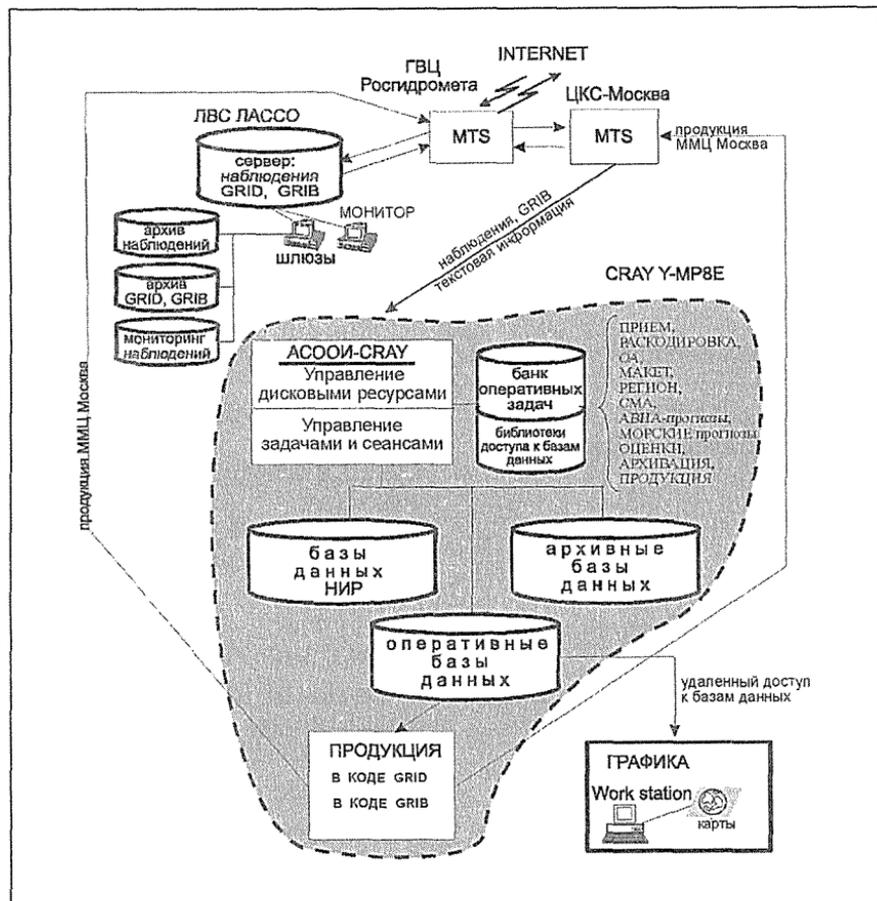


Рис. 1. Общая схема функционирования оперативных технологий обработки гидрометеорологической информации на ЭВМ CRAY Y-MP8E.

поступающие в ЦКС-Москва, должны декодироваться и размещаться в базе данных наблюдений; на каждом этапе обработки должна производиться запись результатов в соответствующие базы данных или файловые системы и по возможности их архивация; пользователи должны иметь возможность получить интересующие их данные любого уровня обработки в зависимости от степени доверия к результатам предыдущих программ обработки, при этом интерфейс должен быть достаточно дружелюбным.

Раскодированные данные наблюдений размещаются в специализированной реляционной базе данных наблюдений. Эта база оперативно используется в задачах системы объективного анализа полей метеовеличин, глобальной системой усвоения данных, а также в ряде исследовательских задач.

Для выборки данных используется библиотека операторов доступа, с помощью которой осуществляется гибкая выборка данных из программ пользователей в соответствии с указанным запросом. Состав базы данных наблюдений приведен в таблице.

Таблица

Информационный состав базы данных наблюдений

Имя отношения	Включенные коды наблюдений
ZAGL — заголовки телеграмм	SYNOP-FM12, BUOY, BATHY, TESAC, TEMP(A-D), PILOT-SHIP(A), PILOT(A), TEMP-SHIP(A-D), AIREP, SATOB, KN-24, KN-15, TEMP-DROP(A-D), SATEM(A), SHIP-FM13
SYNP — наблюдения на поверхности земли и океана	SYNOP-FM12, BUOY, BATHY, TESAC, SHIP-FM13
TEMP — аэрологические наблюдения	TEMP(A - D), TEMP-SHIP(A - D), TEMP-DROP(A - D), PILOT-SHIP(A), PILOT(A)
SHIP — синоптика моря	SHIP-FM13, BUOY
BATHY — океанография	BATHY, TESAC, BUOY
HYDR — гидрология	KN-15, KN-24
SATM — спутниковые наблюдения	SATEM(A)
SATB — спутниковые наблюдения	SATOB

Поля анализов и прогнозов, поступающие от зарубежных центров в коде GRIB, и поля, производимые системой объективного анализа и прогностическими моделями Гидрометцентра России, размещаются в группе баз данных прямого доступа с именованными ключами и общим интерфейсом (типа SHOT). Состав баз данных и период хранения данных определяются потребностями прогностических задач. В базах содержатся постоянно хранящиеся и циклически обновляемые данные, представленные в виде полей или многолетних рядов полей метеовеличин, а также нормализованных таблиц различного содержания.

