

Информационно-вычислительные ресурсы ИРНОК: инфраструктура, данные, приложения

И. В. Бычков, Т. И. Маджара, А. П. Новопашин, Е. С. Фереферов*,
А. Г. Феоктистов, Р. К. Федоров

Институт динамики систем и теории управления им. В.М. Матросова СО РАН, 664033,
Иркутск, Россия

*Контактный автор: Фереферов Евгений Сергеевич, e-mail: fereferov@icc.ru

Поступила 21 апреля 2023 г., доработана 24 апреля 2023 г., принята в печать 28 апреля 2023 г.

В статье представлены этапы развития информационно-вычислительной и телекоммуникационной инфраструктуры Иркутского научно-образовательного комплекса в период с 2008 по 2022 г. Показано развитие инфраструктуры: сети передачи данных, систем хранения данных, виртуализации, вычислительных ресурсов, в том числе оборудования Иркутского суперкомпьютерного центра.

Ключевые слова: информационно-вычислительные и телекоммуникационные ресурсы и платформы, высокопроизводительные вычислительные системы, системы хранения данных, системы виртуализации, геопортал, инструментальные средства.

Цитирование: Бычков И.В., Маджара Т.И., Новопашин А.П., Фереферов Е.С., Феоктистов А.Г., Федоров Р.К. Информационно-вычислительные ресурсы ИРНОК: инфраструктура, данные, приложения. Вычислительные технологии. 2023; 28(3):117–135. DOI:10.25743/ICT.2023.28.3.008.

Введение

В институтах СО РАН, расположенных на территории г. Иркутска и области, накопена и постоянно актуализируется уникальная информация, относящаяся к различным отраслям наук: математике, химии, наукам о земле, солнечно-земной физике, сферам энергетики, биологии и биоразнообразию, социально-эколого-экономическому состоянию территорий. Особое значение придается формированию знаний и данных об оз. Байкал и Байкальской природной территории (БПТ).

Несмотря на большой объем, накопленные информационные ресурсы ввиду их локализации в разных институтах не всегда могли быть эффективно использованы при проведении междисциплинарных фундаментальных и прикладных научных исследований, что потребовало внедрения современных форм коллективного использования информационно-вычислительных и телекоммуникационных ресурсов и технологий, в том числе создания и развития систем распределенных вычислений, систем хранения и обработки данных, а также устойчивой и высокоскоростной телекоммуникационной инфраструктуры.

Развиваясь в данном ключе уже в течение 25 лет, Интегрированная информационно-вычислительная сеть Иркутского научно-образовательного комплекса (ИИВС ИРНОК)

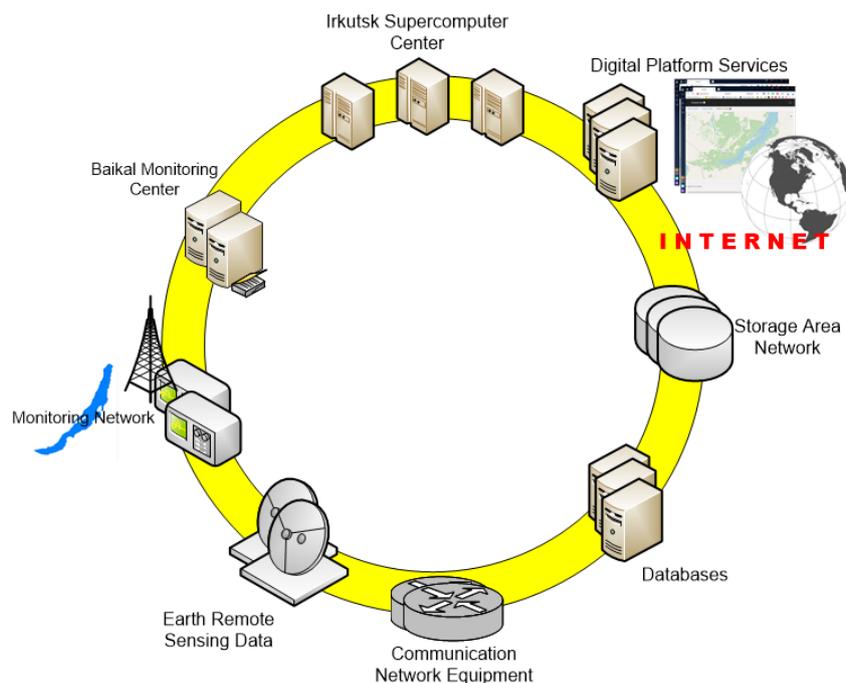


Рис. 1. Концептуальная схема ИИВС

Fig. 1. The conceptual scheme of IICN IREC

является базовой технологической платформой для большинства междисциплинарных исследований, проводимых институтами региона (рис. 1).

Началом создания инфраструктуры ИРНОК считается 1994 г., когда Институтом динамики систем и теории управления (ИДСТУ) Сибирского отделения РАН был разработан системный проект ИИВС ИРНОК и начата его реализация [1, 2]. На начальных этапах основной задачей проекта являлось объединение корпоративных информационно-вычислительных ресурсов научных и образовательных учреждений Иркутска (до 2002 г.), затем Байкальского региона (до 2006 г.) высокоскоростными каналами связи, а также поддержка и развитие доступа в Интернет. Проект был реализован за счет финансовых средств Министерства науки и технологий РФ, РФФИ, СО РАН, а также внебюджетных средств ИДСТУ СО РАН, академических институтов ИНЦ СО РАН и ряда государственных вузов. В результате к 2006 г. ИРНОК объединил 12 академических институтов СО РАН, 4 института СО РАМН, 5 вузов Иркутска, обеспечил высокоскоростной доступ к корпоративным сетям Бурятского научного центра (г. Улан-Удэ) и приобрел статус регионального узла связи.

С момента создания в 2005 г. на базе ИДСТУ СО РАН Центра коллективного пользования (ЦКП) «Иркутский суперкомпьютерный центр СО РАН» (ИСКЦ) [3] ИИВС ИРНОК стал играть ключевую роль в обеспечении доступа к высокопроизводительным вычислительным ресурсам центра.

1. Развитие технологической платформы ИИВС ИРНОК

Сеть связи. В 2009–2010 гг. проводится глубокая модернизация всей кабельной инфраструктуры ИИВС ИРНОК. Выработавшие на тот момент свой ресурс многомодовые оптические линии связи выводятся из штатной эксплуатации и заменяются на более перспективный одномод. С учетом новых требований значительно изменяется и топология сети, традиционная звездообразная структура сменяется более надежной

кольцевой. В ядре сети передачи данных (СПД) устанавливается новое магистральное оборудование, уровень доступа поэтапно переводится на пропускную способность в 1 Гбит/с. В 2009 г. происходит реструктуризация внешних подключений сети, существенно расширяются каналы доступа в Интернет и в Сеть передачи данных СО РАН, организуется новый канал связи до Байкальского музея СО РАН в пос. Листвянка (60 км от Иркутска).

Необходимо отметить, что вплоть до реформы государственных академий наук 2013 г. и четырех последующих за этим лет значительную роль в развитии и поддержке внешних каналов связи играл Институт вычислительных технологий СО РАН, фактически обеспечивающий их централизованную закупку и техническое сопровождение в интересах всех институтов Сибирского отделения РАН. С 2018 г. ИДСТУ СО РАН уже самостоятельно продолжает эту работу, сохранив единое адресное пространство ИИВС ИРНОК и увеличив внешний канал связи для ее участников в два раза — до 320 Мбит/с (табл. 1).

В целях более глубокой интеграции с научно-образовательным сообществом в 2022 г. ИИВС ИРНОК присоединилась к Научно-исследовательской компьютерной сети России (НИКС), крупнейшей в стране научно-образовательной телекоммуникационной сети, являющейся опорной сетью национального уровня, которая обладает протяженной высокоскоростной магистральной инфраструктурой (более 5000 км) и международными каналами, обеспечивающими интеграцию с зарубежными научно-образовательными сетями. Канал, объединивший ИИВС ИРНОК и НИКС, имеет пропускную способность до 10 Гбит/с. Вместе с подключением к НИКС пропускная способность основной магистрали ИИВС также увеличилась до 10 Гбит/с. Актуальное состояние логической структуры сети представлено на рис. 2.

Платформа виртуализации. К 2013 г. успешно функционирующая по традиционной схеме инфраструктура ИИВС ИРНОК, отвечая новым задачам, значительно трансформируется, принимая вид, представленный на рис. 3. Ключевым компонентом в ней становится система виртуализации, отделяющая операционные системы и функционирующие на ее базе сервисы и порталы непосредственно от вычислительного оборудования, систем хранения данных, сетей и систем инженерно-технического обеспечения,

Т а б л и ц а 1. Этапы развития каналов связи
Table 1. Stages for development of communication channels

Год	Пропускная способность каналов, Мбит/с
1994–2002	0.128 – 0.256 – 0.512 (Релком)
2002	2 (СПД СО РАН-RBNet)
2003	4 = 2×2 (СПД СО РАН-RBNet)
2005	14 = 10 + 2×2 (СПД СО РАН-RBNet)
2006	16 = 10 + 2×2 (СПД СО РАН-RBNet) + 2 (БНЦ*)
2007	22 = 20 (СПД СО РАН-RUNNet) + 2 (БНЦ)
2009	52 = 40 (Synterra) + 10 (СПД СО РАН) + 2 (БНЦ)
2010	90 = 80 (Ростелеком) + 10 (СПД СО РАН)
2011	138 = 80 (Synterra) + 4+54 (СПД СО РАН)
2013	160 (Транстелеком, Милеком)
2018	320 (ЭР-Телеком)
2022	10 320 = 320 (ЭР-Телеком) + 10 000 (НИКС)

* Бурятский научный центр

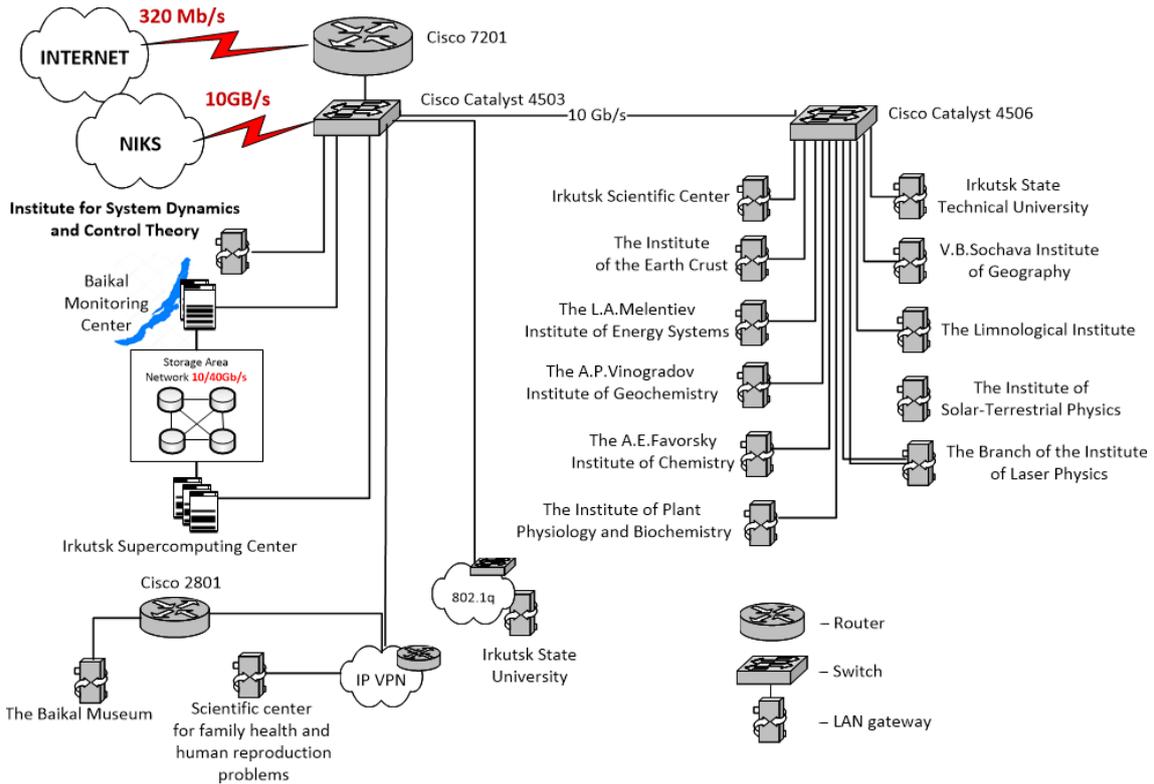


Рис. 2. Логическая структура сети ИИВС ИРНОК
 Fig. 2. The logical structure of the IICN IREC network



Рис. 3. Технологическая инфраструктура ИИВС с 2013 г.
 Fig. 3. The technological infrastructure of IICN IREC since 2013

существенно упрощая процессы разработки, тестирования и эксплуатации сетевых приложений и позволяя более эффективно распределять дорогостоящие ресурсы. Перенос сервисов и приложений в виртуальные машины также позволил осуществлять техническое обслуживание и модернизацию новой инфраструктуры без остановки функционирования системы в целом. Программно-аппаратный комплекс виртуализации на стартовом этапе предоставляет 48 вычислительных ядер и 448 Гб оперативной памяти.

В 2015 г. с приобретением дизель-генераторной установки NIMOINSA HDW-525 мощностью 500 кВА получает полную автономность система энергоснабжения Блока ЭВМ ИДСТУ СО РАН, где развернута основная часть оборудования ИИВС ИРНОК, а также оборудование доступа к внешним каналам связи.