Для доступа к базам данных типа SHOT в режиме чтения и записи пользователям предоставляется библиотека операторов доступа, единая для всех оперативных и архивных баз этого типа.

Объем оперативных баз данных составляет ~ 5 Гбайт. Их пользователями являются сотрудники оперативных и научных подразделений Гидрометцентра России, а также ряда НИУ Росгидромета.

Централизованный архив Гидрометцентра России на ЭВМ CRAY Y-MP8E включает: систему архивных баз данных полей объективного анализа метеовеличин, который ведется в Гидрометцентре России с 1986 г.; базу данных, содержащую исторические ряды среднемесячных полей метеовеличин (некоторые из них начинаются с 1898 г.); базу данных макетов проконтролированных данных наблюдений (с 2000 г.). Централизованный архив размещается на магнитных носителях роботизированной ленточной системы ЭВМ CRAY Y-MP8E и CD-дисках.

Доступ к данным, размещенным в оперативных и исследовательских базах данных типа SHOT, возможен как из программ пользователей, выполняющихся на ЭВМ CRAY Y-MP8E, так и непосредственно из программ, выполняющихся на ПЭВМ и рабочих станциях ЛВС CRAY. Для этого используются специализированные средства удаленного доступа — сервис remDB. Сервис remDB — это сетевое приложение, выполняющее обмен данными между разнородными операционными средами с использованием сетевого протокола TCP/IP и программного интерфейса Windows Sockets в технологии клиент-сервер. Сервер реализован на CRAY Y-MP8E, а клиент имеет несколько разновидностей в зависимости от используемой вычислительной платформы. Сервис remDB

широко применяется в информационном обеспечении оперативной и исследовательской деятельности, в частности, в задачах визуализации полей метеовеличин на рабочей станции в среде графического пакета PV-WAVE и технологии доведения до ПЭВМ синоптиков-прогнозистов некоторых видов текстовой информации, получаемой из УГМС.

Для выполнения ряда прогностических задач не требуются вычислительные ресурсы CRAY Y-MP8E, а достаточны мощности ПЭВМ при наличии полноценного информационного обеспечения. Поэтому развивается технология публикации данных непосредственно на рабочие места пользователей в соответствии с их локальными информационными потребностями. При этом может быть достигнута максимально высокая информационная поддержка. Сервис gemDB и технология публикации данных обеспечивают распределенную обработку данных в ЛВС CRAY.

Технология оперативной обработки информации

Для осуществления регулярного счета различных прогностических задач разработана, эксплуатируется и развивается автоматизированная система оперативной обработки информации (АСОИ-CRAY), которая обеспечивает следующие основные возможности:

- организация взаимодействия задач посредством общих (разделяемых) данных, размещенных в базах данных;
- группирование взаимосвязанных задач по принципу сеансов последовательной обработки данных;
- автоматический запуск сеансов согласно заданному сценарию в пакетном режиме, при этом возможно как последовательное, так и параллельное выполнение сеансов;
- протоколирование процесса счета сеансов;
- организация системы хранения и предоставления авторам листингов прикладных задач;
- централизованное сопровождение банка оперативных задач.

В текущей версии банк задач содержит 65 прогностических и специализированных задач, которые объединены в 23 различных сеанса. Всего за сутки запускается 37 сеансов и 169 задач. На рис. 2

изображена временная диаграмма функционирования оперативных технологий на ЭВМ CRAY Y-MP8E.

Связь по данным между прогностическими задачами осуществляется посредством использования системы информационных баз данных, а связь по управлению — посредством автоматизированной системы управления АСОИИ-CRAY, которая является сложным многофункциональным программным комплексом с собственной управляющей базой данных.