В 2020 г. в связи с необходимостью обеспечения выполнения задач по крупному научному проекту (КНП) “Фундаментальные основы, методы и технологии цифрового мониторинга и прогнозирования экологической обстановки Байкальской природной территории” программно-аппаратный компонент системы виртуализации значительно расширяется: в два раза возрастает количество базовых серверов, скорость их доступа к сетевой подсистеме увеличивается до 10 Гбит/с. После проведенной модернизации программно-аппаратный комплекс имеет в своем составе 128 вычислительных ядер и 960 Гб оперативной памяти.

Системы хранения данных. В 2009 г. в инфраструктуре ИИВС появляется первая система хранения данных “сырым” объемом 62 Тб. Построение новой системы хранения на базе топологии сетей SAN (storage area network) во многом определяет дальнейшее развитие этого неотъемлемого компонента ИИВС. В период с 2014 по 2020 г. в рамках работ по ряду интеграционных (междисциплинарных) проектов объем сети хранения данных (СХД) возрастает до 195 Тб, ее инфраструктура дополняется новым оборудованием и функционалом. Дисковый ресурс, предоставляемый сетью хранения, в зависимости от задач гибко распределяется между серверами системы виртуализации и стационарным серверным сегментом, обеспечивая три уровня хранения по скорости доступа: кэш, оперативный, архивный. В составе сети хранения появляются гибридные системы, позволяющие предоставлять дисковые ресурсы по набирающему популярности протоколу iSCSI. В 2020 г. для размещения данных дистанционного зондирования Земли (ДЗЗ), а также результатов их обработки в целях КНП сеть хранения данных получает дальнейшее развитие. Установка новой гибридной системы хранения большого объема доводит общую емкость сети до 992 Тб (табл. 2).

GPU-кластер. Помимо качественного развития сети хранения данных КНП положил начало развитию еще одного компонента ИИВС — GPU-кластера (рис. 4), распределенной вычислительной системы, использующей в качестве вычислительного ресурса графические процессоры (GPU). Вычислительные ресурсы и программное обеспечение GPU-кластера предоставляют широкие возможности по обработке данных (в том чис-

Т а б л и ц а 2. Динамика развития сети хранения данных ИИВС ИРНОК

Table 2. Development dynamic of storage systems of IICN IREC

Год	Общий ресурс сети хранения данных, Тб
2009	62
2014	90.8
2017	133
2020	548
2021	992

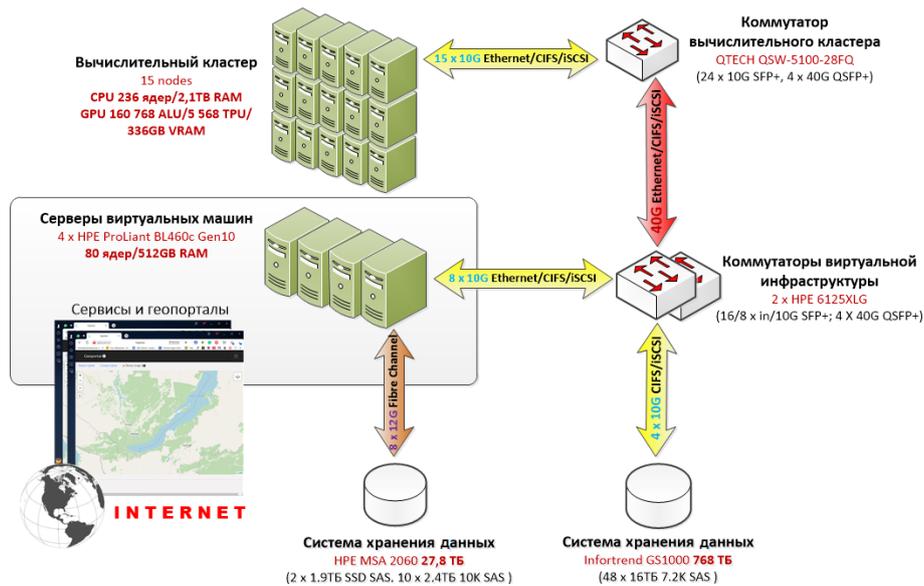


Рис. 4. GPU-кластер
Fig. 4. The GPU cluster

ле ДЗЗ) с использованием перспективных технологий нейронных сетей. Пропускная способность магистрального сетевого сегмента кластера составляет 40 Гбит/с.

Ситуационный центр. Результаты применения современных высокопроизводительных вычислительных систем и специально разработанного программного обеспечения для поддержки комплексного научного онлайн-мониторинга различных экологических показателей, анализа и прогнозирования состояния окружающей среды в рамках крупного научного проекта “Фундаментальные основы, методы и технологии цифрового мониторинга и прогнозирования экологической обстановки Байкальской природной территории” можно наблюдать в открытом в 2022 г. на базе ИИВС Ситуационном центре цифрового мониторинга и прогнозирования экологической обстановки Байкальской природной территории. Ситуационный центр является востребованной мультимедиа-площадкой для проведения тематических семинаров, видеоконференций и совещаний по вопросам, связанным с целями как КНП, так и смежных научных проектов.

2. Развитие вычислительных ресурсов

С 2005 г. на базе ИДСТУ СО РАН функционирует Центр коллективного пользования “Иркутский суперкомпьютерный центр СО РАН” [3]. Центр зарегистрирован в федеральном каталоге “Научно-технологическая инфраструктура Российской Федерации” [4], соответствует требованиям к центрам коллективного пользования научным оборудованием, функционирование которых обеспечивается с привлечением бюджетных средств.

Деятельность ЦКП ИСКЦ соответствует принципам открытости, доступности имеющегося оборудования и направлена на развитие, поддержку и предоставление в коллективное пользование высокопроизводительных вычислительных ресурсов для выполнения расчетов (вычислений). В соответствии с регламентом приоритетным является предоставление ресурсов научным и образовательным организациям, подведомственным Минобрнауки России, для выполнения НИР в рамках государственных заданий.

По результатам экспертной оценки, организованной Научно-координационным советом при ФАНО России в 2015 г., ИСКЦ вошел в число суперкомпьютерных центров первой категории. При составлении рейтинга оценивались состояние имеющейся инфраструктуры, ее востребованность при проведении фундаментальных и прикладных исследований, перспективы развития. На протяжении последующих лет ИСКЦ сохраняет за собой статус одного из ведущих академических суперкомпьютерных ЦКП. Обновление оборудования центра осуществляется с периодичностью в 4–6 лет, что является нормой в условиях стремительных изменений в IT-отрасли.

В ИСКЦ за время существования сменилось несколько поколений вычислительных систем коллективного пользования (рис. 5).

В 2005–2012 гг. эксплуатировались вычислительный кластер МВС-1000 производства ФГУП НИИ “Квант” и ИПМ им. М.В. Келдыша РАН, а также вычислительный кластер собственной сборки Blackford [2].

В 2011–2012 гг. при поддержке Президиума СО РАН и Приборной комиссии СО РАН в ИДСТУ СО РАН реализован проект создания вычислительного кластера нового поколения, в дальнейшем названного именем первого директора-организатора института — академика В.М. Матросова (рис. 6). Основу масштабируемой вычислительной инфраструктуры вычислительного кластера “Академик В.М. Матросов” составили серверные системы производства российской компании “Т-Платформы”, объединенные коммуникационной сетью QDR Infiniband. Для обеспечения непрерывной работы кластера установлен комплекс инженерного оборудования, включающий системы энергоснабжения, бесперебойного питания, охлаждения, автоматического газового пожаротушения и др. Вычислительный кластер “Академик В.М. Матросов” на момент ввода в эксплуатацию имел в своем составе 110 вычислительных узлов T-Blade V205S, 220 16-ядерных процессоров AMD Opteron 6276 “Bulldozer”; вычислительный узел с графическими процессорами NVidia C2070 “Fermi”. Пиковая производительность кластера составляла 33.7 Тфлопс, реальная производительность на тестах HPL — 25.12 Тфлопс (26-е место в 16-й редакции ТОП-50 [5]).

Характеристики	«МВС-1000»	«Blackford»	«Академик В.М. Матросов»	«Ampere»
Тип архитектуры	Кластер SMP-узлов	Кластер SMP-узлов	Кластер SMP-узлов	GPU-сервер
Модель и характеристики CPU / GPU	 Intel Xeon DP (Prestonia), 1C, 2.667 GHz, 130 nm	 Intel Xeon E5345 (Clovertown), 4C, 2.33 GHz, 65 nm	 AMD Opteron 6276 (Interlagos), 16C, 2.3 GHz, 32 nm  Intel Xeon E5-2695 v4 (Broadwell-EP), 18C, 2.1 GHz, 14 nm	 Intel Xeon Gold 6326 (Ice Lake-SP), 16C, 2.9 GHz, 10 nm  Nvidia A100 80GB PCIe, 7 nm
Количество вычислительных узлов	16	20	120	1
Количество CPU / ядер CPU	32 / 32	40 / 160	240 / 4080	2 / 32
Количество GPU / тензорных ядер GPU	- / -	- / -	- / -	4 / 1728
Вычислительная (системная) сеть, пропускная способность, Gbps	Myrinet 2000, 2	GigaEthernet, 1	QDR Infiniband, 40	-
Пиковая (теоретическая) производительность FP64, Tflops	0,171	1,491	90,24	80,97
Реальная (HPL benchmark) производительность FP64, Tflops	0,097	0,924	77,51	59,05
Наивысшее место в ТОП-50 России и СНГ	-	41	26	-

Рис. 5. Характеристики вычислительных ресурсов ИСКЦ

Fig. 5. Characteristics of the ISCC computing resources



Рис. 6. Вычислительный кластер “Академик В.М. Матросов”
Fig. 6. The computing cluster “Academician V.M. Matrosov”

В 2017 г. в рамках пилотного этапа модернизации суперкомпьютерных центров научных учреждений, подведомственных ФАНО России, введен в эксплуатацию второй сегмент вычислительного кластера “Академик В.М. Матросов”, построенный на базе серверов-лезвий Supermicro SuperBlade SBI-7228R и 18-ядерных процессоров Intel Xeon E5-2695 v4 “Broadwell”. В результате модернизации общее число доступных вычислительных узлов кластера увеличилось до 120, их суммарная пиковая производительность достигла 90.24 Тфлопс, реальная — 77.51 Тфлопс.

В конце 2022 г. вычислительная инфраструктура ИСКЦ пополнилась высокопроизводительным сервером на базе графических процессоров NVidia A100 “Ampere”, ориентированным на решение задач численного моделирования и искусственного интеллекта. Пиковая производительность сервера составила 80.97 Тфлопс (FP64) / 624 Тфлопс (TF32), реальная — 59.05 Тфлопс (на тестах HPL) / 232 Тфлопс (на тестах HPL-AI).

Ресурсы ИСКЦ подключены к Национальной исследовательской компьютерной сети России [6] по каналу с пропускной способностью 10 Гбит/с. К настоящему времени



Рис. 7. География пользователей ИСКЦ
Fig. 7. Geography of ISCC users

в ИСКЦ зарегистрировано более 150 пользователей из 20 научных и образовательных организаций, географически распределенных от Санкт-Петербурга до Владивостока (рис. 7).

3. Данные и приложения

Для поддержки междисциплинарных исследований (сбор, обработка и представление данных) в ИДСТУ СО РАН разработана технология развертывания оригинальных предметно-ориентированных геопорталов [7], реализующих общие для информационных систем функции, достаточные для поддержки большинства исследований. Для каждого научного проекта геопортал разворачивается в облачной инфраструктуре ИИВС ИРНОК. Создание геопортала производится на основе заранее сконфигурированных шаблонов виртуальных машин. Геопортал реализован с использованием стандартов OGC, что позволяет формировать удобные сервисы с унифицированными методами обмена данными и гибким пользовательским интерфейсом. В типовом геопортале реализованы инструменты создания сервисов данных и их композиций. Сервисы данных предназначены для организации ввода и редактирования реляционных данных с пространственной привязкой и предоставления данных на основе REST. Для этого пользователь определяет структуру данных, способы отображения и права доступа. Реализована регламентированная передача данных для обработки WPS-сервисами. В рамках геопортала можно создавать новые WPS-сервисы путем композиции существующих сервисов, используя язык JavaScript или JSON спецификации DAG. Реализованные технологии обеспечивают формирование и передачу данных, запуск сервисов, организацию вычислительного процесса в облачной и кластерной инфраструктуре, представление результатов. Некоторые из функций вынесены на отдельные узлы облачной среды.

Сеть хранения данных является важнейшим компонентом ИИВС ИРНОК и обеспечивает для всех участников надежное хранение научных данных. Хранящиеся в СХД данные являются результатами исследований как отдельных научных организаций, так и междисциплинарных интеграционных проектов. В основном это пространственные данные в виде векторных карт, данные дистанционного зондирования Земли и тематические базы данных. Так, в результате интеграционного проекта “Фундаментальные исследования и прорывные технологии как основа опережающего развития Байкальского региона и его межрегиональных связей” (2016–2019 гг.) были собраны данные:

- минерализация придонного слоя воды на основе данных измерений удельного электрического сопротивления грунта методами электротомографии, полученных в рамках гидрологической съемки (с помощью CTD-зонда Sea-Bird Electronics SBE-37);
- химический состав воды в реках, ручьях, лунках пляжей и литорали залива “Лиственничный”;
- химический состав придонной воды на трансектах в заливе “Лиственничный” оз. Байкал;
- химический состав подземных вод в скважинах и колодцах;
- содержание антропогенных загрязнителей в природных средах, которые находятся в прямой связи между собой, в рамках ландшафтно-геохимических потоков;
- описание и координаты мест отбора проб воды из рек, лунок и скважин береговой зоны оз. Байкал и его прибрежной акватории в районе пос. Листвянка;
- содержание взвешенного вещества в снеговом покрове;

- определяемые показатели в почвах, грунтовых водах, ледовом покрове, поверхностных водах, подледной воде и воде Байкала, репрезентативных видах растительного покрова;
- данные ДЗЗ WorldDem на район пос. Листвянка (31 650 км²);
- результаты 3D-моделирования подводного и надводного рельефа в районе пос. Листвянка.