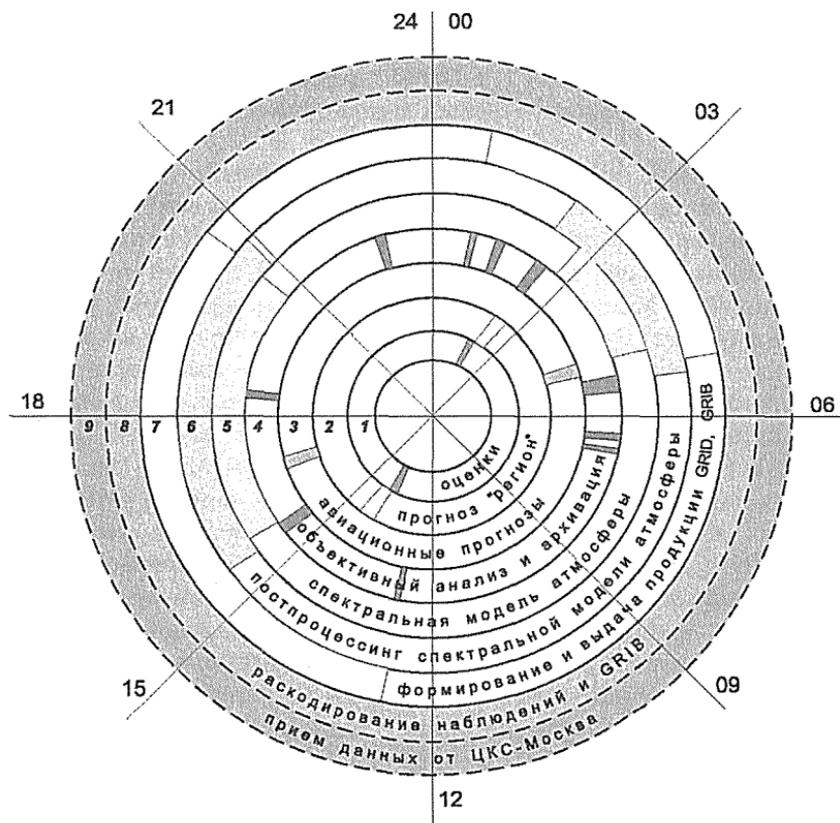


Рис. 2. Диаграмма функционирования оперативных технологий на ЭВМ CRAY Y-MP8E.

В настоящей версии за сутки формируется ~ 1460 сообщений в коде GRID и ~ 3600 сообщений в коде GRIB.

Технология АСОИИ-CRAY разработана как открытая система, которая позволяет унифицированным способом добавлять новые прогностические и специализированные задачи в банк задач, модифицировать существующие задачи, конфигурировать новые и существующие сеансы обработки данных. Средства администрирования дают возможность администратору системы и операторам ЭВМ полностью контролировать весь вычислительный процесс, протекающий в рамках АСОИИ-CRAY.

Обеспечение CRAY Y-MP8E данными наблюдений и продукцией зарубежных прогностических центров

Обеспечение ЭВМ CRAY Y-MP8E данными наблюдений и продукцией зарубежных центров может производиться по двум независимым технологическим линиям. В одной из них источником данных является сервер ГИС Метео, в другой — непосредственно ЦКС-Москва. Более перспективным является, конечно, второй вариант. Однако жесткие сроки разработки и ввода в оперативную эксплуатацию программного обеспечения технологий, а также отсутствие в период разработки в 1996 г. утвержденного протокола обмена данными между ЦКС-Москва и ЭВМ CRAY Y-MP8E привело к использованию на первом этапе раскодированных данных, размещенных в базе данных сервера ГИС Метео.

Для реализации этой линии разработана технология МОНИТОР, которая обеспечивает главный вычислитель очередной порцией информации по графику, согласованному с расписанием счета сеансов АСОИИ-CRAY. При этом в начале каждого сеанса выполняются задачи усвоения полученных данных в базах данных наблюдений и SHOT. Главным недостатком этой технологической линии является полная зависимость состава и наполнения баз данных АСОИИ от информационной базы ГИС Метео.

Для реализации линии информационного обеспечения ЭВМ CRAY Y-MP8E непосредственно от ЦКС-Москва разработана и в настоящее время находится в режиме оперативного использования технология приема и усвоения потока данных, поступающих из ЦКС. Прием данных выполняется по протоколу, установлен-

ных кодовых форм, запись обработанных данных в базы данных наблюдений и GRIB, а также формирование файла текстовой информации.

Для обеспечения доступа к раскодированной информации пользователям предоставляются:

- библиотека пользовательского интерфейса с базой наблюдений на суперЭВМ CRAY Y-MP8E;

- библиотека доступа к данным, размещенным в базах данных типа SHOT;

- программное обеспечение удаленного доступа к текстовым сообщениям о погоде, а также СГЯ и РИП в архитектуре клиент (ПЭВМ) — сервер (CRAY);

- программное обеспечение рабочего места сотрудника на ПЭВМ в сети CRAY для визуализации, анализа и архивации перечисленных текстовых сообщений.

Для архивации данных наблюдений обеспечивается хранение ежемесячной копии 36-суточной базы наблюдений в роботизированном архиве CRAY.

В рамках технологии реализован доступ ко всем базам данных в режиме разделения времени, т. е. обеспечена процедура чтения данных из программ пользователей на фоне непрерывного прямого усвоения данных в процессе их раскодирования.

Можно констатировать, что в настоящее время в Гидрометцентре России создана достаточно эффективная система управления данными, на основе которой функционируют основные оперативные технологии на суперЭВМ CRAY Y-MP8E и осуществляется выпуск прогностической продукции и проведение научных исследований. Приоритетным направлением дальнейшего развития должно стать создание новых технологий распределенной обработки данных в сети с использованием централизованной системы баз данных, а также совершенствование систем долговременного хранения информации.