Одним из междисциплинарных проектов, интенсивно использующих ресурсы ИИВС ИРНОК, стал крупный научный проект “Фундаментальные основы, методы и технологии цифрового мониторинга и прогнозирования экологической обстановки Байкальской природной территории” [8]. В рамках проекта разработана и развернута в ИИВС ИРНОК цифровая платформа, обеспечивающая сбор данных с распределенной сети комплексного экологического мониторинга (5 направлений мониторинга, интервал регистрации 1–5 с, ежедневно более 20 ГБ разноформатных данных), хранение, обработку больших массивов гетерогенных пространственно-временных данных, а также анализ и прогнозирование экологической обстановки на основе комплекса математических моделей, сервисов и методов машинного обучения.

Цифровая платформа мониторинга БПТ включает средства доставки, хранения пространственно-временных данных в СХД и их обработки на основе цифровых сервис-ориентированных и сквозных технологий, декларативных спецификаций и интеллектуализации, а также инфраструктуру развертывания прикладных систем, средства разработки и интеграции прикладных сервисов участниками мониторинга БПТ для принятия решений. Для повышения эффективности проведения прикладных научных исследований и интеграции прикладных сервисов ведется разработка платформы на основе вычислительной модели композиции сервисов, которая позволяет объединять сервисы, созданные разными коллективами. В рамках модели предложен метод автоматической композиций сервисов на основе статистических данных. Использование метода упрощает работу пользователя, автоматизирует часто повторяемые действия и формирует алгоритмы применения сервисов. Разработан комплекс программных компонентов, позволяющий создавать новые сервисы, готовые для включения в композиции.

Для сбора данных в реляционном виде и публикации в виде таблиц и карт в рамках цифровой платформы экологического мониторинга в ИИВС ИРНОК развернут геопортал. Для сбора данных в виде файлов развернута система Next Cloud (набор клиент-серверных программ для создания и использования хранилища данных), которая предоставляет пользователям систему хранения файлов и обмена данными, при этом каждый пользователь имеет список контактов, возможность предоставить свои данные всем или конкретным пользователям. Для обработки данных применяется среда Jupyter Notebook как наиболее мощный инструмент для анализа в интерактивном виде, который объединяет программный код, математические уравнения и визуализацию в виде одного документа. Для организации многопользовательской работы используется JupyterHub, который позволяет настраивать рабочее пространство: директории пользователей, выбранный Docker-образ, заданные характеристики вычислительных узлов и другие параметры. Интеграция JupyterHub, NextCloud и геопортала происходит на уровне системы хранения данных и общей схемы авторизации. Регистрация пользователей производится в системе NextCloud и предоставляется директория в СХД, в которой пользователь может хранить собственные данные в виде файлов. Эта директория также предоставляется пользователю в геопортале и Jupyter Notebook. В геопортале реализована поддержка открытого протокола авторизации OAuth 2.0, который позволяет

предоставить третьей стороне ограниченный доступ к защищенным ресурсам пользователя без необходимости передавать ей (третьей стороне) логин и пароль.

Особое внимание в рамках КНП уделяется методам мониторинга на основе обработки данных ДЗЗ. Собран значительный объем информации: космоснимки Sentinel-2 на Иркутскую область и Республику Бурятия за три года занимают более 57 ТБ (21 тысяча снимков). Для организации удобного поиска космоснимков и быстрого доступа к ним создан каталог данных ДЗЗ, реализованы возможность отображения снимков на карте и поддержка получения космоснимков земной поверхности из различных источников. Снимки находятся в СХД ИРНОК, их регистрация производится разработанными компонентами сбора данных. Каждый снимок представлен на карте в виде ограничивающего прямоугольника. Отображаются космоснимки в двух режимах (для Sentinel-2 это native RGB и RGB на основе каналов B12, B8A, B04). Реализован поиск снимков с учетом сенсора, облачности, положения и даты.

Для организации экологического мониторинга на больших территориях разработаны технология и Web-сервис классификации мультиспектральных космоснимков Sentinel-2. Классификация космоснимков проводится с использованием нейронной сети ResNet50. Технология позволяет определить набор классов (типов поверхностей), подготовить обучающую выборку, провести верификацию и выполнить классификацию большого количества снимков. Обучающая выборка формируется с помощью Web-интерфейса на основе космоснимков Sentinel-2 и полевых исследований. В качестве входных данных нейронная сеть принимает 13 каналов Sentinel-2 и канал, построенный на основе локальных бинарных шаблонов для учета текстурных характеристик. В рамках тестирования для БПТ выделены 25 характерных классов объектов: болота, голая скала, пастбище, вырубки, переходный лес/кустарники, хвойный лес, редколесье, лиственный лес, смешанный лес, облака, пашня, жилая зона, вода. Размечены 134 космоснимка БПТ, количество полигональных объектов 2246 шт., площадь более 5900 кв. км. Средняя точность классификации составила 95%. Разработанный Web-сервис позволяет повысить оперативность мониторинга больших территорий. Точность результатов показала, что технология классификации может применяться для решения актуальных задач БПТ, в частности для анализа изменений лесного фонда, оценки влияния изменений климата на ландшафт, анализа динамики застройки, инвентаризации сельхозугодий и т. д.

Для поддержки работы с топоосновой разработан оригинальный Web-сервис получения фрагмента карт желаемого масштаба на задаваемый пользователем фрагмент территории. Для реализации сервиса использован код ранее разработанного конвертера из формата SXF в формат Shape, который в ходе работы объединяет данные обрабатываемых исходных трапеций, получая таким образом только несколько десятков результирующих слоев. Для использования в Web-сервисе дополнительно в конвертере были реализованы функции: обрезка фрагмента по полигональному контуру и генерация стилевых файлов QGIS (файлов qml) по информации из классификатора ГИС Панорама. После формирования всех слоев с результатами запроса содержимое временной папки сжимается в ZIP-архив, а сама папка удаляется. Ссылка на ZIP-архив для скачивания возвращается в качестве результатов запроса. Архивы с результатами запроса сохраняются долговременно до тех пор, пока не будет исчерпано выделенное для таких файлов место.

В результате работ по крупным научным проектам созданы и развернуты на базе ИРНОК следующие сервисы:

- получения и обработки данных ДЗЗ;
- предоставления цифровой топоосновы;
- проведения атмосферной коррекции спутниковых данных;
- моделирования динамики лесных ресурсов;
- анализа и прогнозирования риска (опасности) лесного пожара на основе информации о классе пожароопасности лесов, метеоусловий и других факторов;
- обработки данных со спутника Sentinel-5P о загрязнении атмосферы;
- сбора и обработки данных о загрязнении приземных слоев атмосферы БПТ;
- выявления и оценки изменений состояния растительного покрова по временным сериям данных ДЗЗ;
- сейсмического районирования;
- геокодирования мест регистрации укусов иксодовыми клещами;
- оценки влияния лесных пожаров на органы дыхания человека;
- расчета концентрации хлорофилла *a*.

В ходе реализации крупных научных проектов на ресурсах ИИВС ИРНОК были сформированы базы данных:

- абиотических показателей литорали оз. Байкал по данным мониторинговых станций в пос. Большие Коты, Наушки, Листвянка и оз. Арахлей;
- уровней отдельных рек (Иркут, Селенга, Баргузин);
- ледовой обстановки на переправе Татаурово — Турунтаево;
- “Уровень оз. Байкал с 2015 по 2020 г.”;
- “Уровень оз. Байкал с 2015 по 2021 г.”;
- “Содержание SO₂, NO₂, O₃, CO (Ангарск, Иркутск, Шелехов, Байкальск, Селенгинск, Улан-Удэ, Гусиноозерск)”;
- распространения инвазивных видов растений Байкальской Сибири;
- центра хранения и оцифровки гербарных и прочих коллекций на базе ЦКП “Биоресурсный центр” СИФИБР СО РАН;
- распространения видов Красной книги Иркутской области;
- поглощения углерода и продуцирования кислорода лесами БПТ;
- насекомых — вредителей деревьев;
- возбудителей болезней леса;
- спутникового мониторинга термальных точек (лесных пожаров) по данным прибора AVHRR спутников серии NOAA (POES);
- лесонарушений (гарей от пожаров лесного фонда) по данным ДЗЗ;
- мониторинга грозовых разрядов;
- “Содержание стойких органических загрязнителей в почвах побережья оз. Байкал”;
- “Содержание тяжелых металлов, главных ионов и биогенных элементов в стоке оз. Байкал”;
- “Содержание стойких органических загрязнителей в снеговом покрове Прибайкалья в 1994–2021 гг.”;
- результатов мониторинга климатических характеристик (температура и влажность воздуха, направление и скорость ветра, сумма осадков, суммарная приходящая солнечная радиация, температура почвы (до 80 см), влажность почвы);
- данных с атмосферно-почвенных измерительных комплексов на БПТ;
- “Активность иксодовых клещей”.

С целью поддержки процессов подготовки и проведения крупномасштабных экспериментов на основе интегрированного использования ресурсов ЦКП ИСКЦ, грид-сис-

тем и облачных платформ разработаны специализированные инструментальные средства [9]. Данные средства поддерживают технологию автоматизации процесса решения крупномасштабных научных задач, включая разработку, модификацию и комплексирование модулей пакетов для решения разных классов задач, их непрерывную интеграцию, доставку и развертывание в узлах среды, автоматизацию построения и выполнения схем решения задач, мультиагентное управление вычислениями в гетерогенной среде.

Предметно-ориентированная среда создается с помощью инструментального комплекса Orlando Tools, реализующего модульный подход к разработке и применению распределенных пакетов прикладных программ (РППП). Orlando Tools предоставляет пользователям как текстовый, так и графический языки описания предметной области РППП и постановок задач. Постановки задач могут формулироваться в процедурной и непроцедурной форме. В последнем случае производится автоматический синтез плана решения задачи (абстрактной программы). План решения задачи, сформированный на основе ее процедурной постановки, может включать управляющие конструкции ветвления, цикла и рекурсии.

Управление вычислениями в гетерогенной среде реализуется мультиагентной системой с иерархической структурой. Она включает несколько уровней функционирования агентов. На каждом уровне могут функционировать агенты, играющие различные роли и выполняющие функции, которые соответствуют этим ролям. Роли агентов (постоянные или временные) возникают в дискретные моменты времени в связи необходимостью организации коллективного взаимодействия. Агенты автономны. Они способны объединяться в виртуальные сообщества, а управление вычислениями основано на их локальных взаимодействиях посредством сотрудничества или конкуренции. При выполнении заданий агенты осуществляют распределение ресурсов на основе экономических механизмов регулирования их спроса и предложения. В процессе вычислений агенты могут перераспределять свою вычислительную нагрузку среди других агентов. Ключевые этапы управления заданиями — применение системы классификации заданий и параметрическая настройка алгоритмов функционирования агентов, основанная на имитационном моделировании.

Оборудование ИСКЦ используется для решения задач, в том числе относящихся к следующим научным областям: органическая и элементо-органическая химия, химия комплексных соединений, физика твердого тела, моделирование наноструктур, геномика и протеомика, экология биосистем, микробиология и вирусология, физика Солнца, физика плазмы, физика элементарных частиц, физика ядра и физика ускорителей, дискретная оптимизация.

Примеры научных проектов, реализуемых с использованием оборудования ИСКЦ последние три года:

- “Исследование физическими методами и методами квантовой химии структурных особенностей и механизмов реакций новых гетероатомных и элементоорганических соединений на молекулярном и супрамолекулярном уровнях” (ИрИХ СО РАН);
- “Концептуально новые энерго- и ресурсосберегающие, атом-экономные и экобезопасные методы органического синтеза и фундаментальные химические реакции на основе ацетилен и других продуктов газо-, нефте- и углерепереработки в интересах фармацевтики, медицинской химии, высоких технологий и оригинальных импортозамещающих малотоннажных производств” (ИрИХ СО РАН);

- “Материалы и технологии для разработки радиационных детекторов, люминофоров и оптических стекол” (ИГХ СО РАН);
- “Исследование физико-химических процессов в молекулярных и надмолекулярных системах методами математической физики, квантовой химии и молекулярной динамики” (ИХКГ СО РАН);
- “Разработка фундаментальных основ перспективных оптических и магнитных материалов и синхротронных рентгеноспектральных методов исследования вещества” (СФУ);
- “Создание и исследование новых металлических, керамических, интерметаллидных, композиционных материалов и наноструктурных покрытий с высокими физико-химическими и эксплуатационными свойствами” (ХФИЦ ДВО РАН);
- “Физика низкоразмерных структур и полупроводниковых наноматериалов” (ИАПУ ДВО РАН);
- “Разработка и применение новых физико-химических методов для изучения живых организмов и создания биоинспирированных материалов” (ЛИН СО РАН);
- “Исследования вирусных и бактериальных сообществ как основы стабильного функционирования пресноводных экосистем и эффективного ответа в условиях антропогенного воздействия” (ЛИН СО РАН);
- “Исследование трансформаций состояния водоемов и водотоков Восточной Сибири в сезонных и долговременных аспектах в контексте изменений климата, геологической среды и антропогенных нагрузок” (ЛИН СО РАН);
- “От клетки — к экосистеме: Исследование ультраструктуры гидробионтов и их сообществ в эволюционно-экологическом аспекте методами клеточной биологии и геномики” (ЛИН СО РАН);
- “Генетика сообществ байкальских организмов: структура генофонда, стратегии консервации” (ЛИН СО РАН);
- “Исследование новых методов ускорения заряженных частиц” (ИЯФ СО РАН);
- “Суперкомпьютерные технологии решения больших задач естествознания, математические модели, методы анализа и оптимизации сложных информационных систем” (ИВМиМГ СО РАН);
- “Развитие новых методов прогнозирования в системе Солнце — Земля” (ИСЗФ СО РАН);
- “Многокомпонентные исследования Вселенной в крупномасштабных экспериментах TAIGA и Baikal-GVD” (НИИПФ ИГУ);
- “Теоретические основы, методы и высокопроизводительные алгоритмы непрерывной и дискретной оптимизации для поддержки междисциплинарных научных исследований” (ИДСТУ СО РАН).

Приведем примеры решаемых задач (в рамках перечисленных научных проектов) с использованием оборудования Иркутского суперкомпьютерного центра.

ИрИХ СО РАН. Теоретическое изучение строения и свойств органических и элементоорганических соединений (в том числе силатранов, дисиленов, кремнийсодержащих диазенов и др.). Квантово-химическое моделирование стереоэлектронного строения органических молекул. Квантово-химические расчеты констант спин-спинового взаимодействия в спектрах ядерного магнитного резонанса. Разработка новых методов направленного органического и элементоорганического синтеза на базе ацетиленов и его производных (продуктов газо-, нефте- и углепереработки) для получения прекурсоров

лекарственных средств новых поколений, активных против социально значимых заболеваний (ВИЧ, коронавирусов).

ИГХ СО РАН. Моделирование структуры кварцевого стекла методами молекулярной динамики. Моделирование процессов плавления кристаллов кварца и кристобалита методом молекулярной динамики. Квантово-химическое моделирование колебательных спектров природных минералов (в частности, эльпидита и его производных). Исследование механизмов преобразования энергии в кристаллах щелочно-земельных галоидов, активированных редкоземельными ионами, с целью получения эффективного сцинтилляционного детектора для контроля за нераспространением радиоактивных материалов.

ИХКГ СО РАН. Квантово-химическое моделирование электронной структуры и магнитных свойств молекулярных магнетиков на основе органических ди- и трирадикалов. Квантово-химические расчеты термической стабильности высокоэнергетических соединений. Квантово-химическое исследование термоллиза богатых азотом соединений. Исследование механизмов термического разложения фуросана с помощью многоконфигурационных квантовохимических методов.

СФУ. Дизайн новых двумерных материалов для нано- и спинэлектроники, исследование их электронных свойств, исследование свойств уже существующих материалов и их композитов. Квантово-химическое моделирование наноразмерных структур для оптоэлектронных применений (в частности, фторированного фосфорена, диселенида палладия и др.). Квантово-химическое моделирование структуры, свойств и контактных взаимодействий для гетероструктур, сконструированных на основе графена и монослоев дителлурида ванадия.

ХФИЦ ДВО РАН (ИМ ДВО РАН). Квантово-механическое моделирование атомной и электронной структуры каталитически активных перовскитов.

ИАПУ ДВО РАН. Квантово-механическое моделирование свойств перспективных наноструктурированных материалов (термоэлектриков) на основе кремния и антимонида галлия.

ЛИН СО РАН. Исследование геномов, реконструкция эволюционной истории и филогенетических отношений эндемичных байкальских амфипод. Филогенетический и структурный анализ тубулинов диатомовых водорослей. Разработка алгоритма идентификации и делимитации биологических видов по геномным данным на примере видового состава микроводоросли *Choricystis* оз. Байкал. Анализ последовательностей ДНК вирусных, бактериальных сообществ, их таксономической идентификации, статистических характеристик, построения филогенетических деревьев. Изучение таксономического и генетического разнообразия микробных сообществ водной толщи оз. Байкал на основе различных маркерных генов. Исследование структуры микробных сообществ и роли микроорганизмов, участвующих в превращении углеводов.

ИЯФ СО РАН. Численные эксперименты в рамках проекта AWAKE по созданию нового типа ускорителя элементарных частиц на основе плазменного кильватерного ускорения с ведущим протонным пучком. Моделирование параметров электронного пучка, ускоренного в плазменной волне, созданной лазерным импульсом проектируемой установки XCELS. Моделирование разрушения плазменной волны электронным гало.

ИВМиМГ СО РАН. Разработка программного кода для моделирования генерации электромагнитного излучения при взаимодействии релятивистских электронных пучков и лазерных импульсов с плазмой. Разработка алгоритмов движения заряженных частиц в самосогласованных электромагнитных полях.

ИСЗФ СО РАН. Численное моделирование распространения магнитозвуковых волновых возмущений в плазме солнечной короны. Решение численной задачи экстраполяции магнитного поля по данным солнечных векторных магнитограмм. Моделирование трехмерной структуры активных областей Солнца в бессиловом приближении магнитного поля.

НИИПФ ИГУ. Разработка методов классификации широких атмосферных ливней. Численное моделирование работы установок Tunka-Grande, TAIGA-Muon, TAIGA-HiSCORE в составе гамма-обсерватории TAIGA.

ИДСТУ СО РАН. Разработка параллельных алгоритмов решения трудных комбинаторных задач. Разработка параллельных алгоритмов решения трудных экземпляров проблемы булевой выполнимости (SAT). Поиск эффективных декомпозиций трудных экземпляров SAT-задач. Анализ стойкости поточных шифров при помощи методов псевдобулевой оптимизации и параллельных алгоритмов решения SAT-задач. Разработка эффективных алгоритмов кластеризации для задач большой размерности на основе методов целочисленного программирования и математической оптимизации.

Заключение

Внедрение современных информационно-вычислительных и телекоммуникационных технологий в научные исследования на примере Интегрированной информационно-вычислительной сети ИРНОК показало свою эффективность и востребованность. Интеграция разнородных научных данных и возможность их обработки высокопроизводительными вычислительными средствами открывают новые возможности для получения новых прорывных результатов. По данным мониторинга пользователями ИРНОК и ИСКЦ ежегодно публикуются десятки научных работ с упоминанием об использовании оборудования центров, в том числе в ведущих международных научных журналах [10], что является показателем высокой востребованности вычислительных ресурсов центра.

Благодарности. Работа выполнена в рамках гранта № 075-15-2020-787 Министерства науки и высшего образования РФ на выполнение крупного научного проекта по приоритетным направлениям научно-технологического развития (проект “Фундаментальные основы, методы и технологии цифрового мониторинга и прогнозирования экологической обстановки Байкальской природной территории”).

Список литературы

- [1] Vassilyev S.N., Stupin G.B. Integrated information-computing network of Irkutsk scientific center for local and global transport of information. State of the Art and Perspectives: Abstract BBD Symposium Work-Shop on Data-Based and Information System. Japan; 1998: 4.

- [2] **Бычков И.В., Маджара Т.И., Новопашин А.П., Ружников Г.М., Опарин Г.А.** Анализ и перспективы развития инфраструктуры интегрированной информационно-вычислительной сети Иркутского научно-образовательного центра. Вестник НГУ. Сер. Информационные технологии. 2008; 6(1):25–36.
- [3] Иркутский суперкомпьютерный центр СО РАН. Адрес доступа: <http://hpc.icc.ru> (дата обращения: 16.04.2023).
- [4] Научно-технологическая инфраструктура Российской Федерации. Адрес доступа: <https://ckp-rf.ru> (дата обращения: 16.04.2023).
- [5] Суперкомпьютеры. ТОП-50. Адрес доступа: <http://top50.supercomputers.ru> (дата обращения: 16.04.2023).
- [6] Национальная исследовательская компьютерная сеть России. Адрес доступа: <https://niks.su> (дата обращения: 16.04.2023).
- [7] **Fedorov R.K., Shumilov A.S., Ruzhnikov G.M.** Geoportals cloud. CEUR Workshop Proceedings. 2017; (2033):305–308.
- [8] **Бычков И.В., Гладкочуб Д.П., Ружников Г.М.** Фундаментальные основы, методы и технологии цифрового мониторинга и прогнозирования экологической обстановки Байкальской природной территории. Новосибирск: Изд-во СО РАН; 2022: 345. DOI:10.53954/9785604788943.
- [9] **Feoktistov A., Kostromin R., Sidorov I.A., Gorsky S.A.** Development of distributed subject-oriented applications for cloud computing through the integration of conceptual and modular programming. Proceedings of the 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO-2018). Riejska: IEEE; 2018: 256–261. DOI:10.23919/MIPRO.2018.8400044.
- [10] ТОП-25 публикаций пользователей ЦКП ИСКЦ. Адрес доступа: <https://hpc.icc.ru/publications> (дата обращения: 16.04.2023).

Information and computing resources of IREC: infrastructure, data, applications

BYCHKOV IGOR V., MADZHARA TARAS I., NOVOPASHIN ALEKSEI P.,
FEREFEROV EVGENIY S.*, FEOKTISTOV ALEKSANDR G., FEDOROV ROMAN K.

V.M. Matrosov Institute of System Dynamics and Control Theory SB RAS, 664033, Irkutsk, Russia

*Corresponding author: Fereferov Evgeniy S., e-mail: fereferov@icc.ru

Received April 21, 2023, revised April 24, 2023, accepted April 28, 2023.

Abstract

The Integrated information and computing network of the Irkutsk scientific and educational complex (IICN IREC) is the basic technological platform for most interdisciplinary research conducted by scientific organizations of the Irkutsk region. The article presents the stages of development of the information, computing and telecommunications infrastructure of the IICN IREC in the period from 2008 to 2022. The main directions of modernization of IREC were: the development of a communication network, the deployment of a virtualization platform, data storage

system and GPU cluster. The Irkutsk supercomputer center (ISCC) of the SB RAS, which interacts with IREC, has also been developed. The most significant stage in the development of the ISCC was the implementation of the project to create a new generation computing cluster, named after the first organizing director of the institute, Academician V.M. Matrosov. The effectiveness of the created information and computing infrastructure is shown on the example of the implementation of two major integration projects. These are the projects “Fundamental research and breakthrough technologies as the basis for the advanced development of the Baikal region and its interregional relations” (2016–2019) and “Fundamental foundations, methods and technologies for digital monitoring and forecasting of the environmental situation in the Baikal natural territory (BNT)” (2020–2022). More than a dozen institutions have taken part in these projects. To support interdisciplinary research, tools that ensure effective interaction of researchers with the IREC information and computing infrastructure have been developed at the ISDCT SB RAS. Among them are Geoportal, Digital Platform for Monitoring of BNT and Orlando Tools. As part of the projects, a large array of various formats data on the state of the region was collected and computationally complex problems were solved. The integration of heterogeneous scientific data and the possibility of processing them with high-performance computing tools open up new opportunities for obtaining new breakthrough results.

Keywords: information-computing and telecommunication resources and platforms, high-performance computing systems, data storage systems, virtualization systems, geoportal, tools.

Citation: Bychkov I.V., Madzhara T.I., Novopashin A.P., Fereferov E.S., Feoktistov A.G., Fedorov R.K. Information and computing resources of IREC: infrastructure, data, applications. *Computational Technologies*. 2023; 28(3):117–135. DOI:10.25743/ICT.2023.28.3.008. (In Russ.)

Acknowledgements. The work was supported by the Ministry of Science and Higher Education of the Russian Federation, the grant No. 075-15-2020-787 for implementation of Major scientific projects on priority areas of scientific and technological development (the project “Fundamentals, methods and technologies for digital monitoring and forecasting of the environmental situation on the Baikal natural territory”).

References

1. **Vassilyev S.N., Stupin G.B.** Integrated information-computing network of Irkutsk scientific center for local and global transport of information. State of the Art and Perspectives: Abstract BBD Symposium Work-Shop on Data-Based and Information System. Japan; 1998: 4.
2. **Bychkov I.V., Madzhara T.I., Novopashin A.P., Oparin G.A., Ruzhnikov G.M.** Analysis and stages of development and future trends of information-computing telecommunication infrastructure of Irkutsk scientific-educational complex. *Bulletin of NSU. Ser. Information Technology*. 2008; 6(1):25–36. (In Russ.)
3. Irkutskiy superkomp'yuternyy tsentr SO RAN [Irkutsk supercomputer center SB RAS]. Available at: <http://hpc.icc.ru/en> (accessed at April 16, 2023).
4. Nauchno-tekhnologicheskaya infrastruktura Rossiyskoy Federatsii [Scientific and technological infrastructure of the Russian Federation]. Available at: <https://en.ckr-rf.ru> (accessed at April 16, 2023).
5. Superkomp'yutery. TOP-50 [Supercomputers. TOP-50]. Available at: <http://top50.supercomputers.ru> (accessed at April 16, 2023). (In Russ.)
6. Natsional'naya issledovatel'skaya komp'yuternaya set' Rossii [National research computer network of Russia]. Available at: <https://niks.su/en> (accessed at April 16, 2023).
7. **Fedorov R.K., Shumilov A.S., Ruzhnikov G.M.** Geoportal cloud. *CEUR Workshop Proceedings*. 2017; (2033):305–308.
8. **Bychkov I.V., Gladkochub D.P., Ruzhnikov G.M.** Fundamental'nye osnovy, metody i tekhnologii tsifrovogo monitoringa i prognozirovaniya ekologicheskoy obstanovki Baykal'skoy prirodnoy territorii [Fundamentals, methods and technologies of digital monitoring and forecasting of the ecological situation of the Baikal natural territory]. Novosibirsk: Izd-vo SO RAN; 2022: 345. DOI:10.53954/9785604788943. (In Russ.)

9. **Feoktistov A., Kostromin R., Sidorov I.A., Gorsky S.A.** Development of distributed subject-oriented applications for cloud computing through the integration of conceptual and modular programming. Proceedings of the 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO-2018). Riejska: IEEE; 2018: 256–261. DOI:10.23919/MIPRO.2018.8400044.
10. TOP-25 publikatsiy pol'zovateley TsKP ISKTs [TOP 25 publications of users of the CCP SSCC]. Available at: <https://hpc.icc.ru/publications> (accessed at April 16, 2023). (In Russ.)